FoodKG Enabled Q&A Application *

Steven Haussmann[1[0000−0003−0216−1820]], Yu Chen[1[0000−0003−0966−8026]], Oshani Seneviratne[1[0000−0001−8518−917X]], Nidhi Rastogi[1[0000−0002−2002−3213]], James Codella[2[0000−0002−3430−0429]], Ching-Hua Chen[2[0000−0002−1020−0861]], Deborah L. McGuinness[1[0000−0001−7037−4567]], and Mohammed J. Zaki[1[0000−0003−4711−0234]]

[1] Rensselaer Polytechnic Institute, Troy NY, USA
[2] IBM Research, USA

Abstract. We demonstrate the usage of our FoodKG [3], a food knowledge graph designed to assist in food recommendation. This resource, which brings together recipes, nutrition, food taxonomies, and links into existing ontologies, is used to power a cognitive agent that performs knowledge-base question answering, primarily to help improve peoples' diets by guiding them towards better foods. The system demonstration involves three categories of questions: simple queries for nutritional information, comparisons of nutrients between different foods, and constraint-based queries to find recipes matching certain criteria.

1 Introduction

In our resource track paper titled 'FoodKG: A Semantics-Driven Knowledge Graph for Food Recommendation' [3], we described the basis for a food knowledge graph (FoodKG) that brings together recipes, food ontologies, and nutritional data. The FoodKG comprises approximately 7.7 thousand nutrient records (sourced from the USDA), one million recipes (sourced from [4]), and 7.3 thousand classes of food (sourced from [2]). Overall, it has over 67 million triples. Full details related to the motivation, composition, and construction of the FoodKG can be found in the resource track paper [3].

For the sake of brevity, we did not fully describe potential and current applications of our Food knowledge graph. In our demo, as described in this paper, we examine how a knowledge graph can be leveraged by a question-answering system to answer queries related to recipes and nutrition.

2 Knowledge-Base Question Answering

In our demo, we will demonstrate the use of our FoodKG for answering natural language questions - a process known as *knowledge base question answering* (KBQA). The system takes natural language questions, such as "What Indian dishes can I make with chicken and garlic?", and generates an answer based on information stored in the FoodKG.

* Copyright ©2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

We seek to answer three types of questions: *simple*, *comparison*, and *constraint*. The simple questions, e.g., "How much sugar is in cheese, cream, fat free?", are created based on the USDA data and require only one hop reasoning. The comparison question, e.g., "Sesame oil or peanut oil, which has less saturated fat?", can be regarded as a composition of two simple questions, followed by a quantitative comparison on some attribute. Last, we also create questions with constraints, e.g., "What Laotian dishes can I make with sugar, water, oranges?". Table 1 summarizes the types of questions.

Questions	Question Template Examples
Simple	How much {nutrient} is in {ingredient}?
Comparison	{ingredient1} or {ingredient2}, which has less {nutrient}?
Constraint	What {tag} dishes can I make with {ingredient_list}?

Table 1. Examples of the types of questions handled by the Q&A system

2.1 Q&A over the FoodKG

Our KBQA system consists of three components: the *question type classifier*, the *topic entity predictor*, and the *KBQA model*. Given a natural language question, such as "how much sugar is in Cheese, Blue?", the *question type classifier* determines the question type. In this case, the question is *simple*. Then, the *topic entity predictor* is applied to detect the topic entity mentioned in the question and link it to the entity 'Cheese, Blue' in the FoodKG. Finally, the *KBQA model* is called to retrieve answers from the KG subgraph (which we assume contains all the candidate answers) surrounding the topic entity. For our implementation, we deploy our state-of-the-art neural network-based KBQA model called BAMnet [1].

2.2 Towards Personalized Food Recommendation

In order to provide personalized food recommendation to users, we are working to extend our KBQA system to be aware of user preferences. Our current strategy is to treat preferences as additional, implicit constraints. This is done by maintaining a personal KG for user preferences, such as liked/disliked ingredients or nutritional requirements.

Given a query, such as "What are brewing recipes that consist of sugar?", we transform the query into one that incorporates the user's preferences. For a user who is allergic to 'mint leaves', this could result in a query of "What are brewing recipes that consist of sugar, and do not have mint leaves?" A regular KBQA system that is not aware of user preferences will recommend 'Kahlua Popsicles', 'Jasmine, Green Tea, Fresh Lemonade', 'Lemon Liqueur' and 'Tequila Lime Punch'. However, our personalized KBQA system will instead recommend 'Kahlua Popsicles', 'Lemon Liqueur' and 'Tequila Lime Punch', since 'Jasmine, Green Tea, Fresh Lemonade' contains 'mint leaves'. In this way, our system is able to provide different users with tailored food suggestions that suit their needs. Future work in this direction includes how to combine personal KG with FoodKG and how to effectively handle negation in a query.

3 Demonstration

3.1 Design

The KBQA system provides answers to user questions; on top of this, we have built a web application to receive questions and present the results. This is summarized in Figure 1. Boldface text refers to specific components within the figure.

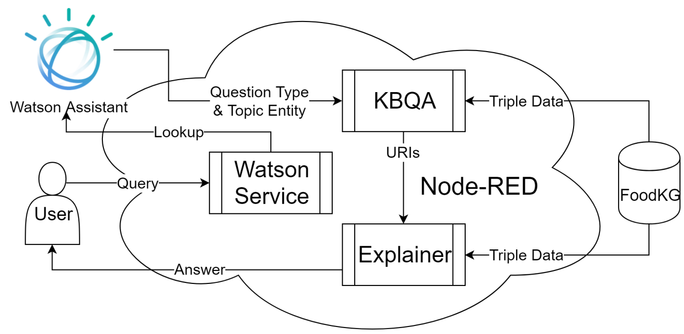

Fig. 1. The architecture of the food information system.

We chose to build a web application in **Node-RED**[3], a flow-based tool to link together logic and services. This made it straightforward to combine the aforementioned **KBQA** system with a simple frontend, along with the necessary services to perform entity recognition display results.

To identify the topic entity and question type of a query, we use IBM's **Watson Assistant**[4] service. This information, along with the original query, is passed to the KBQA service.

Since the answers to the user's queries exist directly in the **FoodKG**, we chose to have the KBQA system respond not with a plain-text response (e.g., "250 mg sodium"), but rather with a URI (e.g., http://idea.rpi.edu/heals/kb/usda-#sodium-01004). This direct reference to the KG is then handed to another service, the **Explainer**. The Explainer service uses a SPARQL query to retrieve data from the FoodKG, and then uses the result of the query to produce human-readable text. Each type of resource – measurement, recipe, nutritional entry – has its own rule. This separation of concerns makes it simpler to extend and modify the Q&A system.

3.2 Examples

To demonstrate the usefulness of the FoodKG, we plan to present the usage of the above system to answer three types of questions. Examples of such questions and their answers are shown in Figure 2.

As noted, we are presently working to incorporate personalization into the question-answering pipeline, leading to markedly different answers to the same

[3] https://nodered.org
[4] https://www.ibm.com/cloud/watson-assistant/

question for different users - depending on personal fitness and preferences. These implicit additions to the question enrich the meaning of each query without requiring a fundamental shift in how the queries are carried out.

You : How much sodium is in butter, salted?
Food Recommender : 643 milligrams Sodium

(a) Simple

You : Which has more carbohydrates, butter, salted or cheese, blue?
Food Recommender : Cheese, blue

(b) Comparison

You : What are some Indian dishes with chicken?
Food Recommender :
- Chicken Tikka Masala (Link)
- Curried Chicken (Link)
- Quick Chicken and Vegetable Indian Curry (Link)

(c) Constraint

Fig. 2. Types of questions

4 Conclusions

There is a vast amount of disconnected food knowledge, which spans an entire spectrum of food topics, from nutrients to recipes. For an individual, numerous biological, psychological, social and environmental factors influence food choices. Therefore, navigating this vast space on food requires meaningful linkages. Furthermore, the provenance of the food facts is equally important in order to trust the data, and eventually, the food that is consumed by an individual. Our unified knowledge graph on food, the FoodKG, has these characteristics, and in our demo, we will illustrate the utility of this resource via a Q&A system.

References

1. Chen, Y., Wu, L., Zaki, M.J.: 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)
2. Griffiths, E.J., Dooley, D.M., Buttigieg, P.L., Hoehndorf, R., Brinkman, F.S., Hsiao, W.W.: FoodON: A Global Farm-to-Fork Food Ontology. In: ICBO/BioCreative (2016)
3. Haussmann, Steven and Seneviratne, Oshani and Chen, Yu and Ne'eman, Yarden and Codella, James and Chen, Ching-Hua and McGuinnes, Deborah and Zaki, Mohammed: FoodKG: A Semantics-Driven Knowledge Graph for Food Recommendation. In: International Semantic Web Conference (2019)
4. Marin, J., Biswas, A., Ofli, F., Hynes, N., Salvador, A., Aytar, Y., Weber, I., Torralba, A.: Recipe1m: A dataset for learning cross-modal embeddings for cooking recipes and food images. arXiv preprint arXiv:1810.06553 (2018)