Envisioning the Design Space of AI-Powered Personal Health Data Interaction

Shriti Raj

School of Information, University of Michigan, shritir@umich.edu

Mark W. Newman

School of Information, University of Michigan, mwnewman@umich.edu

Making personal health data actionable is a nuanced multi-step process wherein system and user agency can be harnessed in varied ways using a combination of computational and interaction techniques. However, personal informatics systems pay inadequate attention to how data-driven reflection and decision-making unfold. Consequently, a systematic understanding of designing for the multi-step process of making data actionable is lacking. In this position paper, we describe a framework called *episode-driven data interaction* and use it as a lens through which we can better envision the design space of AI-assisted data interaction to help make multidimensional personal health data actionable. We briefly describe the framework; and the requirements and challenges it presents for supporting AI-powered personal data interaction.

Additional Keywords and Phrases: personal health informatics, health data interaction, artificial intelligence

ACM Reference Format:

Shriti Raj, and Mark Newman. 2022. Envisioning the Design Space of AI-Powered Personal Health Data Interaction for Chronic Condition Management. This block will be automatically generated when manuscripts are processed after acceptance.

1 INTRODUCTION

Personal informatics (PI) systems have explored various ways of making multidimensional data useful towards objectives, such as general wellbeing, and chronic disease management. Despite the research attention that PI has received, PI systems are often critiqued for limited actionability. For instance, exploratory tools involve considerable cognitive burden in generating actionable insights [5,6,9,13]. Systems that provide recommendations or predictions suffer from interpretation related issues (e.g., misinterpretation, counterintuitive insights, misrepresentation of lived experiences) [7,11,12,14,15,18]. One reason for limited actionability afforded by these systems could be their poor understanding of the process of reflection, that is how data is translated to relevant insights and actions [1,10]. Tools often focus on what insights users want or what data they want to track but not on the nuanced tasks involved in making decisions from the data. As PI tools begin integrating artificial intelligence (AI) capabilities, we must understand what tasks, as carried out by users, need to be automated to support decision-making; and the ways in which human and AI capabilities can cooperate to accomplish those tasks. To better understand and represent the nuances of making data actionable, we present a framework called *episode-driven data interaction* that identifies different tasks involved in making decisions from personal health data in the context of Type 1 diabetes management. We briefly describe the components of this framework. We then discuss the challenges and opportunities for embodying this framework in the design of AI-powered PI systems.

2 FRAMEWORK: EPISODE-DRIVEN DATA INTERACTION

Our framework represents data-driven decision-making involved in managing Type 1 diabetes (T1D), which happens based on three primary types of data – blood glucose (BG), carbohydrates, and insulin. The framework resulted from an analysis of 71 data review sessions between caregivers of teen T1D patients and their clinicians. It is anchored in the concept of episodes, which we define as *phases of suboptimal management* indicated by recurring or one-time events in outcome or behavior data. The framework describes ways in which an episode, as an analytic lens, governed how data was understood and acted upon through four tasks - episode detection, episode elaboration, episode classification, and episode-specific recommendation generation.

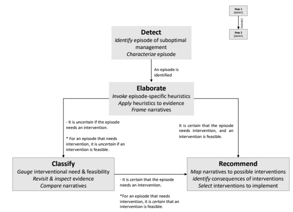

Figure 1: Episode-driven data interaction framework

Episode detection: It involved identifying phases of suboptimal management and characterizing them for severity and concern. An episode could depict a recurring suboptimal outcome (e.g., a high BG event after lunch on several days, a low BG event at bedtime on weekends) or suboptimal behavior (e.g., missed insulin dose at lunch for several days). Episodes could also indicate a one-time occurrence of an unusual outcome or behavior.

Episode elaboration: This involved explaining an episode detected in one data stream (e.g., blood glucose) in terms of other data streams (e.g., carbohydrates, insulin). It happened by creating narratives guided by clinical heuristics that map an episode to patient behaviors and corrective interventions (e.g., overnight high blood glucose is usually caused by insufficient long-acting insulin – increase long-acting insulin), and that guide user attention to different subsets of data.

Episode classification: It involved considering different characteristics of the episode, such as severity and confidence with which it could be explained, to gauge the need and feasibility of fixing it. It resulted in episodes that could be fixed and needed to be fixed, ones that could be fixed but need not be fixed, and ones that should be fixed but cannot be fixed.

Episode-specific recommendation generation: For an episode that was feasible to fix and that required a fix, different types of recommendations could be identified - one-time changes, situational changes, and anticipated changes to regimen or behaviors.

3 DESIGNING FOR AI-POWERED EPISODE-DRIVEN DATA INTERACTION

Designing tools embodying the episode-driven framework would require providing support for each framework component and for bringing the components together. The individual components represent computational problems that are largely considered solved. For instance, pattern recognition and anomaly detection techniques can help identify episodes [20]. Given the role of narratives in episode elaboration, automated narrative generation approaches could be applied [3]. With labelled data, classification algorithms could identify different episode types based on the need and feasibility of an intervention [20]. A knowledge-based recommender technique can be used to generate episode-specific recommendations [8]. The primary research challenge then is not how to automate the different tasks that the framework components represent but to what extent these should or could be automated; and how information generated by them can be brought together and delivered to the user. In this context, we discuss two questions towards the design of AI-powered episode-driven data interaction tools.

3.1 What are the different combinations in which AI and human abilities can operate within and across framework components (or decision-making tasks)?

In choosing to implement a certain combination of human and AI abilities, research can draw from Parasuraman et al.'s model for types and levels of automation [16]. Individual components of the framework can be automated to varying extents and can be brought together through different levels of automation. We contrast two approaches to design for episode-driven data interaction – a human-driven approach and an AI-driven approach.

Human-driven AI-in-the-loop approach: Data savvy users might benefit from an AI-in-the-loop approach to episode-driven data interaction wherein they drive data-driven reasoning by drawing upon AI abilities as helpers in a decision-making scenario. These AI capabilities would correspond to different components that the episode-driven framework outlines and that correspond to different computational tasks – pattern recognition or anomaly detection, automated narrative generation, pattern classification, and recommendation generation. Here, the role of AI would be to moderate data selection and presentation to the user in the context of a decision-making scenario. For instance, given a specific type of episode, AI could extract meaningful data and select the most suitable visualizations to present that data.

AI-driven human-in-the-loop approach: Users who are relatively less data savvy or who want little cognitive burden will benefit from a human-in-the-loop approach wherein AI drives data-based reasoning calling human attention to the results of the different components of the episode-driven data interaction framework when required and to the extent required. Here, the potential role of AI would be to moderate the type of user engagement required to generate value from data, which could adapt depending on the confidence score associated with the computational units corresponding to the different components of the framework. For e.g., with a higher confidence score, the recommendation generation component might simply present a recommendation without the need for the user to interact. With a lower confidence score, it might want to present episodes enhanced with additional data, such as causal events and potential recommendations, while guiding user attention and input in identifying the most suitable recommendation. In guiding users, AI systems can borrow mechanisms used in visual analytics systems for guiding data experts in generating insights, such as orienting the user, directing the user, and prescribing certain tasks or insights to the user [4].

3.2 What types of support could we envision in helping users consume and interact with the results of the computational components – individually or taken together?

Here we describe two challenges and the need to design beyond automation to help users leverage the results generated by an intelligent system.

Handling multiplicity and inaccuracy of results: Computational approaches to provide insights (e.g., recommendations, predictions) from data might result in multiple insights. For instance, in the context of the episode-driven framework, an AI system might result in detection of multiple episodes, provision of multiple recommendations in the context of an episode, or provision of conflicting recommendations for two different episodes. How might AI capabilities help the user select and prioritize relevant insights? Systems could benefit from surfacing examples of data behind AI-generated insights. Additionally, characterizing insights in terms of properties, such as evidence strength, and necessity, might help users in making informed judgments.

Explaining AI-generated results to users: AI-driven insights may not be self-explanatory, more so when they are not representative of users' lived experiences and have scope for misinterpretation (e.g., oversimplification of correlations) [11,12,17]. Systematic techniques are emerging to explain the inner workings of AI systems to users [19]. For instance, providing examples has shown to aid users' understanding of the system [2]. The data associated with the different components of the episode-driven framework can be used as the building blocks for explaining AI-generated results, especially recommendations for an episode. To present relevant data, there is a need to develop techniques for identifying the most representative data in the context of a particular insight (an episode, a recommendation, a narrative); and strategies for presenting it to the user for different objectives, such as explaining a recommendation, comparing recommendations or episodes, and ranking recommendations or episodes.

REFERENCES

1. Eric P.S. Baumer, Vera Khovanskaya, Mark Matthews, Lindsay Reynolds, Victoria Schwanda Sosik, and Geri Gay. 2014. Reviewing Reflection: On the Use of Reflection in Interactive System Design. In *Proceedings of the 2014 Conference on Designing Interactive Systems* (DIS '14), 93–102. https://doi.org/10.1145/2598510.2598598
2. Carrie J. Cai, Jonas Jongejan, and Jess Holbrook. 2019. The effects of example-based explanations in a machine learning interface. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*, 258–262. https://doi.org/10.1145/3301275.3302289
3. David Caswell and Konstantin Dörr. 2018. Automated Journalism 2.0: Event-driven narratives. *Journalism Practice* 12, 4: 477–496. https://doi.org/10.1080/17512786.2017.1320773
4. Davide Ceneda, Theresia Gschwandtner, Thorsten May, Silvia Miksch, Hans-Jorg Schulz, Marc Streit, and Christian Tominski. 2017. Characterizing Guidance in Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 23, 1: 111–120. https://doi.org/10.1109/TVCG.2016.2598468
5. Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche, and Dominikus Baur. 2017. Understanding Self-Reflection: How People Reflect on Personal Data through Visual Data Exploration. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth'17). ACM, New York, NY, USA.*
6. Eun Kyoung Choe, Nicole B. Lee, Bongshin Lee, Wanda Pratt, and Julie A. Kientz. 2014. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1143–1152. https://doi.org/10.1145/2556288.2557372
7. Nediyana Daskalova, Danaë Metaxa-Kakavouli, Adrienne Tran, Nicole Nugent, Julie Boergers, John McGeary, and Jeff Huang. 2016. SleepCoacher: A Personalized Automated Self-Experimentation System for Sleep Recommendations. 347–358. https://doi.org/10.1145/2984511.2984534
8. Robin De Croon, Leen Van Houdt, Nyi Nyi Htun, Gregor Štiglic, Vero Vanden Abeele, and Katrien Verbert. 2021. Health Recommender Systems: Systematic Review. *Journal of Medical Internet Research* 23, 6: e18035. https://doi.org/10.2196/18035
9. Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In *Proceedings of the 2014 conference on Designing interactive systems*, 667–676. Retrieved April 17, 2017 from http://dl.acm.org/citation.cfm?id=2598558
10. Victoria Hollis, Artie Konrad, Aaron Springer, Matthew Antoun, Christopher Antoun, Rob Martin, and Steve Whittaker. 2017. What Does All This Data Mean for My Future Mood? Actionable Analytics and Targeted Reflection for Emotional Well-Being. *Human–Computer Interaction* 32, 5–6: 208–267. https://doi.org/10.1080/07370024.2016.1277724
11. Simon L. Jones. 2015. Exploring correlational information in aggregated quantified self data dashboards. 1075–1080. https://doi.org/10.1145/2800835.2800963
12. Ravi Karkar, Jasmine Zia, Jessica Schroeder, Daniel A. Epstein, Laura R. Pina, Jeffrey Scofield, James Fogarty, Julie A. Kientz, Sean A. Munson, and Roger Vilardaga. 2017. TummyTrials: A Feasibility Study of Using Self-Experimentation to Detect Individualized Food Triggers. 6850–6863. https://doi.org/10.1145/3025453.3025480
13. Rafal Kocielnik, Lillian Xiao, Daniel Avrahami, and Gary Hsieh. 2018. Reflection Companion: A Conversational System for Engaging Users in Reflection on Physical Activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2: 70:1-70:26. https://doi.org/10.1145/3214273
14. Zilu Liang, Bernd Ploderer, Wanyu Liu, Yukiko Nagata, James Bailey, Lars Kulik, and Yuxuan Li. 2016. SleepExplorer: a visualization tool to make sense of correlations between personal sleep data and contextual factors. *Personal and Ubiquitous Computing* 20, 6: 985–1000. https://doi.org/10.1007/s00779-016-0960-6
15. Lena Mamykina, Elizabeth M Heitkemper, Arlene M Smaldone, Rita Kukafka, Heather Cole-Lewis, Patricia G Davidson, Elizabeth D Mynatt, Jonathan N Tobin, Andrea Cassells, Carrie Goodman, and George Hripcsak. 2016. Structured scaffolding for reflection and problem solving in diabetes self-management: qualitative study of mobile diabetes detective. *Journal of the American Medical Informatics Association* 23, 1: 129–136. https://doi.org/10.1093/jamia/ocv169
16. R. Parasuraman, T.B. Sheridan, and C.D. Wickens. 2000. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 30, 3: 286–297. https://doi.org/10.1109/3468.844354
17. Shriti Raj, Joyce M. Lee, Ashley Garrity, and Mark W. Newman. 2019. Clinical Data in Context: Towards Sensemaking Tools for Interpreting Personal Health Data. *IMWUT* Under review.
18. Konrad Tollmar, Frank Bentley, and Cristobal Viedma. 2012. Mobile Health Mashups: Making sense of multiple streams of wellbeing and contextual data for presentation on a mobile device. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on*, 65–72. Retrieved December 13, 2015 from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6240364
19. Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y. Lim. 2019. Designing Theory-Driven User-Centric Explainable AI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3290605.3300831
20. Ashenafi Zebene Woldaregay, Eirik Årsand, Taxiarchis Botsis, David Albers, Lena Mamykina, and Gunnar Hartvigsen. 2019. Data-Driven Blood Glucose Pattern Classification and Anomalies Detection: Machine-Learning Applications in Type 1 Diabetes. *Journal of Medical Internet Research* 21, 5: e11030. https://doi.org/10.2196/11030