

Opening Access to Visual Exploration of Audiovisual Digital Biomarkers: an OpenDBM Analytics Tool

Carla Floricel*

University of Illinois Chicago

Jacob Epifano, Stephanie Caamano, Sarah Kark, Rich Christie, Aaron Masino, Andre D Paredes†

AiCure

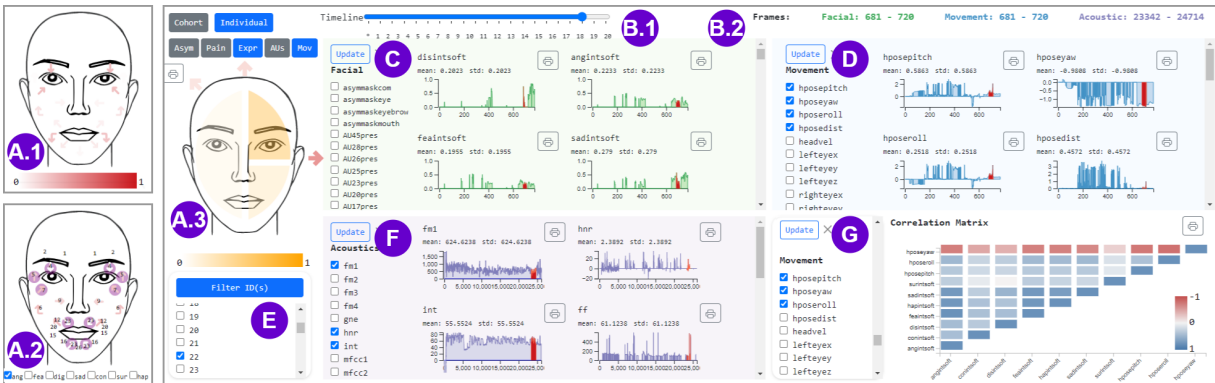


Figure 1: Individual Panel. A) Head Sketch View which uses custom colored masks to display facial asymmetry, pain expressivity, overall expressivity (A.3), action unit (AU) intensity (A.1), and head movement (A.3) biomarkers. In A.2, the AUs' numbers are displayed and the AUs involved in anger expressivity (bottom selection) are marked with purple highlights. B.1) Timeline used to split the data into 20 time intervals. Upon change, the timeline will update the mean values displayed in A and highlight the selected interval in views C,D, and F, while the corresponding frame intervals will be shown in B.2. C, D, F) Facial, Movement, and Acoustics Views that display temporal data distributions for selected biomarker variables. E) ID Query Subpanel where one video can be chosen to display its DBM data. G) Correlation View that shows a pairwise Pearson correlation coefficients for a set of variables.

ABSTRACT

Digital biomarkers (DBMs) are a growing field and increasingly tested in the therapeutic areas of psychiatric and neurodegenerative disorders. Meanwhile, isolated silos of knowledge of audiovisual DBMs use in industry, academia, and clinics hinder their widespread adoption in clinical research. How can we help these non-technical domain experts to explore audiovisual digital biomarkers? The use of open source software in biomedical research to extract patient behavior changes is growing and inspiring a shift toward accessibility to address this problem. OpenDBM integrates several popular audio and visual open source behavior extraction toolkits. We present a visual analysis tool as an extension of the growing open source software, OpenDBM, to promote the adoption of audiovisual DBMs in basic and applied research. Our tool illustrates patterns in behavioral data while supporting interactive visual analysis of any subset of derived or raw DBM variables extracted through OpenDBM.

1 INTRODUCTION

The global market value of digital biomarkers (DBMs) is forecasted to exceed \$7 billion by 2026 [1]. DBMs are objective, quantifiable, physiological, and behavioral data collected and measured using digital devices. Like traditional biomarkers, DBMs have clinical value, such as predicting and diagnosing disease. However, DBMs introduce additional benefits that exceed traditional biomarkers'

constraints, such as capturing longitudinal and continuous measurements that generate large, rich, and complex datasets [2]. Providing clinical researchers with practical tools to derive and interpret DBMs increases their ability to assess changes in health status relevant to healthcare applications [48]. To support the rising demand for DBM adoption by clinical researchers, more practical tools are required to better inform non-technical biomedical researchers on how to use and identify DBMs [16, 26].

In particular, DBMs' increasing role in the therapeutic areas of psychiatric and neurodegenerative disorders provides new-found interest for clinical researchers to explore measurable audiovisual DBM changes to understand better how a patient feels [8, 12, 30, 31, 38, 40, 41, 43, 49]. Growing open source software projects, such as OpenDBM, are lowering the barrier for non-technical clinical researchers to apply quantitative models, including machine learning models, to extract audiovisual features in human speech, voice acoustics, head movement, and facial expressions [21, 22]. However, despite accessible open source tools to extract audiovisual features, clinical researchers are burdened with interpreting large and complex quantitative datasets [26].

Given the novelty of DBMs and their still growing taxonomy and use [14], there is interest among behavioral and biomedical researchers in finding practical tools that can facilitate exploratory analysis for data-informed hypothesis generation. This work aims to improve researchers' understanding of the breadth and scope of the hundreds of audiovisual DBMs available for investigatory adoption. We propose a visual analytics interface for the OpenDBM software¹. Our proposed interface reveals patterns and outliers in facial, head movement, acoustics, and speech DBMs extracted from videos. To our knowledge, this work presents the first audiovisual DBM

*e-mail: cflori3@uic.edu

†e-mail: andre.paredes@aicure.com

¹<https://aicure.com/opendbm/>

interactive visualization tool extracted from and made available through open source software.

2 RELATED WORK AND BACKGROUND

Open Source Audio and Visual Feature Extraction Tools. The system to measure the nature and intensity of vocal and facial expressions is advancing from manual raters to computerized toolkits [23, 42]. These audio and visual toolkits are made possible by leveraging advancements in machine learning and artificial intelligence techniques, such as natural language processing and computer vision [10, 15, 30, 32]. A growing number of open source software projects are starting to make vocal and facial feature extraction toolkits freely available online. For vocal feature extractions, Parselmouth [29], Natural Language Toolkit [34], LexicalRichness [44], and VaderSentiment [56] have been cited for calculating a whole host of speech and acoustic DBM variables. For facial feature extractions, OpenFace is a commonly cited behavior analysis toolkit for detecting and measuring facial landmarks, facial action units, head pose estimation, and gaze estimation [3, 4, 20]. Collectively, these software toolkits provide a rich and diverse suite of extracted features for a more comprehensive analysis of emotional communication behavior over time. However, none of these projects provide visualization tools that can aid data interpretation. The project presented in this paper uses the OpenDBM solution, integrating all of the previously mentioned vocal and facial toolkits for generating a visualization of these extracted audiovisual features collectively.

Visualization in Healthcare. Visualization techniques in healthcare informatics often target cohort data exploration, covering applications in disease evolution from electronic medical records [28, 50, 51, 54], heterogeneous longitudinal clinical data [19, 27], or volumetric patient data [25, 52]. However, prior work in healthcare visual analytics mostly focus on chronic conditions such as cancer [9], stroke [33], diabetes [17], or on infectious disease control due to the COVID-19 pandemic [5, 45], and less on psychiatric and neurodegenerative disorders. Our work incorporates a new approach to promote individual patient data exploration, while incorporating past working approaches for cohort data exploration. There is work in visual analytics for facial activity and head movement and separately, for voice acoustics and speech measurements [13, 36, 46, 47, 55], however, some of it doesn't use video data, and none accounts for all four measurement categories together. We aim to provide efficient tools for psychiatric and neurodegenerative health studies using heterogeneous, audiovisual, behavioral biomarker measurements extracted during clinical assessments.

3 DESIGN PROCESS AND REQUIREMENTS

The design process followed an Activity-Centered-Design approach [35]. Our team held remote meetings for nine weeks with five research groups in DBM therapeutic areas, collectively representing academia, clinics, and industry. While most collaborators were principal investigators with faculty positions, conducting behavioral or biomedical research, all of them were familiar with the OpenDBM software. Throughout this process, the team iteratively gained insight into user approaches to explore mappings between DBMs and conditions and disorders of interest (e.g., major depression and schizophrenia), gathered functional specifications for a DBM interface, and prototyped and evaluated the interface. Due to the large variety in patient behavior for these disorders, we gathered many specific requirements. However, we focused on the following subset of high-level requirements to serve all our collaborators and the open source community:

R1: Provide flexibility in showing details about any subset of DBM variables available through the OpenDBM pipeline. For instance, for early detection of Parkinson's disease, head movement measurements are of greater importance than other DBM, such as voice acoustics. Adaptability to different workflows is an essential

factor in open source. Additionally, analyzing hundreds of variables can be very challenging, and sometimes researchers don't know where to start their analyses. Thus, having the means and the freedom to choose what to explore visually is very important.

R2: Support interactive visualizations for both raw and derived data. Visualizing derived, mean variables is important for getting an effective overview of the cohort data and context for individual patients, while visualizing raw, temporal variables supports in-depth analysis for individual patients. This is critical for checking the quality of the data. For example, researchers might want to exclude from their analyses videos where the audio or the patient's face was not captured.

R3: Emphasize trends and outliers in DBM data. As an example, patients should show negative emotions when talking about unpleasant or uncomfortable subjects. Domain experts should easily observe patterns between patients, which is helpful for further studies. Furthermore, highlighting correlations between biomarkers is fundamental for better understanding these conditions.

4 VISUALIZATION DESIGN

The visual system is open source and can be operated through the OpenDBM Github project² from the visualization interface folder. It is not part of the DBM extraction pipeline, but serves as a complementary application that visualizes the output of the DBM extraction. The interface has two interactive panels: the Cohort Panel and the Individual Panel. These panels are composed of multiple coordinated views that support brushing and linking operations.

4.1 Data Description

Vocal and facial expressions convey emotion and communication behavior and are one of the most researched topics in psychology and related disciplines; as a result, audiovisual DBMs extend from these basic and applied science measurement tools [23]. When a video is processed through OpenDBM, the several vocal and facial feature extraction toolkits combine to present hundreds of unique variable categories relevant to four different audiovisual DBM domains: speech, acoustics, facial expression, and head movement. Each audiovisual DBM domain provides two sets of quantitative variables: raw, captured as a frame-by-frame time sequence measurement, and derived, capturing summary statistics on the total collection of frames. These raw and derived variables provide a wide range of objective behavioral cues, such as transcription and lexical richness for speech, jitter and shimmer for acoustics, eye blink and facial tremor for head movement, and facial action units and facial asymmetry for facial expressions. The proposed interface uses these raw and derived variables to display relevant details and statistics about video cohorts and individual videos using two panels: the Cohort and the Individual Panels. The official documentation² provides the full list of DBM variables extracted by OpenDBM.

4.2 Cohort Panel

The Cohort Panel (Fig. 2) has three main views and functions: provide a cohort overview based on a selected set of variables, observe variable distributions, and find correlations between variables.

Two query subpanels are available for variable and video ID selection, with the variable query subpanel (Fig. 2.A) having three alternative components for each of the three main views (Fig. 2.B,D,E). In the video ID query subpanel (Fig. 2.C), selected IDs are highlighted in the other views while unselected videos can be hidden from the rest. All views have accompanying print buttons to generate plot images that can be used in further studies.

PCA View. This view (Fig. 2.B.1) uses a scatterplot for a cohort overview by arranging videos in 2D based on a selected set of biomarker variables (R1, R2, R3). The axes correspond to the

²https://github.com/AiCure/open_dbm



Figure 2: Cohort Panel. A) Query Subpanel with three alternative components for biomarker variable selection for views B, D, E. B.1) PCA View that uses a scatterplot to display in 2D video data based on the variable selections in A. B.2) The PCA scatterplot is color-coded based on an extra attribute, namely, the task that was performed during each video. C) ID Query Subpanel where IDs can be selected to be highlighted in views B and D, and unselected IDs can be hidden. D) Distribution View showing cohort distributions for four selected variables. E) Correlation View displaying pairwise Pearson correlation coefficients for six selected biomarker variables.

first two components computed by a Principal Component Analysis (PCA) [53]. We employed factor analysis (i.e. PCA) to help researchers get a better sense of underlying structures in the high-dimensional video data while retaining patterns. The view shows trends and outliers, while brushing interactions highlight selected elements in the Distribution View and ID query subpanel.

Distribution View. This view (Fig. 2.D) displays distribution charts for a selected set of biomarker variables (R1, R2, R3). Split into two components, each distribution chart shows one variable distribution across the cohort. The left side uses a scatterplot for easier detection of individual videos, while the right side uses a density plot for a concise cohort overview. Hovering on the scatterplot will highlight corresponding elements in the PCA Scatterplot and ID query subpanel and tooltips will display video IDs and variable values.

Many times, domain experts study biomarker data while trying to find patterns between different cohorts. While gathering specifications for the system, a frequent request was to distinguish between subcohorts with different treatment plans, health conditions, age range, and so on. Thus, if such data is available, the system will color videos (Fig. 2.B.2) based on that list of extra attributes in both the PCA and Distribution Views (R3).

Correlation View. This view (Fig. 2.E) contains a correlation matrix to emphasize interrelationships between a selected set of biomarker variables (R3). The matrix shows the coefficients computed using Pearson’s Correlation [7] method. Accompanying tooltips display the coefficient values for each pair or variables.

4.3 Individual Panel

This panel (Fig. 1) has five coordinated views, showing temporal values for facial, movement, and acoustic variables, as well as a few derived variables, and correlations between raw variables.

The panel features an ID query subpanel (Fig. 1.E), where one video can be chosen to display its DBM data. An interactive timeline (Fig. 1.B.1) splits the raw data into 20 time frames and highlights the corresponding time span in the other views upon change. Similar to the Cohort Panel, each view features print buttons.

Head Sketch View. This head sketch (Fig. 1.A.1,2,3) supports derived value abstractions for facial activity and head pose biomarker data (R2, R3). Four alternative masks that use custom heatmaps can be applied to support facial biomarker visualization. The Asym mask uses colored glyphs to highlight facial features’ asymmetry values. The Pain mask uses a colored glyph to highlight the pain expressivity values. The Expr mask uses colored glyphs to highlight the overall, upper, and lower face expressivity values (Fig. 1.A.3). The AUs mask (Fig. 1.A.1) highlights action units’ intensity values using arrows pointing to the direction of the corresponding facial features’ movements [18]. As emotions are expressed through a combination of action units [42], the AUs mask has a complementary layer (Fig. 1.A.2) that indicates the set of AUs that get activated for each of the seven available emotions. Upon hovering, the action units’ numbers are visible on the facial sketch. Besides the facial activity masks, a head pose (Mov) mask (Fig. 1.A.3) is also available and it indicates yaw, roll, and pitch head movements using colored pairs of arrows for each action. By default, the head sketch masks show derived values. However, this view will show mean values for the selected time frame when the timeline is updated.

Facial, Movement, and Acoustics Views. The Individual Panel features one view for temporal data for each of the three biomarker categories, namely, facial activity (Fig. 1.C), head movement (Fig. 1.D), and voice acoustics (Fig. 1.F) (R1, R2, R3). Each view uses histograms to display temporal distributions for selected variables in the accompanying query subpanels, with the X axis representing time.

The placement of these views facilitate the discovery of patterns among biomarkers. When the timeline is updated (Fig. 1.B1), the corresponding time span is highlighted with red in all histograms while the frame spans are visible for each biomarker category (Fig. 1.B.2). **Correlation View.** This view (Fig. 1.G) uses a correlation matrix to support the same functionality as the Correlation View in the Cohort Panel (R3). However, here, the Pearson Correlation is computed using longitudinal, individual data.

DBMs, as an interdisciplinary tool in AI and medicine, have a far-reaching potential in basic and applied sciences that will continue to drive qualitative domain experts' interest in quantitative DBMs. Therefore, when we started this project, we realized the importance of making the interface accessible to a broad audience, including qualitatively trained domain experts. As a result, we used conventional visual encodings suitable for various visual literacy environments, such as scatterplots, histograms, density plots, and matrices. Additionally, we experimented with custom visualizations, such as the facial and head pose masks, taking advantage of the power of visual mapping techniques and trying to make better sense of the behavioral measurements from facial activity and head movement. We chose to first experiment with these two DBM categories because they were of interest in many use cases during our design requirements interviews.

5 EVALUATION

We evaluated our system using two case studies that involved three domain experts and a video simulated actor dataset. Using simulated actor datasets is a well-accepted practice for capturing the prototypical representation of the multiple and complex emotional states of psychiatric and neurodegenerative disorders [6, 11, 24, 39]. A video simulated actor dataset was generated and used for our case studies. It included 95 videos, all under two minutes, of one adult male actor, instructed to perform five categories of tasks while simulating major depression disorder for some videos. The five tasks included: describe a picture, describe a memory, describe your day, read a passage, and reproduce a vowel sound or a basic facial emotional expression. The evaluation was conducted via video conferencing using screen sharing and the think aloud method.

5.1 Case Study I: Cohort Analysis

For this case study, the domain experts were first interested in discarding videos with no relevant information from the cohort (i.e., videos where the actor don't perform any tasks). The investigation started with the PCA View (Fig. 2.A, B.1), where audio intensity, fundamental frequency, and glottal to noise excitation ratio, all derived audio biomarker variables, were chosen as parameters to display the videos in 2D. Most videos were grouped towards the lower-central part of the view. Next, the previous set of parameters were used for the Distribution View (Fig. 2.D), and after brushing the upper part of the PCA scatterplot, the distribution charts revealed that the selected videos had the lowest audio intensity values from the cohort. When brushing the left or right outliers, the distribution charts showed that the selections belonged to opposite cohort extremities for fundamental frequency and glottal to noise excitation ratio, which aligned with previous research [37]. When using the option to color the videos by the task category (Fig. 2.B.2), the scatterplot showed that the upper and left sides of the view was composed of videos with no audio or videos where the actor reproduced facial expressions or vowels. Lastly, the domain experts wanted to check the correlations between these parameters and speech DBM variables, such as the number of pronouns, verbs, or adjectives used per task (Fig. 2.E). The Correlation View revealed strong positive correlations between speech DBM variables and the audio intensity. This case study showed the interface's ability to show relevant trends and outliers in subsets of derived DBM variables of interest (R1, R2, R3) and helped the evaluators to detect videos with no acoustic DBM data.

5.2 Case Study II: Individual Analysis

This case study started with the exploration of raw data for a video where the actor performs a picture description task (Fig. 1.E). The facial emotion expressions were chosen as parameters for the Facial View (Fig. 1.C). The histograms showed high spikes for most negative emotions, such as distress, anger, fear, and sadness, implying that the task entailed describing a negative impact image. Hence, the domain experts were interested in observing facial action unit changes over time, since they are connected to facial expressions, so they used the Head Sketch View (Fig. 1.A.1) and the timeline for this task (Fig. 1.B.1). After applying the AUs facial mask for action units and the Mov mask for head pose changes (Fig. 1.A.3), they investigated different time frames using the timeline, and observed that, indeed, at most times, the action units for negative emotions were active (Fig. 1.A.2). Next, the evaluators checked correlations between head poses and emotions using the Correlation View (Fig. 1.G), and observed that, surprisingly, roll head movements were negatively correlated with most emotions, while the other head poses didn't show any particularly strong correlations to emotions. Curious, the evaluators watched the actual video, which showed the actor describing a picture of a building on fire. This case study showcases the interface's ability to show patterns in selected raw DBM variables for one video (R1, R2, R3) and helped the evaluators to detect a video where negative emotional impact was present.

6 DISCUSSION AND CONCLUSION

We received positive feedback for the current version of our interface in particular because, previously, domain experts were limited to manual and laborious means of inspecting video data: *"Very cool, so much better to use for the analysis we did last year, huge time saver,"* When asked what they found most useful for their own research studies, most people pointed out to the interface's ability to delineate subcohorts using different colors (Fig. 2.B.2) *"very excited about the color coding,"* as well as the histograms (Fig. 1.C,D,F) *"I had someone looking away from the camera, this is actually picking up their data."* However, one evaluator indicated that it would be useful to *"input our own data [in the interface] to work with the [bio]markers,"* as it could speed up symptom research for different disorders. We plan to address this matter in our future work.

Overall this is a well-received start toward introducing visualization approaches to domain experts in DBM therapeutic fields, but it is still a work in progress. There are several limitations to our current solution, which we plan to improve. For example, the Cohort Panel suffers from scalability issues (i.e., the scatterplots) when it comes to large cohorts of hundreds of videos. Furthermore, we are currently researching missing DBM normative ranges, which better inform domain experts on abnormal patterns and help generate more accurate hypotheses about an individual's health status.

This visualization approach begins to address the need for transparency behind data quality control and quality assurance from these integrated open source toolkits. The open-source space is a natural fit to drive domain expert users to test and ratify best practices. Future work can build upon our visualization approach by further improving visualizations of raw data for checking data quality from new and unvalidated toolkits.

We conclude that our work offers the first visualization analysis tool for domain experts interested in exploring and using audiovisual DBMs, meeting the following five criteria: (1) hundreds of audio and visual DBM variables (2) cohort visualization panel to identify patterns in groups of audiovisual recordings (3) individual panel to identify single patient patterns in vocal and facial expression over time (4) open-source software (5) integration to a growing number of open source audio and visual feature extraction toolkits, via OpenDBM.

REFERENCES

- [1] Digital biomarkers: Global markets (2022). <https://www.bccresearch.com/market-research/biotechnology/digital-biomarkers-market.html>. Accessed: 2022-08-1.
- [2] L. M. Babrak, J. Menetski, M. Rebhan, G. Nisato, M. Zinggeler, N. Brasier, K. Baerenfaller, T. Brenzikofer, L. Baltzer, C. Vogler, et al. Traditional and digital biomarkers: two worlds apart? *Digital biomarkers*, 3(2):92–102, 2019.
- [3] T. Baltrušaitis, P. Robinson, and L.-P. Morency. Openface: an open source facial behavior analysis toolkit. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–10. IEEE, 2016.
- [4] T. Baltrušaitis, A. Zadeh, Y. C. Lim, and L.-P. Morency. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*, pp. 59–66. IEEE, 2018.
- [5] T. Baumgartl, M. Petzold, M. Wunderlich, M. Hohn, et al. In Search of Patient Zero: Visual Analytics of Pathogen Transmission Pathways in Hospitals. *IEEE Trans. Vis. Comp. Graph.*, 27(2):711–721, 2021.
- [6] M. S. Benda and K. S. Scherf. The complex emotion expression database: A validated stimulus set of trained actors. *PLoS one*, 15(2):e0228248, 2020.
- [7] J. Benesty, J. Chen, Y. Huang, and I. Cohen. Pearson correlation coefficient. In *Noise reduction in speech processing*, pp. 1–4. Springer, 2009.
- [8] J. Benoit, H. Onyeaka, M. Keshavan, and J. Torous. Systematic review of digital phenotyping and machine learning in psychosis spectrum illnesses. *Harvard Review of Psychiatry*, 28(5):296–304, 2020.
- [9] J. Bernard, D. Sessler, T. May, T. Schlomm, D. Pehrke, and J. Kohlhammer. A visual-interactive system for prostate cancer stratifications. In *VIS 2014 Workshop on Visualization of Electronic Health Records*, 2014.
- [10] S. Bhadra and C. J. Kumar. An insight into diagnosis of depression using machine learning techniques: a systematic review. *Current Medical Research and Opinion*, 38(5):749–771, 2022.
- [11] H. Cao, D. G. Cooper, M. K. Keutmann, R. C. Gur, A. Nenkova, and R. Verma. Crema-d: Crowd-sourced emotional multimodal actors dataset. *IEEE transactions on affective computing*, 5(4):377–390, 2014.
- [12] A. S. Chandrabhatla, I. J. Pomeraniec, and A. Ksendzovsky. Co-evolution of machine learning and digital technologies to improve monitoring of parkinson’s disease motor symptoms. *NPJ digital medicine*, 5(1):1–18, 2022.
- [13] J. H. Cheong, T. Xie, S. Byrne, and L. J. Chang. Py-feat: Python facial expression analysis toolbox. *arXiv preprint arXiv:2104.03509*, 2021.
- [14] S. Coghlán and S. D’Alfonso. Digital phenotyping: an epistemic and methodological analysis. *Philosophy & Technology*, 34(4):1905–1928, 2021.
- [15] C. M. Corcoran and G. A. Cecchi. Using language processing and speech analysis for the identification of psychosis and other disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 5(8):770–779, 2020.
- [16] V. De Angel, S. Lewis, K. White, C. Oetzmann, D. Leightley, E. Oprea, G. Lavelle, F. Matcham, A. Pace, D. C. Mohr, et al. Digital health tools for the passive monitoring of depression: a systematic review of methods. *NPJ digital medicine*, 5(1):1–14, 2022.
- [17] S. Di Bartolomeo, Y. Zhang, F. Sheng, and C. Dunne. Sequence b raiding: Visual overviews of temporal event sequences and attributes. *IEEE Trans. Vis. Comp. Graph.*, 27(2):1353–1363, 2020.
- [18] P. Ekman and W. V. Friesen. Facial action coding system. *Environmental Psychology & Nonverbal Behavior*, 1978.
- [19] C. Floricel, N. Nipu, M. Biggs, A. Wentzel, G. Canahuate, L. Van Dijk, A. Mohamed, C. D. Fuller, and G. E. Marai. Thalix: Human-machine analysis of longitudinal symptoms in cancer therapy. *IEEE Trans. Vis. Comp. Graph.*, 28(1):151–161, 2021.
- [20] A. Fydanaki and Z. Geradts. Evaluating openface: an open-source automatic facial comparison algorithm for forensics. *Forensic sciences research*, 3(3):202–209, 2018.
- [21] I. Galatzer-Levy, A. Abbas, A. Ries, S. Homan, L. Sels, V. Koesmahargyo, V. Yadav, M. Colla, H. Scheerer, S. Vetter, et al. Validation of visual and auditory digital markers of suicidality in acutely suicidal psychiatric inpatients: Proof-of-concept study. *Journal of medical Internet research*, 23(6):e25199, 2021.
- [22] I. R. Galatzer-Levy, A. Abbas, V. Koesmahargyo, V. Yadav, M. M. Perez-Rodriguez, P. Rosenfield, O. Patil, M. F. Dockendorf, M. Moyer, L. A. Shipley, et al. Facial and vocal markers of schizophrenia measured using remote smartphone assessments. *medRxiv*, 2020.
- [23] M. Geiger and O. Wilhelm. Computerized facial emotion expression recognition. In *Digital Phenotyping and Mobile Sensing*, pp. 43–56. Springer, 2023.
- [24] N. R. Giuliani, J. C. Flournoy, E. J. Ivie, A. Von Hippel, and J. H. Pfeifer. Presentation and validation of the duckees child and adolescent dynamic facial expressions stimulus set. *International Journal of Methods in Psychiatric Research*, 26(1):e1553, 2017.
- [25] N. Grossmann, O. Casares-Magaz, L. P. Muren, V. Moiseenko, J. P. Einck, M. E. Gröller, and R. G. Raidou. Pelvis runner: Visualizing pelvic organ variability in a cohort of radiotherapy patients. In *VCBM*, pp. 69–78, 2019.
- [26] R. H. Birk and G. Samuel. Can digital data diagnose mental health problems? a sociological exploration of ‘digital phenotyping’. *Sociology of Health & Illness*, 42(8):1873–1887, 2020.
- [27] T. A. Harbig, S. Nusrat, T. Mazor, Q. Wang, A. Thomson, H. Bitter, E. Cerami, and N. Gehlenborg. Oncothreads: visualization of large-scale longitudinal cancer molecular data. *Bioinformatics*, 37(Supplement.1):i59–i66, 2021.
- [28] C. W. Huang, R. Lu, U. Iqbal, S. H. Lin, et al. A richly interactive exploratory data analysis and visualization tool using electronic medical records. *BMC Med. Infor. and Dec. Making*, 15:92, 2015.
- [29] Y. Jadoul, B. Thompson, and B. De Boer. Introducing parselmouth: A python interface to praat. *Journal of Phonetics*, 71:1–15, 2018.
- [30] Z. Jiang, M. Luskus, S. Seyedi, E. L. Griner, A. B. Rad, G. D. Clifford, M. Boazak, and R. O. Cotes. Utilizing computer vision for facial behavior analysis in schizophrenia studies: A systematic review. *PLoS one*, 17(4):e0266828, 2022.
- [31] A. Kumar, T. Jaquenoud, J. H. Becker, D. Cho, M. R. Mindt, A. Federman, and G. Pandey. Can you hear me now? clinical applications of audio recordings. *medRxiv*, 2022.
- [32] R. Li, X. Wang, K. Lawler, S. Garg, Q. Bai, and J. Alty. Applications of artificial intelligence to aid detection of dementia: a scoping review on current capabilities and future directions. *Journal of Biomedical Informatics*, p. 104030, 2022.
- [33] M. H. Loorak, C. Perin, N. Kamal, M. Hill, et al. Timespan: Using visualization to explore temporal multi-dimensional data of stroke patients. *IEEE Trans. Vis. Comp. Graph.*, 22(1):409–418, 2015.
- [34] E. Loper and S. Bird. Nltk: The natural language toolkit. *arXiv preprint cs/0205028*, 2002.
- [35] G. E. Marai. Activity-centered domain characterization for problem-driven scientific visualization. *IEEE Trans. Vis. Comp. Graph.*, 24(1):913–922, 2017.
- [36] D. McDuff, R. El Kaliouby, K. Kassam, and R. Picard. Acume: A new visualization tool for understanding facial expression and gesture data. In *2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, pp. 591–596. IEEE, 2011.
- [37] D. Michaelis, T. Gramss, and H. W. Strube. Glottal-to-noise excitation ratio—a new measure for describing pathological voices. *Acta Acustica united with Acustica*, 83(4):700–706, 1997.
- [38] M. L. Miller, I. M. Raugh, G. P. Strauss, and P. D. Harvey. Remote digital phenotyping in serious mental illness: Focus on negative symptoms, mood symptoms, and self-awareness. *Biomarkers in Neuropsychiatry*, p. 100047, 2022.
- [39] H. O’Reilly, D. Pigat, S. Fridenson, S. Berggren, S. Tal, O. Golan, S. Bölte, S. Baron-Cohen, and D. Lundqvist. The eu-emotion stimulus set: a validation study. *Behavior research methods*, 48(2):567–576, 2016.
- [40] A. Piau, K. Wild, N. Mattek, J. Kaye, et al. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild alzheimer disease and implications for clinical care: systematic review. *Journal of medical Internet research*, 21(8):e12785, 2019.

- [41] M. L. B. Pulido, J. B. A. Hernández, M. Á. F. Ballester, C. M. T. González, J. Mekyska, and Z. Smékal. Alzheimer's disease and automatic speech analysis: a review. *Expert systems with applications*, 150:113213, 2020.
- [42] E. L. Rosenberg and P. Ekman. *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, 2020.
- [43] A. Ševčík and M. Rusko. A systematic review of alzheimer's disease detection based on speech and natural language processing. In *2022 32nd International Conference Radioelektronika (RADIOELEKTRONIKA)*, pp. 01–05. IEEE, 2022.
- [44] L. Shen. LexicalRichness: A small module to compute textual lexical richness, 2022. doi: 10.5281/zenodo.6607007
- [45] M. Sondag, C. Turkay, K. Xu, L. Matthews, S. Mohr, and D. Archambault. Visual analytics of contact tracing policy simulations during an emergency response. In *Computer Graphics Forum*, vol. 41, pp. 29–41. Wiley Online Library, 2022.
- [46] A. Tifentale and L. Manovich. Selfiecity: Exploring photography and self-fashioning in social media. In *Postdigital aesthetics*, pp. 109–122. Springer, 2015.
- [47] S. Ueng, C. Luo, T. Tsai, and H. Chang. Voice quality assessment and visualization. In *2012 Sixth International Conference on Complex, Intelligent, and Software Intensive Systems*, pp. 618–623. IEEE, 2012.
- [48] S. Vasudevan, A. Saha, M. E. Tarver, and B. Patel. Digital biomarkers: Convergence of digital health technologies and biomarkers. *NPJ digital medicine*, 5(1):1–3, 2022.
- [49] R. Voleti, J. M. Liss, and V. Berisha. A review of automated speech and language features for assessment of cognitive and thought disorders. *IEEE journal of selected topics in signal processing*, 14(2):282–298, 2019.
- [50] T. D. Wang, C. Plaisant, A. J. Quinn, et al. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In *Proc. SIGCHI Conf. Human Factors in Computing Systems*, p. 457–466, 2008.
- [51] T. D. Wang, C. Plaisant, B. Shneiderman, N. Spring, et al. Temporal summaries: Supporting temporal categorical searching, aggregation and comparison. *IEEE Trans. Vis. Comp. Graph.*, 15(6):1049–1056, 2009.
- [52] A. Wentzel, P. Hanula, T. Luciani, B. Elgohari, H. Elhalawani, G. Canahuate, D. Vock, C. D. Fuller, and G. E. Marai. Cohort-based t-ssim visual computing for radiation therapy prediction and exploration. *IEEE Trans. Vis. Comp. Graph.*, 26(1):949–959, 2019.
- [53] S. Wold, K. Esbensen, and P. Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52, 1987.
- [54] K. Wongsuphasawat and D. Gotz. Outflow: Visualizing patient flow by symptoms and outcome. In *IEEE VisWeek Workshop Vis. Anal. in Health.*, p. 25–28. American Medical Informatics Association, 2011.
- [55] Y. Yamashita, M. Onodera, K. Shimoda, and Y. Tobe. Visualizing health with emotion polarity history using voice. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, pp. 1210–1213, 2019.
- [56] C. K.-C. Yu. Automated analysis of dream sentiment—the royal road to dream dynamics? *Dreaming*, 2022.