Chapter 10
Semantic Technologies for Clinically Relevant Personal Health Applications

Ching-Hua Chen, Daniel Gruen, Jonathan Harris, James Hendler, Deborah L. McGuinness, Marco Monti, Nidhi Rastogi, Oshani Seneviratne, and Mohammed J. Zaki

Abstract Despite recent advances in digital health solutions and machine learning, personal health applications that aim to modify health behaviors are still limited in their ability to offer more personalized decision support. Moreover, while many personal health applications cater to general health and well-being, there remains a significant opportunity to increase the clinical relevance of the insights being generated. This chapter describes the motivation for, and illustrative applications of, semantic technologies for enabling clinically relevant personal health applications. We present two use cases that demonstrate how semantic web technologies, in combination with machine learning and data mining methods, can be used to provide personalized insights to support behaviors that are consistent with nutritional guidelines for people with diabetes.

Keywords Semantic web · Knowledge graphs · Artificial intelligence · Personal health · Health behavior · Diabetes self-management · Consumer health · Decision support systems

C.-H. Chen (✉)
Center for Computational Health, IBM Research, Yorktown Heights, NY, USA
e-mail: chinghua@us.ibm.com

D. Gruen · J. Harris · J. Hendler · D. L. McGuinness · N. Rastogi · O. Seneviratne · M. J. Zaki
Rensselaer Polytechnic Institute, Troy, NY, USA
e-mail: gruend2@rpi.edu; harrij15@rpi.edu; hendler@cs.rpi.edu; dlm@cs.rpi.edu; nidhi.rastogi@rit.edu; senevo@rpi.edu; zaki@cs.rpi.edu

M. Monti
Cognitive AI and Analytics, IBM Global Business Services, Circonvallazione Idroscalo, Segrate, Milan, Italy
e-mail: marco.monti@it.ibm.com

P.-Y. S. Hsueh et al. (eds.), *Personal Health Informatics*, Cognitive Informatics in Biomedicine and Healthcare, https://doi.org/10.1007/978-3-031-07696-1_10

Decision Support for Health Behavior Change

Health outcomes are known to be driven by a combination of medical, genetic and lifestyle factors. In the United States, a disproportionate emphasis is placed on medical treatment, as compared to lifestyle modifications (Bipartisan Policy Center 2012). Where the former is primarily delivered in reaction to poor health status, the latter is often used as a form of disease prevention and/or health maintenance. As such, efforts to implement lifestyle modifications often rest on the shoulders of patients (or more generally, health consumers) with sufficient means, skills and motivation. The process of behavior change is well-studied. Yet, sustained behavior change remains challenging to intervene effectively on (Bouton 2014). While behavior change is recognized by experts as being a complex process involving dynamic and stochastic factors that span the psychological, social and physical domains, popular misconceptions are that changing one's behavior requires no more than 'common sense' or a good marketing campaign, and that most people will rationally process relevant knowledge and information (Kelly and Barker 2016). Arguably, interventions that are adaptive and sensitive to an individual's psychological, social and environmental context, are in a better position to address behavior change than those that are static, or adopt a 'one-size-fits-all' approach.

With the rapid adoption of mobile phones and wearable sensing technologies, most people now have access to mobile applications that can provide real-time sensing and feedback to their users. This trend has led to the development and study of several 'context-aware' digital technologies for tackling health behavior change (Thomas Craig et al. 2020). Common approaches of incorporating contextual awareness into digital behavior change interventions include the use of statistical and machine learning models to generate feedback based on user generated data (e.g., step counts and other forms of physical activity, food logs, sleep logs), as well as the use of rule-based dialog systems, or chat bots, that provide deterministic responses to user textual inputs that conform to anticipated patterns. However, most mobile health applications, while popular among patients, have not seen significant levels of acceptance from clinicians (Gordon et al. 2020). This lack of clinical acceptance is partly explained by factors pertaining to regulations, payment systems, and clinical workflows. It may also be explained by the limited incorporation of evidence-based, clinical guidelines into the function and design of mobile health applications.

Semantic technologies are well-suited for representing clinical knowledge that has been curated by medical and health experts. When semantic technologies that can represent and reason over clinical knowledge are used together with machine learning methods that learn from and adapt to the 'big data' that is continuously generated by activities of daily living, there is the potential to improve the clinical relevance of personal health applications. With a few notable exceptions (Michie et al. 2017; Dragoni et al. 2020; Chen et al. 2021) there has been limited work exploring the use of semantic technologies for health behavior change. This chapter aims to introduce readers to semantic technologies and the potential benefits that they present for enhancing the personalization, interpretability and clinical utility of personal health applications.

The objective of this chapter is to provide an introduction to semantic technologies to health informatics researchers and practitioners, and to demonstrate their application in combination with other artificial intelligence methods (e.g., data mining and machine learning) via exemplary use cases pertaining to people with diabetes. These use cases were selected to highlight how clinical and health knowledge can be combined with "big data" sources of personal behaviors and personal context, to provide insights that are relevant to both health consumers and the clinicians who serve them. The remainder of this chapter is organized as follows: In Sect. "Semantic Technologies and the Personal Health Knowledge Graph", we provide an introductory overview of semantic technologies, highlighting key concepts related to knowledge graphs and defining Personal Health Knowledge Graph (PHKG). In Sect. "Combining Learning and Logic for Personal Health Applications", we explain how methods that combine machine learning and semantic technologies are able to exploit the best of machine learning and knowledge graphs, allowing computers to simultaneously tap into deep data and deep knowledge. To ground our discussion in a personal health application, Section "Nutrition Self-Management for People with Type 2 Diabetes" describes the experiences of people with type 2 diabetes who are engaging in self-management behaviors, and includes two examples of how semantic technologies have been used in conjunction with machine learning and data mining to generate personalized and context-aware meal recommendations. We close our chapter with a discussion of the many opportunities we see for using semantic technologies in the pursuit of improving personal health applications for health consumers.

Semantic Technologies and the Personal Health Knowledge Graph

Semantic technologies are used to enable computers to process data in ways that leverage the meaning of terms, such as through the use of logical reasoning. At the heart of these technologies is the *knowledge graph* (KG), which has been defined as "as a graph of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of interest and whose edges represent relations between these entities." (Hogan et al. 2022) DBpedia, YAGO and Wikidata are examples of public knowledge graphs generated from content available in Wikipedia, a crowd-sourced encyclopedia available on the Internet (Ringler and Paulheim 2017; Abián et al. 2018; Pillai et al. 2019). Knowledge graphs inherit from classic artificial intelligence such formalisms as semantic networks and description logics (Baader et al. 2007). The advantages of using knowledge graphs to represent knowledge are that they are amenable to the linking of knowledge across multiple sources and domains (through identifying overlapping semantic concepts across ontologies). To be regarded as high-quality, knowledge represented in knowledge graphs should be consistent, and feature a certain degree of completeness, accuracy and timeliness (i.e., degree to which knowledge is kept up-to-date) along with containing provenance content (where the knowledge came from). *Semantic reasoners* are capable of inferring new knowledge from the data contained in knowledge graphs. These reasoners may be based on logic (e.g., first-order logic, predicate logic, non-monotonic logic), fuzzy logic, or machine learning. The use of machine learning methods for reasoning over KGs has been of rising interest in the artificial intelligence community, due to the rapid and parallel growth in availability of very large, electronic data sets and access to computing power. Section "Combining Learning and Logic for Personal Health Applications" of this chapter discusses the advantages of combining machine learning and semantic technologies, and our subsequent use cases in Sect. "Nutrition Self-Management for People with Type 2 Diabetes" demonstrate a combined use of both types of approaches for personal health applications. For a recent survey on methods for reasoning over knowledge graphs, the reader is referred to Chen et al. (Chen et al. 2020).

Semantic technologies are often at the core of interactive decision-support systems that have to deal with complex knowledge. They are useful for addressing key challenges in knowledge management such as finding, summarizing or answering questions pertaining to information contained in electronic medical records, legal documents and scientific literature. Typical functions performed using semantic technologies include: entity summarization, faceted search, and question answering. Entity summarization involves generating a concise description of what is known about an entity, such that it satisfies users' information needs (Liu et al. 2021; Cheng et al. 2020). Faceted search is a method of finding information that allows users to progressively navigate towards more relevant results using filters that are meaningful within the search domain (e.g., searching for recipes based on filters for nutritional content, cuisine, preparation time, etc.) (Arenas et al. 2016). Question and answering over knowledge bases allows users to seek answers (from the knowledge graph) to questions posed in natural language (Arenas et al. 2016; Moschitti et al. 2017). Before the invention of the World Wide Web (WWW), semantic technologies were used within large organizations with significant institutional knowledge bases, and wherein knowledge representation could be centralized (Pan et al. 2017). With the invention of the WWW, the potential for semantic technologies to enable intelligent agents that could 'traverse' globally linked knowledge became an exciting and real proposition (Berners-Lee et al. 2001; Hendler 2003). Applications using the KG should be able to provide a set of knowledge services, which should be feature high reliability (e.g., fast response time, and high fault tolerance) and high usability (e.g., good learnability).

When constructing knowledge graphs, the usual assumption is that the entities and the relationships between entities are shaped by domain experts, who define an *ontology*. The ontology defines the vocabulary that is used to describe the various concepts, relations and axioms that need to be represented in the knowledge graph.

The knowledge graph then uses the terms from that ontology when representing assertions regarding individuals, instances within the domain of interest. The ontology may be partially or entirely contained within the knowledge graph itself. The process of ontology engineering (Kendall and McGuinness 2019) lies in capturing necessary and sufficient conditions for including terms, and connections between terms, in the ontology. Complementary to such a 'top-down' approach is a 'bottom-up' approach, wherein new knowledge is generated (through descriptive statistics and/or logical inference) from specific instances of the data. Using this approach, new categories and related concepts can be derived, resulting in the creation of new knowledge. Since the construction of large knowledge graphs can be time consuming, various efforts exist to increase the degree of automation of knowledge graph construction. For example, the Semantic Data Dictionary (SDD) approach is able to facilitate automatic creation of knowledge graphs by semantically annotating tabular data with concepts from existing ontologies (Rashid et al. 2020). Furthermore, the automated knowledge base construction community has been employing natural language processing techniques to develop knowledge graphs (Suchanek et al. 2013a), and these efforts have given rise to the Automated Knowledge Base Construction workshop series (Suchanek et al. 2013b), that has now become a full-fledged conference (https://www.akbc.ws), which supplements parallel efforts by the larger semantic web community.

The World Wide Web Consortium (W3C) has established standards for implementing semantic technologies. The Resource Description Framework (RDF) is the basic mechanism through which basic statements can be made. The RDF data model is based upon the idea of making statements about resources in expressions of the form *subject–predicate–object*, known as an RDF *triple*. The *subject* denotes the resource, and the *predicate* denotes traits or aspects of the resource and expresses a relationship between the subject and the *object*. For example, one way to represent the statement "The lasagna contains meat" in RDF is as the triple: a subject denoting "the lasagna", a predicate denoting "contains", and an object denoting "meat". RDF triples can be serialized using several alternative syntaxes, including *N-Triples*, *Turtle*, *RDF/XML*, and *JSON-LD*. Examples of how the triple for "the lasagna"-"contains"-"meat" using the alternative RDF data formats are shown below.

Using *N-Triples* syntax:

```
<http://example.com/exampleOntology#Lasagna>
<http://example.com/exampleOntology#contains>
<http://example.com/exampleOntology#Meat> .
```

Using *Turtle* syntax:

```
@prefix ex: <http://example.com/exampleOntology#> .
ex:Lasagna ex:contains ex:Meat .
```

Using *RDF/XML* syntax:

```
<?xml version="1.0" encoding="utf-8" ?>
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-
      ns#" xmlns:ns0="http://example.com/exampleOntology#">

<rdf:Description rdf:about="http://example.com/
    exampleOntology#Lasagna">
<ns0:contains rdf:resource="http://example.com/
    exampleOntology#Meat"/>
</rdf:Description>
</rdf:RDF>
```

Using the *JSON-LD* syntax:

```
[
  { "@id":"http://example.com/exampleOntology#Lasagna",
    "http://example.com/exampleOntology#contains":[
            {"@id":"http://example.com/exampleOntology#Meat"}
    ]
  },
  {"@id":"http://example.com/exampleOntology#Meat"}
]
```

While RDF is a way of representing knowledge graphs, languages such as the RDF Schema (RDFS) language and the Web Ontology Language (OWL) can be used to define ontologies. While OWL is more expressive than RDFS, it is also more complex to use. Both OWL and RDFS are recommended standards by the W3C. For more on semantic modeling in RDFS and OWL, readers are referred to an introductory text by Allenmang and Hendler (Allemang et al. 2020) and Ontology Engineering text by Kendall and McGuinness (Kendall and McGuinness, 2019).

The predominant query language for RDF graphs is SPARQL, (pronounced *spahr- kuhl*, and it is the recursive acronym for SPARQL Protocol And Query Language) is an SQL-like query language for RDF that has been standardized by the W3C. The following is an example of a SPARQL query to show all foods contained within a menu named `italian_menu`, using a fictional ontology called `exampleOntology`:

```
PREFIX ex: <http://example.com/exampleOntology#>
SELECT ?food ?menu
WHERE {
    ?x ex:foodname ?food ; ex:isContainedin ?y .
    ?y ex:menuname ?menu ; ex:isInMenu ex:italian_menu .
}
```

While most popular knowledge graphs capture entities that are of global relevance (i.e., of interest to the general population), knowledge graphs that capture data that is relevant only to a particular individual (i.e., a *personal knowledge graph*), can also be useful. Given the large amount of data that is now being tracked and recorded from personal activities, and increased consumer demand for more personalized services, in particular for health and wellness, reasoning over a personal knowledge graph presents an opportunity for generating insights highly relevant to the person whose data is represented in the knowledge graph. Moreover, if data in a personal knowledge graph is linked to data in general knowledge graphs, a reasoner could generate insights that relate a personal experience to those in the general population. Balog and Kenter (Balog and Kenter 2019) present the concept of the personal knowledge graph and how it differs from general knowledge graphs. They note an increased, but fragmented amount of research relating to personal knowledge graphs, and propose a research agenda for personal knowledge graphs. Meanwhile, Gyrard et al. (Gyrard et al. 2018) specifically consider the concept of a personal knowledge graph for health, which integrates and represents all health information specific to an individual, including their medical history and health behaviors, as well as relevant socio-environmental factors that the individual may be exposed to. They also identify several research challenges for advancing the state-of-the-art in personal knowledge graphs for health, including how to model and integrate general health and personal health knowledge, and how to analyze data from the Internet-of-Things (IoT) to produce meaningful contextual information for supporting health behavior change. For additional perspectives on personal knowledge graphs for health, the reader is referred to (Rastogi and Zaki 2020) and (Shirai et al. 2021). In this chapter, we consider a *Personal Health Knowledge Graph* (PHKG) to be a knowledge graph representation of a person's health and wellness data. This data may come from various sources (e.g., physical activity trackers, digital food logs, personal health records). In Sect. "Populating a Personal Health Knowledge Graph with Personalized Assessments of Dietary Needs and Preferences" we will describe how a PHKG can be automatically constructed from a user's temporal food log data, and how the PHKG can be used (along with general health knowledge) to derive a user's dietary needs and preferences. Then, in Sect. "Personalizing Dietary Recommendations" we describe how to identify recipes that satisfy these needs and preferences.

Combining Learning and Logic for Personal Health Applications

In the context of personal health, the combination of knowledge graphs and machine learning opens up new possibilities for designing effective digital health assistant applications (Thomas Craig et al. 2020). The use of conversational agents in digital health applications is a popular design choice because it supports natural language queries from the user. These natural language queries need to be converted into SPARQL queries if one wants to answer the query by retrieving information from a knowledge graph. While SPARQL is well-suited to retrieve factual information stored in the knowledge graph, and also to infer answers via reasoning, it is not well-suited for answering ranking based queries (i.e., multiple answers that are sorted in order of relevance) that arise in recommendation settings. Indeed, in personal health applications, users may seek recommendations and/or facts to support decisions about what health behaviors to engage in. In recommendation settings, the answer to the user's query should ideally be personalized to take into account a user's intent, context and constraints. As it turns out, such personalized responses can be provided via machine learning based methods, such as knowledge base question answering (KBQA).

Learning-based methods have the advantages of discovering and leveraging implicit semantics, and can scale to large datasets. However, learning is data-intensive, can produce trivial or known insights and insights are often difficult to explain. Knowledge-based methods have the advantages of being able to explicitly represent and use knowledge without requiring "big data", and this knowledge is easier to transfer between projects. On the other hand, capturing knowledge is labor intensive and logical inference can be computationally intensive. The best of both approaches can be captured via a hybrid approach that injects semantics within machine learning methods, and on flip side, leverages machine learning to scale up semantic approaches. In this section we will highlight the interaction between logic and learning for answering personalized user queries.

Since knowledge graphs store high quality information in a structured format, they are well-suited for answering factual queries by leveraging the underlying semantics. For example, a query like "What are some physical exercises I can try?" can be converted into the following SPARQL query.

```
SELECT DISTINCT ?exerciseName
WHERE {
    ?exercise <http://purl.org/dc/terms/subject> <http://dbpedia.
org/resource/Category:Physical_exercise>;
     <http://www.w3.org/2000/01/rdf-schema#label> ?exerciseLabel .
BIND (STR(?exerciseLabel) AS ?exerciseName)
}
```

Interpreted as a factual retrieval question, this query would return a list of physical exercises, which can then be displayed to the user. However, it is clear that returning a long list of exercises is probably not what the user intends as the response. Rather, the user's context and preferences should be taken into account while answering such a query. For example, taking into consideration the fact that the user might be at the gym, or taking into account their health goals (e.g., lose weight) and their exercise preferences and also their physical ability, and so on. Going even further, this query can be interpreted as asking for recommendation of physical exercises, e.g., "what are some physical exercises I can try that are good for

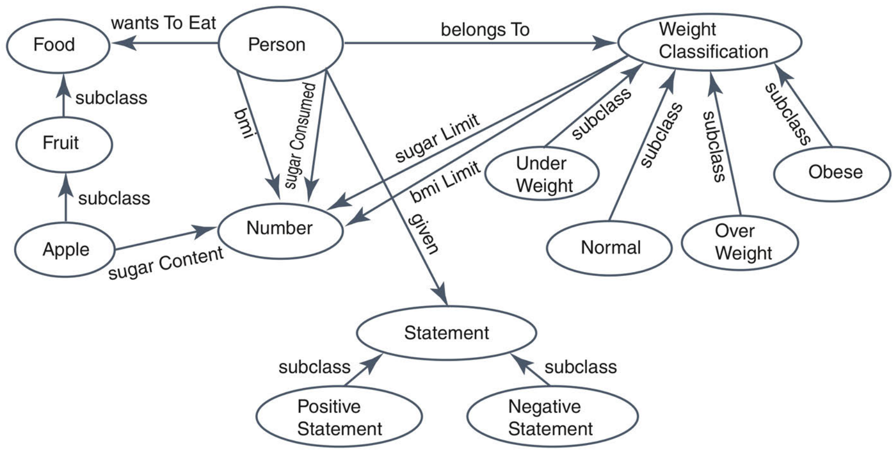

Fig. 10.1 Example of the inference rules and ontology for answering the query "Can I eat a Gala apple?"

me?" Instead of simple retrieval this may require the system to compare alternatives, and then suggest the most beneficial activities at that given place, time, and context, potentially along with an explanation of the suggestion.

As another example, consider the query "Can I eat a Gala apple?" To answer this question well, the system should recognize that there could be an implicit context at play. Namely, the user may be concerned about weight management, or other relevant underlying health conditions. To answer this query we need to rely on a reasoning engine over the personal knowledge graph, as illustrated in Fig. 10.1.

The logic for the inference required to answer this query is captured by the inference rules below.

```
Rule 0:
Subclass Transitivity
Rule 1:
(Person and
(person:bmi > WeightClassification:bmiLimit))
=>
Person belongsTo
[owl:equivalentClass WeightClassification] .

Rule 2:
(Person and
(WeightClassification and (Person:wantsToEat Food) and
(Person:sugarConsumed + Food:sugarContent > WeightClassification:s
ugarLimit))))
=>
Person given NegativeStatement .
```

In this example, the focus is on comparing the sugar limit for the person based on their health condition and status, who are returning that they cannot eat the Gala apple if they exceed the sugar intake limit. This example also illustrates the challenges associated with inferring the user intent and health conditions. Additional constraints besides sugar intake may have to be considered to answer this question adequately. Furthermore, there is the question of automatically deducing the inference rules. So far, we have assumed that an expert provides these. However, this approach is not scalable, and is a challenge that machine learning-based approaches are in a good position to address.

Machine learning can help construct sets of inference rules for reasoning over the KG. They can also help in automatically converting natural language queries to SPARQL queries. Furthermore, learning can help infer the set of active constraints to consider when answering a query—these would span the user's preferences, health guidelines, and all other relevant information. In general, learning is required to hone in on the user intent, as well as to evaluate the relevance of the input constraints and responses. On the other hand, machine learning methods can benefit tremendously from the structured knowledge in the knowledge graphs by leveraging the underlying semantics of the concepts and relationships. For example, knowledge graph embedding methods (Bordes et al. 2011) can be employed to learn concept and relationship embeddings, or representations, that can be used in a deep learning framework to answer user queries.

The combination of semantics and machine learning is even more important when dealing with queries that involve providing recommendations. For example, a user may ask "What is a good breakfast for me?" To answer this type of query, the machine learning framework would have to leverage the interlinked knowledge graphs such as their personal health knowledge graph, a medical guidelines knowledge graph, and a food knowledge graph.

If all of the constraints (e.g., food preferences, allergies, ingredient availability, etc.) are treated as mandatory constraints, the answer is likely to be a null set. While SPARQL provides the OPTIONAL clause to allow for optional constraints, the resulting answer set is not trivial to rank based on relevance to the query (Feyznia et al. 2014). Such queries can be answered by KBQA based methods such as BAMNET (Chen et al. 2019), which is an end-to-end bidirectional attention memory network for complex question answering over a knowledge graph. Readers are referred to Fu et al. (Fu et al. 2020) for an in-depth review of KBQA methods. In more recent work, we have developed a novel system for personalized food recommendation, called *pFoodReq (*Chen et al. 2021*)* that uses constrained question answering over a food knowledge graph to help users search for relevant recipes. We describe *pFoodReq* in detail in Sect. "Personalizing Dietary Recommendations".

Nutrition Self-Management for People with Type 2 Diabetes

Diabetes is a chronic health condition that affects approximately 10.5% of the United States population (National Diabetes Statistics Report 2020). People with diabetes are typically advised to engage in several self-management behaviors in order to improve their health outcomes. People newly diagnosed with diabetes or pre-diabetes, and advised to modify their diet face numerous challenges. In addition to understanding which specific dietary guidelines apply to them, they must also understand how these guidelines translate into specific actions and food choices they can make. Then they must actually implement these changes. Successful behavior change requires understanding and knowledge of the guidelines and nutritional content of different foods and their impact. It requires introspection on their current dietary behavior to understand what changes need to be made and the relative importance of making those changes. Beyond understanding what to do, changing one's diet is notoriously difficult. It can require changing long-term habits, eschewing foods one enjoys, and avoiding foods that are prominent in social gatherings or play an important role in their cultural cuisine. As such, the challenges are both informational—understanding (and remembering) what changes to make and specifically how to implement them, and motivational—providing messages, options, and specific suggestions to encourage making good choices and making doing so as appealing and non-disruptive as possible.

Our current efforts aim to address these challenges by surfacing the health guidelines relevant to a specific user, explaining why those guidelines apply to them, and suggesting foods the user could eat. We also aim to help the user understand their current dietary behavior to see where they are successfully adhering to the guidelines and what changes would be most beneficial to make.

We describe two use cases in the following subsections. In the first use case described in Sect. "Populating a Personal Health Knowledge Graph with Personalized Assessments of Dietary Needs and Preferences", we review a user's food log (i.e., a daily diary of meals consumed) through the lens of a set of relevant dietary guidelines, and generate semantic expressions in the OWL language to represent the gaps between their actual and expected food consumption patterns. In this use case, we combine semantic technologies with data mining methods. In the second use case described in Sect. "Personalizing Dietary Recommendations", suggests specific foods that will fit a user's dietary guidelines and food preferences. In this use case we combine semantic technologies with machine learning. An essential knowledge resource common to both use cases is the Food Knowledge Graph (FoodKG). The FoodKG was constructed by Haussmann et al. (Haussmann et al. 2019) and integrates recipe data from the Recipe1M+ data set (Marín et al. 2021) with ingredient nutritional information from the United States Department of Agriculture's National Nutrient Database for Standard Reference (Haytowitz et al. 2019). The FoodKG uses the FoodOn ontology (Dooley et al. 2018). Resources and instructions for constructing the FoodKG are provided at `https://foodkg.github.io/foodkg.html`.

Populating a Personal Health Knowledge Graph with Personalized Assessments of Dietary Needs and Preferences

In this section, we demonstrate how semantic technologies can be combined with data mining techniques to generate semantic expressions of a user's dietary needs and preferences. In this example, the user's dietary needs are assessed by comparing the user's recent eating patterns with relevant health guidelines set by the American Diabetes Association (ADA) (American Diabetes Association 2021). Any gaps between the user's behaviors and the guidelines are considered to represent the user's current dietary needs. The user's dietary preferences can also be discovered from their reported eating patterns. These dietary needs and preferences can be captured in the user's personal health knowledge graph (PHKG) and queried by downstream applications. Unlike most efforts for automatic KG population, which extract entities and relationships from unstructured text using natural language processing methods, we discover relevant patterns from time-series data in our use case. To support this use case, we created the Personal Health Ontology (PHO) based on a set of interviews conducted with 21 people who declared themselves to be within five years of being diagnosed with type 2 diabetes. Using a semi-structured interview style, we asked participants to describe their eating patterns and probed specifically about the contextual, health and lifestyle factors that influenced their eating behaviors. The PHO differs from existing efforts such as (Puustjarvi and Puustjarvi 2011), which have put a focus on interoperability of various e-health tools through a shared vocabulary. In contrast, our focus was on capturing the personal behavioral preferences. The essential steps involved in populating a PHKG with the user's dietary needs and preferences are four-fold: (i) relevant eating patterns need to be discovered from temporal food log data (ii) eating patterns need to be mapped to a personal health ontology (iii) eating patterns need to be assessed against medical nutrition therapy guidelines (iv) semantic 'directives' for health needs need to be inferred. These steps are depicted in Fig. 10.2.

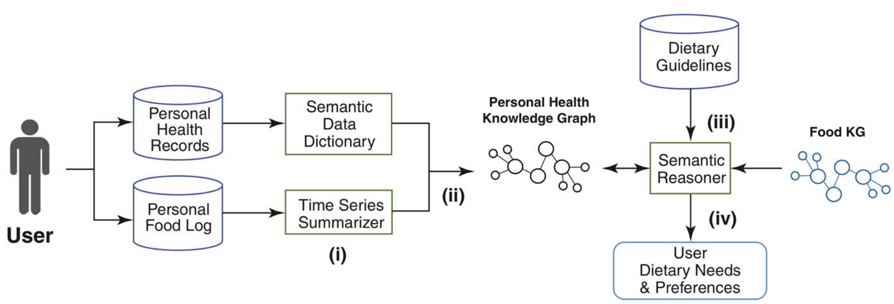

Fig. 10.2 Illustration of how personal health data from the user is transformed by the Time Series Summarizer (Harris et al. 2021) and Semantic Data Dictionary (Rashid et al. 2020) into RDF triples that populate a PHKG. A semantic reasoner is used to generate expressions of the users dietary needs and preferences based on the PHKG and clinical dietary guidelines

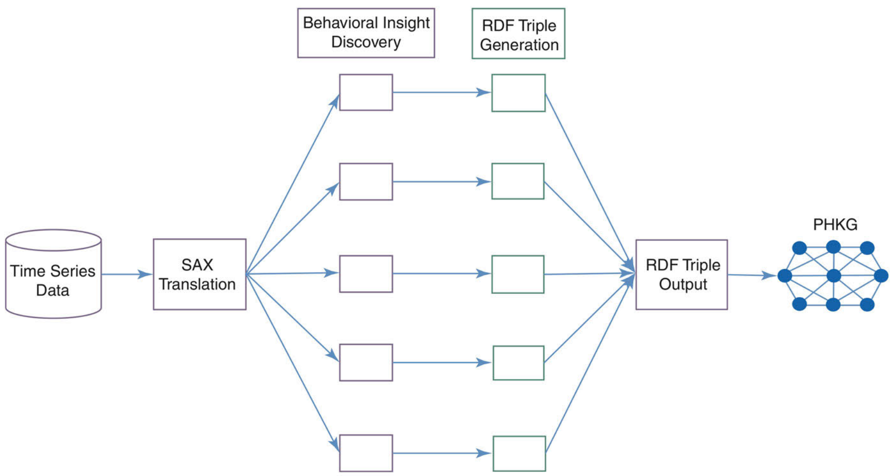

Fig. 10.3 Workflow to discover behavioral insights within a user's food log data and generate RDF triples to populate the PHKG

To implement the use case depicted in Fig. 10.2, we customized an existing Time-Series Summarization (TSS) framework (Harris et al. 2021) to generate RDF triples representing a user's temporal personal health data (e.g., digital food diaries, personal wearables logs). The TSS applies advanced data mining approaches to discover patterns within time-series data. In order to identify 'interesting' patterns, the TSS framework relies on a dimensionality reduction algorithm called Symbolic Aggregate Approximation (SAX) (Lin et al. 2007) to translate the raw time-series data into a string of alphabetical letters (e.g., `'abbbacdae'`). Each of these letters can represent different time granularities (e.g., `'a'` can represent a day or a week in the data). Data mining algorithms, such as the frequent item-set mining tool called SPADE (Zaki 2001) and the categorical clustering algorithm called *Squeezer (*He et al. 2002*)*, are used to search the data for patterns once the data is translated into categorical data. Once a pattern is retrieved, it is represented as a template-based natural language summary, or 'protoform'. An example protoform is "On <*quantifier*><*sub-time window (plural)*> in the past <*time window (singular)*>, your <*attribute*> was <*summarizer*>." Within this protoform, there are five defined placeholders that are each filled with words/phrases chosen from a predefined vocabulary. An example of how this example protoform could be filled is "On *most of the days* in the past *week*, your *calorie intake* was *high*." This framework uses extended versions of rule-based linguistic summarization algorithms that use fuzzy logic to select the correct words/phrases for a protoform (Zadeh 2002; Zadeh 1983; Zadeh 1975; Kacprzyk et al. 2002). TSS was customized to produce RDF triples that would conform with the PHO. The TSS workflow is shown in Fig. 10.3.

An example set of TSS PHKG triples with respect to the user's carbohydrate intake is as follows[1]:

```
:Alice a prov:Person;
        sio:has-attribute :AliceInsulinMedicationDosage,
       :AliceCarbIntakePattern .
:AliceInsulinMedicationDosage a pho:FixedMedicationDosage.
:AliceCarbIntakePattern a pho:ConsistentPattern;
       sio:has-attribute        chebi:carbohydrate,
       :AliceCarbIntakePatternSumm,
:AliceCarbIntakePatternCV,
:AliceCarbIntakeTimeWindow.
:AliceCarbIntakePatternSumm a pho:Summarizer;
       sio:has-value "considerably" .
:AliceCarbIntakePatternCV a stato:CoefficientofVariation;
       sio:has-value "0.99" .
:AliceCarbIntakeTimeWindow a pho:TimeWindow;
       sio:has-value sio:week .
```

Once the RDF triples from the user's daily personal logs have been generated, we implemented a semantic reasoner to evaluate the generated graph against guidelines that determine whether the user has complied with the applicable medical and dietary guidelines. To that end, we modeled several ADA guidelines related to diet and activity into a computable form using OWL. As an example, consider the following ADA guideline recommendation (American Diabetes Association 2021), which we will refer to as 'Dietary-Guideline-01':

> For individuals whose daily insulin dosing is fixed, a consistent pattern of carbohydrate intake with respect to time and amount may be recommended to improve glycemic control and reduce the risk of hypoglycemia.

The guideline contains a *rule* portion that indicates the necessary and sufficient conditions, and a *directive* that indicates what action to take if the rule was evaluated to be true. A semantic reasoner can ingest such ADA guidelines implemented as *rules* and the PHKG triples output from the TSS to recommend a course of action in the form of a *directive*,

---

[1] The full form of the prefixes used in the code listings are as follows:

- **chebi** http://purl.obolibrary.org/obo/chebi#
- **owl** http://www.w3.org/2002/07/owl#
- **pho** http://idea.rpi.edu/heals/pho#
- **prov** http://www.w3.org/ns/prov#
- **rdf** http://www.w3.org/1999/02/22-rdf-syntax-ns#
- **rdfs** http://www.w3.org/2000/01/rdf-schema#
- **sio** http://semanticscience.org/resource/
- **stato** http://purl.obolibrary.org/obo/stato.owl#

The *rule* (i.e., Dietary-Guideline-01) is represented in OWL as follows. First, this rule applies an OWL property restriction on any instances of the `pho:FixedMedicationDosage` on its *has-attribute* property (i.e., our patient should be taking a fixed medication dose for this rule to take effect). Then we check to see if the patient has been following a `pho:ConsistentPattern` of `chebi:carbohydrate` consumption, which is again implemented as an OWL property restriction.

```
pho:Dietary-Guideline-01 rdf:type owl:Class ;
   owl:equivalentClass
     [ owl:intersectionOf (
        [ rdf:type owl:Restriction ;
             owl:onProperty sio:has-attribute ;
             owl:someValuesFrom pho: FixedMedicationDosage ]
        [ rdf:type owl:Restriction ;
             owl:onProperty sio:has-attribute ;
             owl:someValuesFrom [
                owl:intersectionOf (
                    pho:ConsistentPattern
                       [ rdf:type owl:Restriction ;
                         owl:onProperty sio:has-attribute ;
                         owl:someValuesFrom chebi:carbohydrate ] ) ;
                           rdf:type owl:Class]
                         ] ) ;
             rdf:type owl:Class] ;
    rdfs:subClassOf pho:DietaryGuideline ;
    rdfs:label "For a diabetic individual, if their daily insulin
dosing is fixed, and there is a consistent pattern of carbohydrate
intake with respect to time and amount, that pattern should be main-
tained." .
```

Note that concepts such as `pho:FixedMedicationDosage`, `pho:ConsistentPattern`, and `chebi:carbohydrate` that are mentioned in the rule are defined in the corresponding ontologies, i.e. Personal Health Ontology (PHO) and Chemical Entities of Biological Interest (ChEBI). For example, the consistent pattern is defined as follows, which indicates that a `pho:ConsistentPattern` should consist of some `pho:TimeWindow` (i.e., week, day, month, etc.) and a `pho:Summarizer` (slightly, considerably, etc.):

```
pho:ConsistentPattern rdf:type owl:Class ;
    owl:equivalentClass [
    owl:intersectionOf (
        [ rdf:type owl:Restriction ;
              owl:onProperty sio:has-attribute ;
              owl:someValuesFrom pho:TimeWindow ]
          [ rdf:type owl:Restriction ;
              owl:onProperty sio:has-attribute ;
              owl:someValuesFrom pho:Summarizer ] ) ;
    rdf:type owl:Class] ;
    rdfs:subClassOf pho:TemporalPattern .
```

The *directive* is represented in OWL in the following manner. This is a custom declaration to suit our specific application, which simply states that if a certain PHKG instance is conforming to the above *rule*, that instance would be classified under `Dietary-Guideline-01` and has an associated `pho:hasDirective Representation` that provides the python programmatic representation for the *constraints* (i.e., the lower and upper limits of the carbohydrate intake along with the daily total limit) that would be plugged into KBQA as an input.

```
pho:ConsistentCarbIntakeDirective rdf:type owl:Class ;
    owl:equivalentClass [ rdf:type owl:Restriction ;
        owl:onProperty sio:has-attribute ;
        owl:someValuesFrom pho:LowCarb] ,
        [ rdf:type owl:Restriction ;
        owl:onProperty sio:is-associated-with ;
        owl:allValuesFrom pho:Dietary-Guideline-01] ;
        rdfs:subClassOf pho:Directive ;
        pho:hasDirectiveRepresentation
        “””{‘carbohydrate’ :
        { ‘unit’: ‘g’, ‘meal’ : { ‘type’: ‘range’, ‘lower’ : ‘30’,
        ‘upper’: ‘45’}, ‘daily total’ : ‘150’}}""" ;
        rdf:label "Baseline carbohydrate level should be 30g - 45g
per carbs per meal and for the whole day 150g max.".
```

Using the available set of OWL formalizations for ADA guidelines and the PHKG, a semantic reasoner can be used to infer whether our user, i.e., *Alice*, has been adhering to behaviors consistent with the guidelines. A corresponding set of rules can be created to capture any cases of guideline violations. Then the semantic representations allow us first to identify the eventual deviation and then to provide evidence-based recommendations based on their lifestyle and diabetes condition.

Ongoing and future work is focused on expanding the set of addressable queries and integrating personal health records with the PHO. Therefore, our ongoing work includes: (1) expanding the PHO to further accommodate concepts important for comparing behaviors to ADA guidelines, (2) applying the semantic data dictionary (Rashid et al. 2020) approach to the conversion of personal health records into RDF triples that are consistent with the PHO, and (3) linking the PHKG to other semantic resources such as the Healthy LifeStyle (HeLiS) ontology (Dragoni et al. 2018).

Personalizing Dietary Recommendations

In this section, we present the personalized food recommender *pFoodReq* (Chen et al. 2021), a recommender system for answering questions that seek relevant food recommendations (e.g., “What is a Chinese dish with beef that does not include ginger?”). *pFoodReq* frames the recommender problem as that of performing knowledge base question and answering (KBQA). While recommender systems (De Croon et al. 2021) need not use semantic technologies, KBQA methods do, by definition, require a knowledge graph. Specifically, KBQA systems assume that a subset of the nodes in the knowledge graph contains answers to a general class of questions, and that the relationships between graph entities are useful for identifying good answers. Typically, questions are posed in natural language and the function of the KBQA system is to efficiently and effectively identify relevant and correct answers to the question, from the knowledge graph. Here, we will describe the specific approach used by *pFoodReq* to retrieve recipes from the FoodKG. Figure 10.4 shows the core elements of the *pFoodReq* system. At the heart of the system is a KBQA component that retrieves recipes from the FoodKG. In Fig. 10.4, the “User Dietary Needs & Preferences” could be extracted from users’

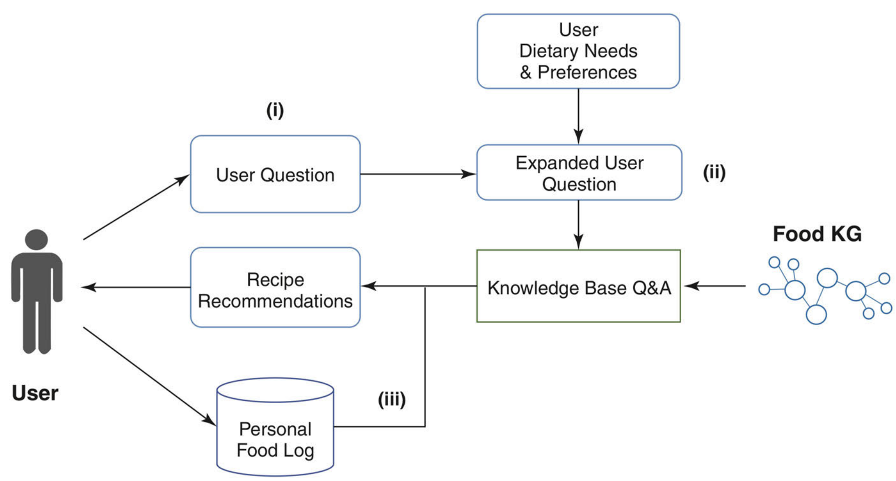

Fig. 10.4 Illustration of how a user’s question is combined with a directive regarding the user’s general dietary needs and preferences to produce an expanded question that is provided to a KBQA model, resulting in a set of recipes retrieved from the Food Knowledge Graph

PHKG by a semantic reasoner, as presented in Sect. “Populating a Personal Health Knowledge Graph with Personalized Assessments of Dietary Needs and Preferences”.

The KBQA model in *pFoodReq* has been trained to retrieve recipes from the FoodKG as answers to questions that are expressed as a combination of ‘positive’ (i.e., attributes to be included) and ‘negative’ (attributes to be excluded) constraints. Attributes that can be accommodated include recipe ingredients (e.g., mushrooms, peanuts), nutritional content (e.g., carbohydrates, fat), and cuisine/diet/dish type (e.g., Korean, vegan, dessert). Examples of these questions are: *What are jellies recipes that contain orange? What turkish or dinner-party recipes can I cook without milk? Can you recommend low protein russian recipes which have onions?* Although the user’s question (refer to (i) in Fig. 10.4) represents the user’s immediate preferences, people with diabetes also have long-term health needs. For example, according to the ADA guidelines (American Diabetes Association 2021), diabetics may need to control their caloric intake, target high fiber foods, or avoid carbohydrates with high protein content. Since any recipes recommended by *pFoodReq* would be expected to accommodate these needs, *pFoodReq* may expand the user’s question (refer to (ii) in Fig. 10.4) to include constraints related to these needs, even though the user does not include them in their question. For example, the user’s question “*What turkish or dinner- party recipes can I cook without milk?*” would be expanded by *pFoodReq* to become “*What turkish or dinner-party recipes can I cook without milk and includes carbohydrates within the desired range of 5 g to 30 g*?” In general, if a user typically avoids certain foods, these foods can also be appended to the user’s question as a negative constraint.

Rather than semantically parsing the user’s natural language questions and converting them into SPARQL queries, *pFoodReq* adopts an information retrieval approach that relies on a large training set of ‘ground-truth’ questions and answers to train a deep learning model that learns how to locate good answers to a question from the *FoodKG*. However, questions with positive and negative constraints are not easily represented in deep learning models. Hence, a new approach for using deep learning to handle these positive and negative constraints was implemented in the KBQA model in *pFoodReq*. Intuitively, the deep learning model learns associations between words in the question sentence and the corresponding answer (recipe) entity, or entities, and its nearby (recipe and non-recipe) entities and relationships in the knowledge graph. Unlike a SPARQL query (without an OPTIONAL clause), which would treat all answers satisfying the query as equally relevant, *pFoodReq*’s deep learning KBQA approach produces a continuous, scalar score for each candidate answer, allowing them to be ranked in priority of ‘relevance.’ These rankings are generated by comparing learned representations of the candidate recipe answers (Li and Zaki 2020) with recipes in the user’s historical food log and assigning higher scores to recipe answers that are more semantically similar to recipes in the food log (refer to (iii) in Fig. 10.4). The full details of the deep learning model used for KBQA in *pFoodReq* are provided in an earlier methodological paper by Chen, Wu and Zaki (Chen et al. 2019).

Summary

In this chapter, we have described how semantic technologies and machine learning can be used to bring both logic and learning to personal health applications. We have shared two use cases related to supporting dietary behaviors for people with diabetes. In our first use case, we showed how semantic technologies and data mining were used to extract, represent and reason over applicable dietary health guidelines and past user behaviors, resulting in a personal health knowledge graph that contains knowledge about a user's health preferences and needs. In our second use case, the user's health preferences and needs provide context to a user's question, allowing the recommendations to the user to consider the user's immediate and ongoing interests. There remains a significant opportunity to enhance and expand upon the ideas presented in this chapter. For example, the scope of the PHO is still limited, as are the number of ADA guidelines represented in OWL and the types of RDF triples that the TSS generates. Several dimensions of the user context could also be incorporated, such as geographical location, social context and financial constraints. Additionally, KBQA remains an active area of research from both a methodological and an application-oriented perspective.

A major advantage of using semantic models (particularly in comparison to machine learning models) is that they are inherently interpretable, and therefore amenable to providing explanations for their results. Readers are referred to works by Dragoni et al. (Dragoni et al. 2020; Dragoni et al. 2018), which present state-of-the-art applications of semantics for explainable, personalized health insights. Further, extending the work described in this chapter, we are modeling an ontology for food and diet recommendation explanations, called the Food Explanation Ontology (FEO) (Padhiar et al. 2021). FEO can be used to generate various types of explanations, such as contextual, contrastive, and counterfactual. Many of these facets can supplement the clinically relevant personal health applications in promoting effective behavior change through suitable explanations. Ideally, a personal health application would be able to provide explanations for any suggestions and recommendations, to improve their overall understanding of their health condition, the health guidelines, and their behaviors.

To keep the scope of this chapter amenable to readers new to semantic technologies, we have limited our discussion to the most salient and essential ideas and trends. There is a vast body of literature on semantic technologies, such as various reasoning techniques (higher order (Eiter et al. 2006), probabilistic (Giugno and Lukasiewicz 2002), causal (Gudivada et al. 2008)), OWL2 profiles (Motik et al. 2009), and linking data using protocols such as Linked Data Platform (Mihindukulasooriya et al. 2013), all of which are quite useful when considering the next generation personal health applications powered by semantics. The software and ontologies described in this chapter are available at `https://github.com/semantics-for-personal-health/semantics-for-personal-health.github.io`. This work was conducted as part of the

Health Empowerment by Analytics, Learning and Semantics (HEALS) project (HEALS 2017). The primary goal of the HEALS (Health Empowerment by Analytics, Learning, and Semantics) project is to apply advanced cognitive computing capabilities to help people understand and improve their own health conditions.

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

References

Abián D, Guerra F, Martínez-Romanos J, Trillo-Lado R. Wikidata and DBpedia: a comparative study. In: Szymański J, Velegrakis Y, editors. Semantic keyword-based search on structured data sources. Cham: Springer International Publishing; 2018. p. 142–54.

Allemang D, Hendler J, Gandon F. Semantic web for the working ontologist: effective modeling for linked data, RDFS, and OWL. 3rd ed. New York, NY: Association for Computing Machinery; 2020.

American Diabetes Association Standards of Medical Care in Diabetes—2021; 2021. https://doi.org/10.2337/dc21-S005.

Arenas M, Cuenca Grau B, Kharlamov E, Marciuška Š, Zheleznyakov D. Faceted search over RDF-based knowledge graphs. Journal of Web Semantics. 2016;37–38:55–74. https://www.sciencedirect.com/science/article/pii/S1570826815001432.

Baader F, Calvanese D, McGuinness D, Nardi D, Patel-Schneider P. The description logic handbook: theory, implementation and applications. 2nd ed. Cambridge University Press; 2007.

Balog K, Kenter T. Personal knowledge graphs: a research agenda. In: Proceedings of the ACM SIGIR International Conference on the Theory of Information Retrieval (ICTIR); 2019.

Berners-Lee T, Hendler J, Lassila O. The semantic web. Sci Am. 2001;284(5):34–43. http://www.jstor.org/stable/26059207.

Bipartisan Policy Center. What makes us healthy vs. what we spend on being healthy [Internet]; 2012 [cited April 15, 2021]. https://www.bipartisanpolicy.org/ report/what-makes-us-healthy-vs-what-we-spend-on-being-healthy/.

Bordes A, Weston J, Collobert R, Bengio Y. Learning structured embeddings of knowledge bases. In: Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence. AAAI'11. AAAI Press; 2011. p. 301–306.

Bouton ME. Why behavior change is difficult to sustain. Prev Med. 2014;68:29–36.

Chen X, Jia S, Xiang Y. A review: knowledge reasoning over knowledge graph. Expert Syst Appl. 2020;141:112948. https://www.sciencedirect.com/science/article/pii/S0957417419306669

Chen Y, Subburathinam A, Chen CH, Zaki MJ. Personalized food recommendation as constrained question answering over a large-scale food knowledge graph. In: Proceedings of the 14th ACM International Conference on Web Search and Data Mining. WSDM '21. New York, NY, USA: Association for Computing Machinery; 2021. p. 544–552. https://doi.org/10.1145/3437963.3441816.

Chen Y, Wu L, Zaki MJ. Bidirectional attentive memory networks for question answering over knowledge bases. In: Annual Conference of the North American Chapter of the Association for Computational Linguistics; 2019.

Cheng G, Gunaratna K, Kharlamov E. Entity summarization in knowledge graphs: algorithms, evaluation, and applications. In: Companion Proceedings of the Web Conference 2020. WWW '20. New York, NY, USA: Association for Computing Machinery; 2020. p. 301–302: https://doi.org/10.1145/3366424.3383108.

De Croon R, Van Houdt L, Htun NN, Štiglic G, Vanden Abeele V, Verbert K. Health Recommender systems: systematic review. J Med Internet Res. 2021;23(6):e18035. https://www.jmir.org/2021/6/e18035

Dooley D, Griffiths E, Gosal G, Buttigieg P, Hoehndorf R, Lange M, et al. FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration. NPJ Sci Food. 2018;2:23.

Dragoni M and Bailoni T et al. Helis: An ontology for supporting healthy lifestyles. In: ISWC. Springer; 2018. p. 53–69.

Dragoni M, Donadello I, Eccher C. Explainable AI meets persuasiveness: translating reasoning results into behavioral change advice. Artif Intell Med. 2020;105:101840. https://www.sciencedirect.com/science/article/pii/S0933365719310140.

Eiter T, Ianni G, Schindlauer R, Tompits H. Effective integration of declarative rules with external evaluations for semantic-web reasoning. In: European Semantic Web Conference. Springer; 2006. p. 273–287.

Feyznia A, Kahani M, Zarrinkalam F. COLINA: A Method for Ranking SPARQL Query Results through Content and Link Analysis. In: Proceedings of the 2014 International Conference on Posters & Demonstrations Track – Volume 1272. ISWC-PD'14. Aachen, DEU: CEUR-WS.org; 2014. p. 273–276.

Fu B, Qiu Y, Tang C, Li Y, Yu H, Sun J. A survey on complex question answering over knowledge base: Recent advances and challenges. arXiv preprint arXiv:2007.13069. 2020 Jul 26.

Giugno R, Lukasiewicz T. P-SHOQ (D): a probabilistic extension of SHOQ (D) for probabilistic ontologies in the semantic web. In: JELIA. vol. 2. Springer; 2002. p. 86–97.

Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. npj Dig Med. 2020;3:14.

Gudivada RC, Qu XA, Chen J, Jegga AG, Neumann EK, Aronow BJ. Identifying disease- causal genes using semantic web-based representation of integrated genomic and phenomic knowledge. J Biomed Inform. 2008;41(5):717–29.

Gyrard A, Gaur M, Thirunarayan K, Sheth AP, Shekarpour S. Personalized Health Knowledge Graph. In: CKGSemStats@ISWC; 2018.

Harris JJ, Chen CH, Zaki MJ. A framework for generating explanations from temporal personal health data. ACM Trans. Comput. Healthcare. 2021;02(21):1–43.

Haussmann S, Chen Y, Seneviratne O, Rastogi N, Codella J, Chen CH, et al. FoodKG enabled Q&A application. In: ISWC Satellites; 2019. p. 273–276.

Haytowitz DB, Ahuja JKC, Wu X, Somanchi M, Nickle M, Nguyen QA, et al. USDA National Nutrient Database for Standard Reference, Legacy Release [Data File]; 2019. [Cited 02 Apr 2021]. Nutrient Data Laboratory, Beltsville Human Nutrition Research Center, ARS, USDA. https://data.nal.usda.gov/dataset/usda-national-nutrient-database-standard-reference-legacy-release.

He Z, Xu X, Deng S. Squeezer: an efficient algorithm for clustering categorical data. J Comput Sci Technol. 2002;09(17):611–24.

HEALS: Health empowerment by analytics, learning, and semantics [Internet]; 2017 [updated April 2021; cited April 15, 2021]. https://idea.rpi.edu/research/projects/heals.

Hendler J. Science and the semantic web. Science. 2003;299(5606):520–1. https://science.sciencemag.org/content/299/5606/520.

Hogan A, Blomqvist E, Cochez M, d'Amato C, de Melo G, Gutierrez C, et al. Knowledge graphs. ACM Comput. Surv. 54(4), Article 71 (May 2022), 37 pages. https://doi.org/10.1145/3447772.

Kacprzyk J, Yager RR, Zadrozny S. In: Abramowicz W, Zurada J, editors. Fuzzy Linguistic summaries of databases for an efficient business data analysis and decision support. Boston, MA: Springer US; 2002. p. 129–52. https://doi.org/10.1007/0-306-46991-X_6.

Kelly MP, Barker M. Why is changing health-related behaviour so difficult? Public Health. 2016;136:109–16. https://www.sciencedirect.com/science/article/pii/S0033350616300178.

Kendall E, McGuinness D. Ontology engineering. In: Synthesis Lectures on The Semantic Web: Theory Technology. vol. 9; 2019. p. i–102.

Li D, Zaki MJ. RECIPTOR: an effective pretrained model for recipe representation learning. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. KDD '20. New York, NY, USA: Association for Computing Machinery; 2020. p. 1719–1727. https://doi.org/10.1145/3394486.3403223.

Lin J, Keogh JE, Wei L, Lonardi S. Experiencing SAX: a novel symbolic representation of time series. Data Min Knowl Disc. 2007;08(15):107–44.
Liu Q, Cheng G, Gunaratna K, Qu Y. Entity summarization: state of the art and future challenges. Journal of Web Semantics. vol. 69; May 2021.
Marín J, Biswas A, Ofli F, Hynes N, Salvador A, Aytar Y, et al. Recipe1M+: a dataset for learning cross-modal embeddings for cooking recipes and food images. IEEE Trans Pattern Anal Mach Intell. 2021;43(1):187–203.
Michie S, Thomas J, Johnston M, Aonghusa P, Shawe-Taylor J, Kelly M, et al. The human behaviour-change project: harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation. Implement Sci. 2017;12(1):121.
Mihindukulasooriya N, Garcia-Castro R, Gutiérrez ME. Linked Data Platform as a novel approach for Enterprise Application Integration. In: COLD; 2013.
Moschitti A, Tymoshenko K, Alexopoulos P, Walker A, Nicosia M, Vetere G, et al. In: Pan JZ, Vetere G, Gomez-Perez JM, Wu H, editors. Question answering and knowledge graphs. Cham: Springer International Publishing; 2017. p. 181–212. https://doi.org/10.1007/978-3-319-45654-6_7.
Motik B, Grau BC, Horrocks I, Wu Z, Fokoue A, Lutz C, et al. OWL 2 web ontology language profiles. W3C Recommendation. 2009;27:61.
National Diabetes Statistics Report. Atlanta, GA; 2020.
Padhiar I, Seneviratne O, Chari S, Gruen D, McGuinness DL. Semantic modeling for food recommendation explanations; 2021. https://arxiv.org/abs/2105.01269.
Pan JZ, Vetere G, Gomez-Perez JM, Wu H, editors. Exploiting linked data and knowledge graphs in large organisations. Springer International Publishing; 2017.
Pillai SG, Soon LK, Haw SC. Comparing DBpedia, Wikidata, and YAGO for web information retrieval. In: Piuri V, Balas VE, Borah S, Syed Ahmad SS, editors. Intelligent and interactive computing. Singapore: Springer Singapore; 2019. p. 525–35.
Puustjarvi J, Puustjarvi L. Personal health ontology: towards the interoperation of e-health tools. Int J Electron Healthc. 2011;6(1):62–75.
Rashid SM, McCusker JP, Pinheiro P, Bax MP, Santos HO, Stingone JA, et al. The semantic data dictionary–an approach for describing and annotating data. Data Intelligence. 2020;2(4):443–86.
Rastogi N, Zaki MJ. Personal health knowledge graphs for patients. arXiv:2004.00071; 2020. https://doi.org/10.48550/arxiv.2004.00071.
Ringler D, Paulheim H. One knowledge graph to rule them all? Analyzing the differences between DBpedia, YAGO, Wikidata & co. In: Kern-Isberner G, Fürnkranz J, Thimm M, editors. KI 2017: advances in artificial intelligence. Cham: Springer International Publishing; 2017. p. 366–72.
Shirai S, Seneviratne O, McGuinness DL. Applying personal knowledge graphs to health. arXiv:2104.07587; 2021. https://doi.org/10.48550/arxiv.2104.07587.
Suchanek F, Fan J, Hoffmann R, Riedel S, Talukdar PP. Advances in automated knowledge base construction. SIGMOD Records journal, March. 2013a.
Suchanek FM, Riedel S, Singh S, Talukdar PP. AKBC 2013: third workshop on automated knowledge base construction. In: Proceedings of the 22nd ACM international conference on Information & Knowledge Management; 2013b. p. 2539–2540.
Thomas Craig KJ, Morgan LC, Chen CH, Michie S, Fusco N, Snowdon JL, et al. Systematic review of context-aware digital behavior change interventions to improve health. Transl Behav Med. 2020:10. https://doi.org/10.1093/tbm/ibaa099.
Zadeh LA. The concept of a linguistic variable and its application to approximate reasoning–I. Inf Sci. 1975;8(3):199–249.
Zadeh LA. A computational approach to fuzzy quantifiers in natural languages. Comput Math Appl. 1983;9(1):149–84.
Zadeh LA. A prototype-centered approach to adding deduction capability to search engines-the concept of protoform. In: IEEE Symposium on Intelligent Systems; 2002.
Zaki MJ. SPADE: an efficient algorithm for mining frequent sequences. Mach Learn. 2001;42(1):31–60.