Towards Adaptive and Personalised Recommendation for Healthy Food Promotion

Diana Nurbakova, Felix Bölz, Audrey Serna, Jean Brignone

To cite this version:

Diana Nurbakova, Felix Bölz, Audrey Serna, Jean Brignone. Towards Adaptive and Personalised Recommendation for Healthy Food Promotion. 2023. hal-04261745

HAL Id: hal-04261745
https://hal.univ-lorraine.fr/hal-04261745
Submitted on 27 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Towards Adaptive and Personalised Recommendation for Healthy Food Promotion

Diana Nurbakova\textsuperscript{1,}\textsuperscript{*}, Felix Bölz\textsuperscript{1,2}, Audrey Serna\textsuperscript{1} and Jean Brignone\textsuperscript{1}

\textsuperscript{1}Univ Lyon, INSA Lyon, CNRS, UCBL, LIRIS, UMR5205, Villeurbanne, F-69621, France
\textsuperscript{2}University of Passau, Passau, 94032, Germany

Abstract
This paper presents ongoing work on an adaptive persuasive system to promote healthy eating habits. It exploits and extends the idea of a constrained question answering (QA) system over a knowledge graph proposed by Chen \textit{et al.} [1]. In particular, we introduce the way to model personalised challenges, a key component of gamified behaviour change techniques, and additional constraints allowing to handle meal plans by keeping a track on the distribution of daily intakes across meals, repetitive recommendations, constraints related to nutritional labels of the recipes. To access rich nutrition and user-item interaction data, we use HUMMUS [2] instead of FoodKG [3].

Keywords
healthy food recommendation, behaviour change, constrained recommendation, knowledge graph

1. Introduction
Healthy food recommendation is a crucial domain in the recommender system field as malnutrition has become a global issue [4]. However, an efficient healthy food promotion system goes beyond a traditional recommender as it requires rich nutrition and health data from both, user and item sides, but most importantly, as it implies a behaviour change. Incorporating the latter into recommendation process is a challenging task. The existing works mainly focus on only one aspect. In [5], we have presented our general idea of adaptive and privacy-preserving strategies for healthy food promotion. In this paper, we focus on the recommendation of healthy recipes tailored to the user’s personal constraints and behaviour change stage. To make use of rich food data, a knowledge graph can be used (see Figure 2 for an example).

Our main research question can be formulated as follows: \textit{How can we model adaptive persuasive healthy food recommendation?} Thus, we define the addressed problem as follows: Given user-recipe interaction data, user profile (including health profile and dietary preferences), healthy food guidelines, current challenges for this user or other behaviour change techniques, and rich recipe data, extract the recipes from the knowledge graph that satisfy the set of input requirements. A result of the recommendation can be a single meal or a meal plan.
A recent work of Chen et al. [1] has proposed pFoodReQ, a constrained question answering (QA) system over a large food knowledge graph, FoodKG [3]. It introduces a set of constraints allowing to model user’s dietary preferences (likes/dislikes and allergies), health guidelines and explicit requirements from a user query formulated in a natural language. These constraints are then used to extract recipes from FoodKG using a knowledge based question answering system (KBQA). Our contribution can be summarised as follows. We extend the idea of such a constrained QA system for healthy food promotion by (a) modelling challenges, a key component of gamified approaches to behaviour change, as additional constraints; (b) introducing new types of constraints (based on Nutri-score, daily intakes, repetitive recommendations); (c) creating automatic queries for the ease of integration into a mobile application; (d) switching from FoodKG to HUMMUS [2] which allows to access a richer nutrition and user-item interactions (ratings, reviews) data. The latter also enables to further enhance our proposed approach by adding collaborative filtering techniques. In this paper, we present our ongoing work on this modelling.

2. Related Work

The recommendation of healthy food has been attracting the attention of the research and industrial communities for a while. However, it remains much less popular than recommendations for tourism, movies, etc. or simple food recommendation, i.e. without consideration of the health dimension (e.g. [6, 7, 8, 9]). In our opinion, this is due to the absence of health-related data, especially user data. Earlier works on healthy food recommendation focus mainly on calorie intake (e.g. [10, 11]). In contrast to that, Toledo et al. [12] suggest a multi-criteria approach that takes into account the levels of protein, sodium, cholesterol, and saturated fats, as well as user preferences. With the spread of food traffic light systems, in particular on packaged food, such an idea was adopted in healthy food recommendation ([13, 14, 15]). However, such a single score of healthiness is quite limited as it expresses generalised information without any personalisation and adaptation. Thus, Chen et al. [1] stress out that most of the existing recommenders are unable to take into account such valuable information as user’s allergies and nutrition needs, and do not make use of rich food knowledge available through semantic data.

They propose pFoodReQ, a constraint-based question answering (QA) system reasoning over FoodKG [3], a large-scale food knowledge graph. A set of constraints has been introduced to handle in a unified way three main aspects: user query, user’s dietary preferences (likes and allergies), and health guidelines (nutrition needs). Personalisation is attained via query expansion. More precisely, user’s dietary preferences and health guidelines extend the initial query. The use of nutrient and micronutrient budgets can be adjusted to users’ needs. pFoodReQ is rather independent form KBQA system operating over a knowledge graph, but for the experiments the authors use BAMnet [16]. In terms of recommendation task, pFoodReQ can be seen as a content-based approach. As user-recipe interactions (ratings, reviews, likes, etc.) are missing in FoodKG, the authors simulate such interactions by adding dishes log to a user’s query and exploiting similar recipes. Recipe popularity, ratings or collaborative information has not been explored. Moreover, all constraints should be satisfied and can therefore, be seen as hard constraints which can be limiting in real-world scenario, e.g. not all ingredients among available ones should necessarily be present in the recommended recipe.
3. Methodology

Our solution lies on pFoodReq [1]. In this work, we extend pFoodReq in several ways:

1. Creation of automatic queries based on user profile (Section 3.1).
2. Switch from FoodKG to HUMMUS in terms of a knowledge graph (Section 3.2).
3. Additional constraints (Section 3.3).
4. Incorporation of challenges and behaviour change strategies into queries (Section 3.4).

Figure 1 provides a general overview of our method. We detail each point in what follows.

3.1. Automatic Query based on User Profile

pFoodReq [1] is a QA system that provides a list of recipes as a response to a query expressed in natural language containing explicit requirements (e.g. type of cuisine, type of dish, ingredients to use or exclude, etc.). pFoodReq provides personalised answers based on a user query. It does not handle user profiles in a proper sense but is generated, even though in Chen et al. notation persona data is used for personalisation, i.e. allergies, ingredients likes/dislikes, particular health guidelines. Interaction data is introduced as food logs, containing a list of past recipes. To simulate user queries, Chen et al. have defined multiple query templates allowing to generate such queries. Our system has rich user data. In a typical scenario, the meal plans or single meals are suggested to a user without explicit query. To do so, we use the idea of templates that a system will fill in automatically based on user information, past behaviour and goals. However, a possibility of explicit search for recipes will be enabled.

3.2. From FoodKG to HUMMUS

We use the HUMMUS dataset [2] which extends FoodKG [3]. The motivation behind that is two-fold: (1) the presence of recipe healthiness scores in HUMMUS, namely Nutri-score [17, 18],
WHO [4] and FSA [19, 20, 21]; (2) the availability of user-item (recipe) interactions missing in the original FoodKG. Moreover, from the technical perspective, we consider that the use of SPARQL queries can be beneficial to reason over a knowledge graph. Refer to Figure 2 to get an example of the graphs’ structure.

3.3. Additional Constraints

pFoodReq takes into account several kinds of constraints:

1. User query based (generated based on templates): positive ingredient constraints, negative ingredient constraints, nutrient based constraints, cuisine based constraints.
2. Dietary preferences (generated randomly): ingredient likes/dislikes, ingredient allergies.

Note that multiple guidelines per user are enabled by pFoodReq. In our work, we extend these constraints. In particular, we describe how to take into account recipe Nutri-score value, daily budgets, repetitive recommendations, and challenges for users to undergo.

Nutri-score Nutri-score [17, 18] is a 5-colour nutrition label (A, B, C, D, E) rating the overall nutritional value of food products (see Figure 3). It was originally used for packaged food similar to the FSA score [19, 20], another front-of-package nutrition label. But nowadays, thanks to its simplicity and easy visual interpretation (sort of a traffic-light system where green light means a healthy option, amber/yellow means medium, and red means bad for health), it has become a

**Figure 2**: Example view of a recipe in HUMMUS using GraphDB.
Figure 3: Nutri-score labels from A (the best) to E (the worst), image source: [17]

rather common way to ‘measure’ the food healthiness in Europe (and in France, in particular). Moreover, in a recent work [23], El Majjodi et al. have shown that the presence of nutrition labels in the personalised recommenders decrease the perceived choice difficulty among users. For those reasons, we believe that incorporating the information about recipe Nutri-score in our system can be beneficial. It can be done in several levels:

- **recipe visualisation**: the information about recipe Nutri-score is displayed to a user;
- **challenge**: a challenge to eat dishes rated at least B for a week or another time period can be proposed to a user (see Listing 1 for a SPARQL query extracting all recipes rated at least B in terms of their Nutri-score);
- **goal setting**: a user may set their goal in terms of Nutri-score;
- **general health guidelines**: the value of the Nutri-score of the recipes can be used as a general healthy nutrition guidelines.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX foodkg: <http://idea.rpi.edu/heals/kb/>
PREFIX hummus: <https://www.hummus.uni-passau.de/>

SELECT DISTINCT ?recipe ?title WHERE { ?recipe a foodkg:recipe; rdfs:label ?title; hummus:has_nutri_score ?nutri_score. FILTER (?nutri_score = 'A' || ?nutri_score = 'B' || ?nutri_score = 'C') # Nutri score range: Between A and E, with A as best. }
```

Listing 1: SPARQL query to extract recipes of at least nutri-score category C.

**Daily intake** In terms of nutrients and micronutrients, the health guidelines of pFoodReq contain daily total. However, the work does not detail how this daily budget is used. Consider the following example from pFoodReq code (file data_builder/src/config/data_config.py):

```python
{
    'protein':
        {
            'unit': 'g',
            'meal':
                {
                    'type': 'range',
                    'lower': '15',
                    'upper': '40'},
            'daily total': '60'}
    }
```
Table 1
Distribution of calories across eating occasions for adults (19-59 y.o.) in target calorie level (# kcal) and percentage of Total Calories (% TCal) [25]. Note: kcal = calories

<table>
<thead>
<tr>
<th>Total Calories (kcal)</th>
<th>Breakfast</th>
<th>Lunch</th>
<th>Dinner</th>
<th>Snack 1</th>
<th>Snack 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># kcal</td>
<td>% TCal</td>
<td># kcal</td>
<td>% TCal</td>
<td># kcal</td>
</tr>
<tr>
<td>2000</td>
<td>440</td>
<td>22</td>
<td>640</td>
<td>32</td>
<td>640</td>
</tr>
</tbody>
</table>

The left guideline suggests that a good protein intake per meal should be between 15g and 40g and should not exceed the daily total of 60g. The right guidelines suggest that proteins should constitute between 10% and 25% of total calories. The multiplier value means that one gram of protein has 4 calories (e.g. [24]). The additional constraints are then added.

In our work, we go further. As we store user’s historical data, including their food log, our system can suggest meal plans based on the distribution of calorie intake across meals, and adjust the recommendation based on the consumed meal (e.g. if a person eats less for breakfast, this overhead can be redistributed to the lunch and/or dinner budget). Thus, the US Institute of Medicine of the National Academies [25] has proposed a distribution of calories across meals for the adults, 19-59 years old (see Table 3.3). Based on that, we can introduce new constraints for each meal (eating occasion). For the main three meals, the constraints can be given as:

```json
{ 'breakfast': { 'percentage': 'calories', 'meal': 22 },
  'lunch': { 'percentage': 'calories', 'meal': 32 },
  'dinner': { 'percentage': 'calories', 'meal': 32 }
}
```

Note that the total amount of calories is calculated based on the user profile (i.e. weight, height, physical activity, goals, etc.).

Recipe Repetition and Interaction History When it comes to food, the dishes we eat usually repeat. So, in contrast to traditional recommendation, where a system recommends to a user only unseen items, healthy food recommendation task relaxes this constraint. Thus, the same item even with which a user interacted in the past can be recommended again. However, one should avoid returning the same items over and over again. To do so, different constraints can be applied to an item. They can take different forms:

- a time budget defining a number of days between re-recommending it, e.g.:

```json
{ 'rec_recipe':
  { 'unit': 'days', 'recommended': 7, 'followed': 14 },
}
```

- a penalty or a decay coefficient;
- a maximal number of repetitions.
This is possible as our system allows to track the consumed and already recommended items thanks to our recipe repetition and interaction data module.

### 3.4. Incorporating Challenges

Challenges constitute an important part of a gamified approach to behaviour change. Thus, Oyebode et al. [26] point out several design recommendations such as tailoring the content and allowing users to set their own goals, offering suggestions on both how to set effective goals and how to reach the goals. In our system, we will propose several challenges aiming at promoting and developing healthy eating habits. To fuse them with recommendation, we suggest to model them as additional constraints used in the recommendation process.

There is a common ‘5-a-day’ recommendation, according to which one should eat 5 fruits and vegetables per day. It originates from the WHO’s healthy diet guide [4]. More precisely, the latter states that a healthy diet for an adult includes:

“**At least 400 g (i.e. five portions) of fruit and vegetables per day** [27], excluding potatoes, sweet potatoes, cassava, and other starchy roots.”

Godman [28] provides a table of the correspondence of the number of fruits and vegetables to what is considered to be “one serving”, e.g.: apricots (1 fresh, 1/2 cup canned, or 5 dried), carrots (1/2 cup cooked, 1/2 raw carrot, or 2–4 sticks). The more beneficial in terms of health are fruits and vegetables rich in vitamin C and beta carotene (e.g. carrots, citrus, berries) and leafy green vegetables (e.g. spinach, kale). According to the British Dietetic Association (BDA) [29], different forms of fruits and vegetables count in a 5-a-day plan, i.e. fresh (raw), frozen, dried, and canned. However, a recent study [30] has shown in an example of potatoes that starch breakdown and release of sugars in the body depend significantly on the cooking method.

Thanks to the rich semantic data contained in the knowledge graph, we can select fruits and vegetables to include as a constraint. We can also exclude starchy vegetables, keeping only the “beneficial” vegetable list. Listing 2 provides an example of a SPARQL query that retrieves all recipes containing at least two non-starchy vegetables. Lines 12-25 correspond to filters of starchy vegetables (e.g. cassava, plantain, potatoes, turnip, etc.).

### 4. Conclusion

We have presented a model allowing to incorporate gamified challenges and additional constraints (recipe Nutri-score, distribution of calories across meals, recipe repetition) into existing KBQA framework called pFoodReq [1] over HUMMUS [2] knowledge graph. The queries are generated automatically based on rich user profile available via a mobile application.

The current modelling suffers from some limitations though. First, as a recommendation item is a recipe, the system cannot recommend just an apple or any other fruit as a snack as it is not a recipe (dish). To overcome this, it is possible to introduce a new tag, e.g. *snack*, and assign it to fruits, making it possible to use them in recommendation. Second, now, the constraints should be satisfied in a rather hard manner. A means to soften the constraints should be considered.
Moreover, currently, we do not make full use of the user-item interaction data available in HUMMUS. We leave these limitations to our future work.

In addition, we identify the following future directions of our work. First, we will evaluate the described model in two stages: (1) with simulated data (similar to the setting of Chen et al. [1]) and (2) with real users data via our mobile application which is under the development. Then, we will focus more on the recommendation of meal plans. We will also explore the use of photos of available ingredients (food products) as a soft constraint on proposed recipes.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX foodkg: <http://idea.rpi.edu/heals/kb/>
PREFIX foodon: <http://purl.obolibrary.org/obo/FOODON_>

SELECT DISTINCT ?recipe (COUNT(DISTINCT ?food_on_ingredient) AS ?n_ing)
WHERE {
  ?recipe a foodkg:recipe;
    rdfs:label ?title;
    foodkg:uses/owl:equivalentClass/rdfs:subClassOf ?food_on_ingredient.

  FILTER EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00001261.} #vegetables
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:03411016.} #starch-producing plant
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00001692.} #cassava
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00002159.} #plantain (musa)
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00002368.} #yam
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00002310.} #taro
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00002297.} #sweet potato
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00001142.} #maize (corn)
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00001148.} #potato
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:03317199.} #green pea (whole)
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00002981.} #butternut squash (whole, raw)
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:03303946.} #butternut squash (frozen)
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00003599.} #acorn squash (whole, raw)
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:03315223.} #acorn squash (whole, raw)
  FILTER NOT EXISTS {?food_on_ingredient rdfs:subClassOf* foodon:00002324.} #turnip
}
GROUP BY ?recipe HAVING (?n_ing >= 2)
```

Listing 2: SPARQL query to extract recipes containing at least 2 non-starchy vegetables. This query returns 290656 results.
References


[12] R. Yera Toledo, A. A. Alzahrani, L. Martinez, A Food Recommender System Considering


[29] BDA, Fruit and vegetables - how to get five a day, 2020. URL: https://www.bda.uk.com/resource/fruit-and-vegetables-how-to-get-five-a-day.html.