Realizing the Potential of Personal Health Informatics Through A Personal Semantic Health Knowledge Graph

DANIEL M. GRUEN, Rensselaer Polytechnic Institute, USA
OSHANI SENEVIRATNE, Rensselaer Polytechnic Institute, USA
CHING-HUA CHEN, IBM T. J. Watson Research Center, USA
SHRUTHI CHARI, Rensselaer Polytechnic Institute, USA
SOLA SHIRAI, Rensselaer Polytechnic Institute, USA
DEBORAH L. MCGUINNESS, Rensselaer Polytechnic Institute, USA

Supporting the effective use of personal informatics for health requires semantic capabilities that combine and connect data points in meaningful ways, and support their integration with other systems and information sources. This combination of semantic structure and specific health data can form what we describe as a personal health knowledge graph. Creating such a graph includes the typical challenges of collecting and integrating data from multiple sources, combining and harmonizing different semantic schemas, and supporting effective and user-understandable mechanisms for privacy and control. In addition, to be maximally useful, a truly personal health knowledge graph must include information on the users' goals and motivations, mental models, knowledge and beliefs.

Additional Key Words and Phrases: Knowledge Graph, Personal Informatics, Semantic Web, Explainable AI

ACM Reference Format:
Daniel M. Gruen, Oshani Seneviratne, Ching-Hua Chen, Shruthi Chari, Sola Shirai, and Deborah L. McGuinness. 2022. Realizing the Potential of Personal Health Informatics Through A Personal Semantic Health Knowledge Graph . In . ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

Supporting the effective use of personal informatics in applications such as those aimed at improving health-related behaviors requires semantic capabilities that combine and connect data points in meaningful ways, and support their integration with other systems and information sources [1, 5]. These include medical guidelines, information on nutritional values of foods, findings from scientific studies, and insights from population data. This combination of semantic structure and specific data points can form what we describe as a personal health knowledge graph. Creating such a graph includes the typical challenges with any data federation endeavor of collecting and integrating data from multiple sources, and harmonizing multiple semantic schemes. In addition, to be maximally useful, a truly personal health knowledge graph must include information on the users' goals and motivations, and on their own mental models, knowledge and beliefs.

Such a truly personal health knowledge graph will be valuable in enabling user facing systems to provide analytical insights, notifications, and explanations that are:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2022 Association for Computing Machinery.
Manuscript submitted to ACM

Appropriately Informative - understood by the user, leveraging what they already know while providing additional information not redundant to what they know;

Goal and Task Relevant - useful in helping the user accomplish what they want to (or should) do, for example by focusing on items that are actionable and changeable by the user, with messages delivered at most appropriate times;

Important - selecting, from the vast amount of potential data, statistical trends and analytic insights - those to which a user's attention should be drawn - based on an understanding of which items are most critical to achieving desired outcome. This could be, for example, because they represent a critical risk factor for the user or because they support presenting a message that the system believes will be maximally motivational to the user in driving desired behavioral change.

2 Challenges

To create such a comprehensive, personal knowledge graph, enabling the meaningful integration of data on user medical conditions and behaviors, scientific knowledge, with patients' goals, motivations, knowledge and beliefs, requires that numerous challenges be overcome. These include:

(1) Inferring and representing a user's assumed prior knowledge, including predicting what information would be redundant, what is already known and can be referenced and built on, what would be surprising and needing explanation [3];
(2) Identifying and representing user's goals (either by explicit identification such as by the user, a clinician or caregiver; by appealing to explicit guidelines; by inference from other data collected such as on user behavior; by comparing user data to trends seen among similar users.);
(3) Identifying and representing user's motivations (such as by analyzing user behavior over time and identifying which messages and interventions are correlated with subsequent behavioral changes);
(4) Appropriately integrating with information from the personal health knowledge graphs of others (population data) while preserving the flexibility of the individual's data representation
(5) Determining appropriate; sub-populations (collections of other users) with whom to compare data for predictions and inferences;
(6) Accomplishing the above in ways that preserve privacy while enabling analyzing data and deriving insights across users;
(7) Accomplishing the above in ways that insure equity, inclusiveness, and respect for cultural differences in the data collected and semantic schemes used, while preserving individual differences and avoiding computational stereotyping.

3 Lessons from Two Health Related Projects

Our thoughts on the value of a comprehensive knowledge graph for deriving maximum value from personal informatics are informed by our experiences with two health-related applications, specifically, a patient-facing health-coach application and a machine-learning based disease risk prediction application for use by clinicians.

Project 1: Our personal health coach was focused on helping people who are pre-diabetic or newly diagnosed with diabetes know which specific dietary guidelines applied to them, understand the extent to which their existing eating behaviors matched those guidelines, and then modify their existing dietary behaviors to better adhere to the guidelines [4][1]. It's design was guided by interactions with professional dieticians, and combined information on the patient's

[1] A demo our personal healthcoach application is available at https://foodkg.github.io/demo.html.

diagnosis, health measures, food log, ADA guidelines and related medical sources. A comprehensive personal knowledge graph would enable capabilities we identified as important, including:

- Identifying insights from the patient's behavior most important to show because of their intrinsic value to health (e.g. stop drinking soda);
- Identifying insights from behavior seen to be most motivational to that specific user (e.g. for a user who responds to positive messages, "you have stayed within your caloric level on 4 of the last 5 weekends");
- Addressing suspected possible gaps in personal knowledge (e.g. fruit juice is high in sugar);
- Providing healthier suggestions likely to be appealing (e.g. recommending a food very similar to, but significantly healthier than, a food the user consumes often).

Project 2: Our clinical risk prediction support tool focused on surfacing the results of an ML model which predicted the risk of a patient developing chronic kidney disease, along with additional related information of various types to explain the results and make them more useful to clinicians [2]. This work was guided by interactions with a panel of clinicians, and pointed to the importance of rich semantic graphs in guiding the generation and presentation of results to support capabilities seen as important, for example:

- Focusing in actionable insights, e.g. items changeable by clinician and user, taking into accounts specifics of the patient including their proximity to treatment centers, financial capabilities, and insurance;
- Focusing on new information - e.g explaining results that might be surprising to a general practitioner;
- Drawing attention to information on the patient the clinician may not have noticed, e.g. correlation with a condition appearing a long time ago in patient history.

4 Conclusion

As the quantity and variety of data collected on people and their behavior increases, and as new relationships between social and environmental factors and health continue to be discovered, a personal knowledge graph that represents and connects heterogeneous information related to a patient in principled ways is important. Such a knowledge graph can support the federation of information from different systems, enriched with relevant domain knowledge. Furthermore, to support effective human interaction with computational health systems, such a knowledge graph should include information on users' knowledge and beliefs. While challenges exist to creating a truly comprehensive personal health knowledge graph, our work on multiple projects has shown the value of taking steps in that direction.

Acknowledgments

This work is partially supported by IBM Research AI through the AI Horizons Network.

References

[1] Krisztian Balog and Tom Kenter. 2019. Personal knowledge graphs: A research agenda. In *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval*. 217–220.

[2] Shruthi Chari, Prithwish Chakraborty, Mohamed Ghalwash, Oshani Seneviratne, Elif K Eyigoz, Daniel M Gruen, Fernando Suarez Saiz, Ching-Hua Chen, Pablo Meyer Rojas, and Deborah L McGuinness. 2021. Leveraging Clinical Context for User-Centered Explainability: A Diabetes Use Case. *KDD Applied Data Science in Healthcare workshop, arXiv preprint arXiv:2107.02359* (2021).

[3] Daniel Gruen, Shruthi Chari, Oshani Seneviratne, and Deborah L McGuinness. 2021. A personal health-knowledge knowledge graph. From KGC'21 PHKG Workshop. https://https://www.slideshare.net/DanGruen/kgc-21-personal-healthknowledge-knowledge-graphs. Accessed: 2022–02-22.

[4] Nidhi Rastogi and Mohammed J Zaki. 2020. Personal Health Knowledge Graphs for Patients. *arXiv preprint arXiv:2004.00071* (2020).

[5] Sola Shirai, Oshani Seneviratne, and Deborah L McGuinness. 2021. Applying Personal Knowledge Graphs to Health. *arXiv preprint arXiv:2104.07587* (2021).