Ontology-Based Nutritional Recommender System

Dexon Mckensy-Sambola 1, Miguel Ángel Rodríguez-García 2*, Francisco García-Sánchez 3 and Rafael Valencia-García 3

1 Escuela de Informática, Bluefields Indian & Caribbean University (BICU), Bluefields 81000, Nicaragua; dexon.sambola@bicu.edu.ni
2 Departamento de Ciencias de la Computación, Universidad Rey Juan Carlos, 28933 Madrid, Spain
3 Departamento de Informática y Sistemas, Universidad de Murcia, 30100 Murcia, Spain; fgracia@um.es (F.G.-S.); valencia@um.es (R.V.-G.)
* Correspondence: miguel.rodriguez@urjc.es

Abstract: Obesity is considered an epidemic that is continuously growing around the world. Heart diseases, diabetes, and bone and joint diseases are some of the diseases that people who are overweight or obese can develop. One of the vital causes of those disorders is poor nutrition education; there is no raising awareness about eating healthy food and practicing physical activities to burn off the excess energy. Therefore, it is necessary to use new technologies to build methods/tools that help people to overcome these avoidable nutrition disorders. For this reason, we implemented a recommendation engine capable of identifying the different levels of overweight and obesity in users and providing dietary strategies to mitigate them. To do so, we defined the Ontology of Dietary Recommendations (ODR) with axioms to model recipes, ingredients, and a set of diets to assist people who suffer from obesity. We validated the defined model by using a real set of individuals who were anonymized. A panel of advisors evaluated each individual record and suggested the most appropriate diets from those included in the ontology. Then, the proposed system was asked to also provide diet recommendations for each individual, which were compared with those proposed by the advisors (ground truth), reaching a mean accuracy of 87%.

Keywords: diet recommendation; body mass index; optimal nutrition; ontology; knowledge-based systems

1. Introduction

Currently, obesity is considered a “global, complex, multifactorial, and generally preventable disease” [1,2]. According to the World Health Organization (WHO) (https://www.who.int/health-topics/obesity (accessed on 17 December 2021)), overweight and obesity are associated with the abnormal accumulation of fat, which can be potentially harmful. This feeding disorder can be developed by having poor eating habits and an unhealthy lifestyle, and it has an even worse impact on human health, reducing immunity, increasing vulnerability to illnesses, altering physical and sensory development, and reducing productivity [3]. Several diagnostic measures have been employed to assess obesity [4]. Although it has some limitations, the body mass index (BMI) is the most widely used index because of its simplicity and low cost. It measures in kg/m² the abnormal accumulation of fat, considering anthropometric parameters such as weight and height. Other more accurate techniques, such as dual-energy X-ray absorptiometry (DEXA), are more complicated and expensive. For instance, a person is considered as an obese person if his/her BMI is higher than or equal to 30 kg/m², taking into account the table published by the WHO (https://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi (accessed on 17 December 2021)), which establishes different obesity categories depending on the level of the BMI reached.

As any other widespread disease, obesity is considered as an epidemic that can affect people of all ages [5]. The self-management of physical activities and dietary habits
are the primary components not only to have a healthy life by maintaining the ideal weight of the body, but it is also very advisable to lose weight [6]. There are six types of nutrients that the body needs: vitamins, minerals, protein, fats, water, and carbohydrates [7]. It is important to consume them daily to build and maintain a healthy body function (https://www.who.int/publications/m/item/healthy-diet-factsheet394 (accessed on 12 December 2021)). The exact composition of a healthy, balanced, and varied diet depends on the needs of each person (age, sex, life habits, physical exercise), the cultural context, locally available foods, and eating habits [8]. However, according to the WHO (https://www.who.int/news-room/fact-sheets/detail/healthy-diet (accessed on 17 December 2021)), the basic principles of healthy eating are always the same regardless of the factors mentioned above.

With the arrival of the Semantic Web [9], ontologies have become a relevant model for representing knowledge in many scientific fields, including medicine [10], biomedicine [11], bioinformatics [12], and many others [13]. Ontologies are the basis for sharing and reusing knowledge both on the Semantic Web and in scientific and non-scientific fields. It is a conceptualization of a domain in an understandable, machine-readable human format consisting of entities, attributes, relations, and axioms [14]. Ontologies represent a method for formally expressing a shared understanding of information, which can be considered as a prerequisite for the Semantic Web. The W3C states that “ontologies are formalized vocabularies of terms, often covering a specific domain and shared by a community of users. They specify the definitions of terms by describing their relationships with other terms in the ontology” (https://www.w3.org/2007/OWL/wiki/Document_Overview_branch_for.refs (accessed on 17 December 2021)).

In this work, we present an ontology of the food domain related to health self-management in a chronic multifactorial disease such as obesity and the different levels of overweight. Semantic technologies were integrated into the project to provide a formal representation that enables accurately defining concepts such as food allergies, ingredients, recipes, and diets in this particular domain. Reasoning techniques were leveraged to infer users’ health status and make dietetic recommendations accordingly. All these components articulate a knowledge-based recommender system for obesity and overweight conditions. To sum up, the main contributions of this work are as follows:

- We present the Ontology of Dietary Recommendations (ODR), a shareable conceptual model comprising the most relevant elements in the food recommendation domain including diets, recipes, ingredients, anthropometric indexes, and food allergies;
- A knowledge base built on top of the ODR ontology was used to support the recommendation process. The knowledge base was populated with data about specific diets and recipes gathered from scientific publications and other sources;
- The knowledge base and the underlying ontology schema were developed by adopting the Linked Data best practices (https://www.w3.org/TR/ld-bp/ (accessed on 13 December 2021)). In particular, the ontology is easily extensible with new axioms, and the knowledge base can be updated with the latest advancements in the field;
- We propose a knowledge-based diet recommendation system that infers the most suitable diets given the user’s data and returns a list of recipes that comply with those diets’ composition details. In the prototypical implementation described here, the BMI was used as the main indicator to assess the users’ health status (i.e., their obesity level), and the recommendations are focused on helping with weight loss.

The remainder of the paper is organized as follows. In Section 2, background information about existing ontologies that conceptualize the health domain in general and food and nutrition in particular is put forward, along with the state-of-the-art dietary recommendation systems. The proposed ontology-based nutritional recommender system is described in detail in Section 3. The results of the evaluation of the system are discussed in Section 4. Finally, in Section 5, conclusions are presented and future research directions identified.
2. State-of-The-Art

Currently, overweight and obesity are considered nutritional disorders, where inadequate physical activity and abusing high-calorie intakes make their presence more likely [2,15]. There is no single rule applicable for everyone, but most recommendations reduce the intake of calories by day (https://www.nhs.uk/conditions/obesity/treatment/ (accessed on 17 December 2021)). The best way to achieve this reduction depends on removing unhealthy food such as processed food and alcoholic or sugary drinks [16]. However, the self-management of obesity is not an easy task, since it requires re-educating eating, adopting new nourishment habits, and tracking the intake by counting calories [17]. Although people are aware of the importance of having healthy eating habits, a high percentage neglects following those habits due to the cognitive effort required to organize meals [18]. Furthermore, suggesting a proper meal is not an easy task, not only because its implications in terms of personal health, but also because it requires trying to advise what people should eat, considering their personal preferences (e.g., vegan, vegetarian, semi-vegetarian). Consequently, specific ingredients (e.g., tomato, chicken, etc.) play an important role in food recommendations (i.e., dishes recommended as part of a diet) since those ingredients should be linked to each personal taste [19]. A way of modeling the decomposition of recipes into individual ingredients must be defined to enable the recommender engine to suggest recipes that satisfy all personal requirements. In this sense, ontologies enable conceptualizing a domain providing a specific vocabulary [14]. In the literature, ontologies have been applied not only in the food domain [20,21], but also in improving recommendation models [22,23]. In the following sections, a detailed analysis of cutting-edge food conceptual models and semantically-empowered recommender engines is provided.

2.1. Health and Nutrition Ontologies

In health care, an ontology is a formal conceptualization of a health-related domain [24]. Ontologies are used to model the medical domain’s semantics, making its terminology explicit [25,26]. Currently, a great variety of ontologies has been developed covering several domains that go from representing the human disease domain to creating medical terminology, enabling medical information systems to be interoperable and facilitating knowledge interchange across multiple disciplines [27,28]. Concretely, in the health domain, there is a variety of interesting works proposed by Malhotra et al. [29]. They defined the Alzheimer Disease Ontology (ADO) to cover the absence of formal representation in this disease. The ontology was built using various knowledge resources such as the National Center for Biomedical Ontology to extract synonyms and abstracts obtained from PubMed to construct a dictionary or encyclopedia and atlas of the mentioned disease. Following the same domain, He et al. [30] built the coronavirus Infectious Disease Ontology (CIDO) to cover areas related to coronavirus diseases such as etiology, transmission, prevention, and treatment, among others. The CIDO integrates more than 300 ontologies and contains 4000 terms imported from some 20 other ontologies such as ChEBI [31] or the Human Phenotype Ontology [32]. It defines axioms that relate, for instance, the virus SARS-CoV-2 to the produced disease COVID-19 or the phenotype that manifests the disease in infected patients.

Focusing on the nutritional domain, El-Sappagh et al. [33] proposed the Diabetes Mellitus Treatment Ontology (DMTO) as a knowledge base that provides a detailed description of a patient’s conditions, profiles, complications, and symptoms of Type 2 Diabetes Mellitus (T2DB) treatment. In the same line, Yang et al. [34] emphasized the application of the Semantic Web in nutrition research. They proposed the Ontology for Nutritional Epidemiology (ONE), a formal way to specifically relate terms used in this domain. To build the ontology, the authors analyzed multiple articles and ontologies related to the nutritional epidemiology domain. More related to the food recommendation, Gharibi et al. [35] described the process of building FoodKG, a food knowledge graph that links recipes to their ingredients and ingredients to their nutritional data. Following the same line, Çelik
in [36] presented the Ontology-Driven Mobile Safe Food Consumption System (FoodWiki) and the Food Ontology Knowledge Base (FOKB). The system provides a web service that enables food-sensitive users obtain detailed information about their consumed nutrients individually. It does not consider recipes, only individual ingredients. The retrieved data are based on the FOKB, which semantically describes health risk groups, unsuitable components for those groups, and unhealthy ingredients with side effects taking into account users’ health status.

In this work, we implemented an ontology to define diets recommended to users, considering their health status. Diets are related to a set of recipes, and each recipe is semantically described in terms of ingredients. Taking into account that the FOKB defines a detailed specification of ingredients, part of that model has been reused to build our ontology. An in-depth description of the ontology conceived of in this work can be found in Section 3.1.

2.2. Dietary Recommender Systems

Recommender systems are software applications that serve as an information filter to help users navigate large sets of alternatives and find items that fit their preferences [37,38]. In the literature, several techniques have been proposed to predict the users’ interests: (i) collaborative filtering (CF): the most widespread approach consisting of collecting preferences from many users, filtering use strategies that require collaboration, and making automatic predictions; (ii) content-based filtering (CB) utilizes features defined in items liked by users to find similar ones and suggest them; (iii) knowledge-based techniques (KBs) use implicit knowledge about users and products to infer users’ needs and preferences and make the recommendations [39]; (iv) hybrid approaches that are based on a combination of the mentioned techniques. Those techniques have been widely applied in a variety of diverse domains, for instance CF in music [39], CB to retrieve learning resources [40], KB in dialogue systems [41], and hybrid approaches in dating recommender systems [42].

In particular, in the domain of this work, dietary recommendation, Agapito et al. [43] presented DIETOS, a dietary recommender system based on user health conditions and eventual chronic diseases. The system utilizes questionnaires to create users’ health profiles and make the recommendations considering their medical conditions. In the same line, Pecune et al. [44] described three different recommender systems: (i) a preference-based recommender that utilizes a content-based filtering technique to only take into account the users’ preferences to suggest recipes; (ii) a healthy recommender, which only considers healthy diets, ignoring users’ preferences; (iii) a hybrid approach based on a combination of both previous approaches. Following the same idea about preference recommendations, Toledo et al. [45] developed a framework that suggests a personalized daily meal plan. The system incorporates a multi-criteria decision analysis tool for personalized meal recommendation by considering users’ preferences. Similarly, Ting and Chen [46] proposed the Dietary Recommendation System (DRS), a system capable of analyzing users’ situation, inferring whether they are following a balanced diet, and recommending suitable diet recipes to make their lives healthier. The DRS utilizes an ontology to relate nutritional information to recipes. The former is obtained from the Food Composition Database. It provides information about various nutrients, their nutritional values, and the amount that should be incorporated daily in an adult diet. The latter were extracted from two different web pages, where people publish recipes describing the ingredients and the steps to prepare the dish. Finally, Tumnark et al. presented in [47] a recommender engine that provides personalized dietary recommendations for weightlifting athletes. The engine infers the amount of the total energy expenditure (TTE) of athletes, considering their activity level. Then, it leverages this information to recommend a balanced menu with varied foods in appropriate portions to satisfy their nutritional needs. It also uses an ontology to semantically describe specific athlete information such as weight, height, body type, training plans, food items, and nutritional values. The engine utilizes this information to execute rules expressed in the Semantic Web Rule Language (SWRL) [48] to make suggestions.
considering the recommended energy and nutrients needed for specific training types, age, and height.

In this work, we propose a recommender system that suggests recipes complying with certain diet restrictions of users depending on their health status. The recommendation engine was built on top of a domain ontology. In particular, the system analyzes the users’ health situation by taking into account some of the main anthropometric measurements such as height, weight, and BMI, among others. The semantic underpinnings of our approach enable the system to identify the diets best suited for each user and then list recipes whose ingredients and nutritional values match the recommended diets’ constraints.

3. Recommender System Architecture

In this work, we propose a novel knowledge-based tool that provides nutritional recommendations to users based on their profile. While the aim of the conceived framework is to take into account the most relevant features in the users’ profiles for dietary advice, in this prototypical implementation, only the BMI and the corresponding overweight and obesity types are considered. The overall architecture of the ontology-based nutritional recommender system is depicted in Figure 1 and is comprised of three modules, namely the User Data Collector, which was designed to enable users to interact with the system, the Ontology of Dietary Recommendations (ODR), which contains the conceptualization of the diet domain, and the Diet Recommender, which is responsible for providing users with the recipes that best suit the individual users’ diet needs.

![Figure 1. Architecture of the system.](image)

Mainly, the system’s target is to assist people with finding recipes that enable them to lose weight. In brief, the system works as follows. First, a user inserts his/her personal data by using the collector. Details such as weight, age, height, and so on, need to be provided. Then, the system will take advantage of these details to determine the user’s health condition. For instance, if the introduced weight and height are 160 cm and 80 kg, the user is associated with Obesity Type I since the BMI is between 30 kg/m² and 34.9 kg/m². Next, the system will select the dietary plan to reduce the overweight considering the type of nutritional disorder detected. To do so, the reasoner will utilize the ODR. Following the above example, the diets that will be suggested given the identified obesity type are bland, dissociated, Dukan, and Mediterranean. Finally, the outcome of the system will be a list of the recipes that match the nutritional restrictions defined in the health plan of those suggested diets. In the following subsections, each component is described in detail.
3.1. Ontology of Dietary Recommendations

To build the ontology, we used the iterative method proposed by Abdel-Badeeh and Hisahm (Salem, 2012), which comprises several stages to specify the requirements, collect and analyze the data, and finally, construct the initial model. As a baseline, the FOKB [36] was employed to structure the ODR. The following classes were preserved from the original schema: Product, which classifies ingredients by EAN (European Article Number); FoodAllergies, to express relations between food and allergen items; Ingredients, to define the building blocks of the recipes; and finally, Measure, which includes different rates to quantify indicators such as the Basal Metabolic Rate (BMR), the BMI, or Blood Pressure (BP), among others. For the purposes of our work, several elements were added to obtain the resulting ODR schema, namely Person, Recipe, and Diets. Next, the purpose of these added classes is described.

Person: Axioms were included to define the most common variety of eating disorders that can affect users. There are several ways to determine whether an individual is suffering from this dietetic situation. A variety of measures can be considered to specify the grade of obesity. Depending on the measure obtained, the individual will be classified into a set of grades that go from ‘Very Severely Underweight’ to ‘Obesity Type III’. In particular, for this work, the BMI was chosen to assess the grade of obesity because of its simplicity and low cost. In [49], Weir and Jan set the cutoffs to estimate the body fat. Concretely, for obesity type I, they estimated that the BMI must be between 30 kg/m² and 34.9 kg/m². Consequently, to express these cutoffs in the ontology, we defined the has_bmi data property to declare the limits and various classes to represent the different types of obesity. For example, Obesity_Type_I was defined as a class equivalent to Person whose BMI is between the margins previously set. Following the same procedure, the following defined classes are declared: Normal, Obesity_Type_I, Obesity_Type_II, Obesity_Type_III, Overweight, Underweight, Severely_Underweight, and Very_Severely_Underweight. Table 1 contains the semantic description of these defined classes.

Table 1. Axioms defined in the ontology to specify some nutritional profiles.

<table>
<thead>
<tr>
<th>Class</th>
<th>Axioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very_Severely_Underweight</td>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&lt; &quot;15.0&quot;^^xsd:double])</td>
</tr>
<tr>
<td>Severely_Underweight</td>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&gt;= &quot;15.0&quot;^^xsd:double)</td>
</tr>
<tr>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&lt; &quot;16.0&quot;^^xsd:double])</td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&gt;= &quot;16.0&quot;^^xsd:double)</td>
</tr>
<tr>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&lt; &quot;18.5&quot;^^xsd:double])</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&gt;= &quot;18.5&quot;^^xsd:double)</td>
</tr>
<tr>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&lt; &quot;25.0&quot;^^xsd:double])</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&gt;= &quot;25.0&quot;^^xsd:double)</td>
</tr>
<tr>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&lt; &quot;30.0&quot;^^xsd:double])</td>
<td></td>
</tr>
<tr>
<td>Obesity_Type_I</td>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&gt;= &quot;30.0&quot;^^xsd:double)</td>
</tr>
<tr>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&lt; &quot;35.0&quot;^^xsd:double])</td>
<td></td>
</tr>
<tr>
<td>Obesity_Type_II</td>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&gt;= &quot;35.0&quot;^^xsd:double)</td>
</tr>
<tr>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&lt; &quot;40.0&quot;^^xsd:double])</td>
<td></td>
</tr>
<tr>
<td>Obesity_Type_III</td>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&gt;= &quot;40.0&quot;^^xsd:double)</td>
</tr>
<tr>
<td><img src="#" alt="Axiom" /> and (has_bmi some xsd:double&lt; &quot;44.0&quot;^^xsd:double])</td>
<td></td>
</tr>
</tbody>
</table>
Diet: This organizes the types of diets recommended by the system. At the moment, only the normal and hypocaloric diets (i.e., a diet that is low in calories) are defined since those are the ones associated with overweight/obesity problems representing one of the main challenges for dietitians and nutritionists. Eight subclasses were considered representing eight different hypocaloric diets (see Figure 2), namely Alkaline_Diet, Artichoke_Diet, Atkins_Diet, Bland_Diet, Dissociated_Diet, Dunkan_Diet, Mediterranean_Diet, and Pineapple_Diet. This is an initial set of diets chosen based on existing literature [50–52]. Some of these diets have been refuted by health experts given their relationship to nutritional deficiencies, but were included due to their popularity. Please bear in mind that this set of diets is dynamic and can be updated based on the latest scientific evidence.

![Figure 2. Diets' hierarchy defined in the ODR.](image)

Each diet was systematically associated with a nutrition disorder that is related to users' metabolism rates. For example, in Figure 3, the axiom defined to link Alkaline_Diet to Overweight is depicted. Therefore, the recommendation engine will consider such an association when providing dietary suggestions to users depending on their health profile.

![Figure 3. Axiom to link a diet with the user's health status.](image)

Recipe: This contains the entire taxonomy of all the recipes included in the knowledge base. The foodRecSys – V1 dataset (https://www.kaggle.com/elisaxxygao/foodrecsysv1 (accessed on 17 December 2021)), where thousands of recipes have been collected from the Allrecipes.com website, was used to create such a hierarchy. The dataset includes data such as ingredients, nutritional components, and even users’ ratings. We randomly selected 105 recipes from the whole dataset and scrutinized details about the ingredients and nutritional components to integrate each recipe in our knowledge base. To do so, we developed a procedure (see Algorithm 1) that analyzes each recipe, extracts its ingredients and nutritional values, and expresses the extracted information in the form of OWL axioms. The process takes such details and utilizes data properties and object properties to link the retrieved data to entities already defined in the knowledge base. For instance, the recipe “Abuelas_Picadillo” is made of red wine, butter, bell pepper, garlic, and olive oil, among other ingredients, and contains 74 mg of cholesterol, 378 calories, and 24 g of proteins. Then, the procedure utilizes this information to define the axioms that represent the new recipe with the mentioned ingredients and nutritional values.
Algorithm 1 Mapping ingredients.

1: procedure ANALYZE_RECIPE(strRecipe)
2:     while not at end of the recipe do
3:         word ← tokenize(strRecipe)
4:         if ingrDictionary.contains(word) then
5:             // word is an ingredient
6:             newAxiom ← getIngredientAxiom(word)
7:             ontology ← newAxiom
8:         else if isNutritionalValue(word) then
9:             // word is a nutritional value
10:            newAxiom ← getNutritionalAxiom(word)
11:            ontology ← newAxiom
12:        end if
13:     end while
14: end procedure

In addition, in order to associate recipes with the type of diets, we defined a specific subclass of Recipe for each diet type (see Figure 4). The aim of this hierarchy is to clearly state the unique constraints, in terms of both ingredients and nutritional values, that a recipe should satisfy to be valid for a given diet. Essentially, each class defines in OWL language the set of nutritional restrictions that users must follow to implement the diet. Therefore, the number of proteins, calories, or allowed types of the ingredient constitute some of the rules defined in those classes. In this way, when a diet is selected, the recipes will be filtered according to these nutritional restrictions in an automatic way using a reasoning engine. For example, the class Recipe_For_Alkaline_Diet is defined as shown in Figure 5. Since the Alkaline_Diet is characterized by high fruit and vegetable content, avoiding chicken, red meat, and eggs, among others, recipes for such a diet should comply with those restrictions.

Figure 4. Hierarchy of recipes created for each specific diet.

Figure 5. Semantic definition of the class Recipe_For_Alkaline_Diet.

The ontology was developed in OWL 2 [53] and is available at https://agrisemantics.inf.um.es/ontologies/ODR (accessed on 16 December 2021). Table 2 shows some of the most relevant metrics of the ontology. The correctness of the resulting ontology was checked using the following tools: (i) the RDFS Validator (https://www.w3.org/RDF/Validator/, accessed on 16 December 2021) to repair the definitions of concepts, relations, and instances;
(ii) OOPS! (http://oops.linkeddata.es/catalogue.jsp, accessed on 16 December 2021) (the Ontology Pitfall Scanner!) to identify deficiencies in metadata information such as license and version information, among others; (iii) OQuare (https://semantics.inf.um.es/ontology-metrics/, accessed on 16 December 2021) to test the model’s features. Figure 6 depicts the core concepts of the resulting ontology.

Table 2. The ODR’s main metrics.

<table>
<thead>
<tr>
<th>Metrics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Axioms</td>
<td>5733</td>
</tr>
<tr>
<td>Logic Axioms</td>
<td>2815</td>
</tr>
<tr>
<td>Classes</td>
<td>1193</td>
</tr>
<tr>
<td>Object Properties</td>
<td>21</td>
</tr>
<tr>
<td>Data Properties</td>
<td>23</td>
</tr>
</tbody>
</table>

Figure 6. Global hierarchy concepts of the ODR.

3.2. User Data Collector

The user data collector represents the entry point for users to utilize the system. It provides an interface that enables users to enter personal information such as weight, height, and age. The system will use these details to determine the health condition of users. Based on these data, metabolic indicators such as the BMI, the BMR, or the BP, among others, will be calculated. Such information will help the recommendation engine to identify whether users suffer from a nutritional disorder and to suggest the diet that most fits the users’ dietetic problems.

In Figure 7, the process followed to define the axioms in the ontology from the data inputted by the users is depicted. Two parts can be easily differentiated in the figure, the form in which users insert their personal information on the left and the data type properties defined in the ontology on the right. Each property relates a Person entity to each data introduced by the users.
Algorithm 2 shows the method developed to insert a new user in the ontology. `baseURI` stands for the ontology’s namespace, and `personLabel` contains the string “Person”, identifying the name of the parent class, namely, Person. The new user is then included as a new class in the ontology, subclass of the class Person.

**Algorithm 2** Populate ontology with a new user (class Person).

1: procedure INSERTNEWUSER(userName, weight, age, ...)
2: newUserClass ← getOWLClass(baseURI + userName)
3: personClass ← getOWLClass(baseURI + personLabel)
4: newAxiom ← getOWLSubClassAxiom(newUserClass, personClass)
5: ontology ← newAxiom
6: weightAxiom ← getOWLLiteralAxiom(newUserClass, weight)
7: ontology ← weightAxiom
8: ageAxiom ← getOWLLiteralAxiom(newUserClass, age)
9: ontology ← ageAxiom
10: end procedure

### 3.3. Diet Recommender Module

The Diet Recommender module acts as the mediator between the user interface and the knowledge base. It is responsible for systematically converting the data collected by the user interface into an OWL language to ask the reasoner for the users’ health condition. The query was written in Manchester OWL Syntax using classes, the object properties, and the data properties identifiers. As a result, it retrieves the users’ category of obesity, BP level, or the BMR. Similarly, the system will utilize the information inferred to select the diets most recommendable for the users’ health status. Hence, the module works as follows:

- **Step 1:** retrieving the new user—At this stage, user’s personal details are transformed into OWL axioms, and the recommender will retrieve such information to make the diet recommendation;
- **Step 2:** linking diets—Users’ medical conditions are the first factor the system considers for making suggestions. Therefore, the details inserted by users are used to classify their situation and recommend diets accordingly. In Figure 3 above, the axiom utilized for specifying the association between the user’s dietetic profile and diets is shown;
- **Step 3:** analyzing recipes—Once the diet is assigned, a list of recipes can be recommended considering the ingredients more adequate for users. Figure 8 shows the
exemplary taxonomy of Alkaline diet recipes. The classification of such recipes is performed automatically by the reasoner by considering the semantic description of each recipe subclass (see Figure 5 above).

Figure 8. Recipes that comply with the Alkaline diet’s restrictions.

Hence, depending on the user’s taste, different recipes can be combined to follow the recommended diet. For instance, let us suppose that a new user called Robert, age 48, 90 kg, 1.70 cm, aims to receive nutritional advice given his current health status. First, he has to introduce his personal details in the form. The User Data Collector module will harvest the information, compute the measures, and create a new Person entity in the ODR ontology model. Second, considering the information inserted, the Diet Recommender module will utilize an OWL Reasoner to infer the user’s health status. In particular, in this example, the module will consider the BMI to infer the obesity type suffered by the user, if any, and will associate the most advisable diets, accordingly. Finally, the user can choose from various recipe options and organize his menu regime by complying with the chosen diet.

4. Evaluation

In this section, the evaluation dataset is described along with the metrics used to assess the performance of the recommender system. Then, the results obtained are put forward and discussed in comparison with related work.

4.1. Dataset

The dataset used in the evaluation experiment is based on the one created by Palechor and Manotas in [54]. This dataset was chosen because it contains the basic information about individuals required as the input by the proposed system to provide the resulting recommendations. In particular, it consists of 2111 records and 17 attributes, storing a variety of variables that harvest details about eating habits and physical condition. For instance, age, height, weight, frequency of consumption of vegetables, number of main meals, and consumption of food between meals are some of the variables collected by the authors utilizing a survey. From this dataset, a panel of advisors evaluated such attributes for each record and manually suggested a set of diets from those defined in our ontology. When the recommendations were different, they met and analyzed the case independently until there was a general agreement. As a result, a new column was added to the dataset containing the agreed set of diets pre-selected by the experts for each test case (i.e., row corresponding to one of the individuals’ records), which constituted the ground truth for the experiment. The panel of advisors was composed of five students of the Master of Nutrition, Technology and Food Safety of the University of Murcia. The students agreed to be part of this study, which was proposed as a final practice.

Table 3 classifies the number of records of the dataset based on their nutritional categories. The records belonging to “Insufficient” were discarded from the experiment because it was explicitly focused on hypocaloric diets. Therefore, only 1839 records were used for evaluation purposes.
Table 3. Dataset distribution.

<table>
<thead>
<tr>
<th>Type</th>
<th>Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insufficient</td>
<td>272</td>
</tr>
<tr>
<td>Normal</td>
<td>287</td>
</tr>
<tr>
<td>Overweight</td>
<td>580</td>
</tr>
<tr>
<td>Obesity Type I</td>
<td>351</td>
</tr>
<tr>
<td>Obesity Type II</td>
<td>297</td>
</tr>
<tr>
<td>Obesity Type III</td>
<td>324</td>
</tr>
</tbody>
</table>

4.2. Evaluation Metrics

The metrics typically calculated to assess the performance of classification models such as the one described in this work are: accuracy, precision, recall, and F-measure. These metrics have been widely applied to evaluate information retrieval systems [35]. Therefore, these four outcomes were utilized to measure the performance of the recommender. Accuracy is measured as follows:

\[
\text{accuracy} = \frac{\text{correct diets predictions}}{\text{total predictions}}
\]

where \(\text{correct diets predictions}\) defines the number of correctly predicted diets, including the relevant diets correctly retrieved and the non-relevant diets precisely left. Next, the precision is calculated as follows:

\[
\text{precision} = \frac{|\{\text{relevant diets}\} \cap \{\text{retrieved diets}\}|}{|\{\text{retrieved diets}\}|}
\]

where \(\text{relevant diets}\) stands for the set of diets suggested by the panel of advisors and \(\text{retrieved diets}\) represents the set of diets inferred from the ontology. In addition, the recall measure was used to compute the number of relevant diets that were successfully retrieved. The recall formula is as follows:

\[
\text{recall} = \frac{|\{\text{relevant diets}\} \cap \{\text{retrieved diets}\}|}{|\{\text{relevant diets}\}|}
\]

Finally, the harmonic mean of both measures is calculated producing the F-measure. The formula is as follows:

\[
F - \text{measure} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}
\]

4.3. Results

Table 4 shows the results of the experiments for the four measures considered above. The table contains the average of the 1839 records utilized in the evaluation. The overall accuracy of the proposed approach was 0.84, 0.92, 0.86, 0.84, and 0.85 for “Normal”, “Overweight”, “Obesity Type I”, “Obesity Type II”, and “Obesity Type III”, respectively. Given the characteristics of our dataset, where for each class (e.g., “Normal”), only one pool of diets can be considered, the results are shown aggregated by nutritional profile.
Table 4. Results obtained in the experiments.

<table>
<thead>
<tr>
<th>Profile</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.84</td>
<td>0.34</td>
<td>0.77</td>
<td>0.47</td>
</tr>
<tr>
<td>Overweight</td>
<td>0.92</td>
<td>0.84</td>
<td>0.97</td>
<td>0.90</td>
</tr>
<tr>
<td>Obesity Type I</td>
<td>0.86</td>
<td>0.72</td>
<td>0.94</td>
<td>0.81</td>
</tr>
<tr>
<td>Obesity Type II</td>
<td>0.84</td>
<td>0.71</td>
<td>0.93</td>
<td>0.80</td>
</tr>
<tr>
<td>Obesity Type III</td>
<td>0.85</td>
<td>0.71</td>
<td>0.93</td>
<td>0.80</td>
</tr>
<tr>
<td>Overall</td>
<td>0.87</td>
<td>0.7</td>
<td>0.92</td>
<td>0.78</td>
</tr>
</tbody>
</table>

4.4. Discussion

In general terms, the recommender system obtained promising results despite a significant imbalance in the records utilized (the number of test cases for “Overweight” was double the number of test cases for “Normal”). However, a big difference can be observed between the performance for the case of users with “Normal” weight and users with obesity problems. In particular, the worst result was obtained for the “Normal” profile, where the precision was 0.34, meaning that many diets not selected by the panel of advisors were suggested by the system. Nevertheless, recall for this weight profile was significantly better, 0.77. This indicates that, although various incorrect diets were suggested, most of the diets selected by the panel of advisors were included. As shown in the table, such behavior was repeated in the other profiles, where the best precision and recall were obtained for “Overweight”, 0.84 and 0.97, respectively. In fact, the same dynamics is reflected in the aggregate results, where the precision was 0.70, while the recall reached 0.92. Given that the focus of this work is set on obesity and overweight, removing the “Normal” profile test cases resulted in an accuracy, precision, recall, and F-measure of 0.87, 0.76, 0.95, and 0.84, respectively.

As mentioned above, the low values in precision mean that given a test case, the system retrieves false positives, namely it recommends diets that the panel of advisors did not include in their selection for such a case. Conversely, the larger recall exposes that the system suggested a high percentage of the actually relevant diets (i.e., those previously selected by the panel of advisors). This behavior can be partly due to the fact that the information used by the advisors, consisting of the 17 attributes available in the dataset, differed from that considered by the system. In particular, while the system only took into account the attributes such as the BMI, age, weight, and height to calculate the obesity profile and assign the corresponding nutritional plan, the panel of advisors also analyzed other parameters such as the frequent consumption of high caloric food, vegetables, food between meals, or water daily. Therefore, this lack of information might hamper the system’s performance, justifying the limited precision when comparing against the advisors’ recommendations. To overcome such a limitation, we plan to improve the ontology model and its inferencing ability by including in the rules aspects such as calorie consumption, frequency of ingredient type, nutrient reference values, or allergies or intolerances suffered.

In the literature, the approach closest to that presented here is the system proposed by Manoharan [56], which employs statistical and deep learning models such as logistic regression, naive Bayes, multilayer perceptrons, or recurrent neural networks to suggest diets for diverse patients with different types of disease. The highest precision achieved by the system was 0.95, and the recall was 0.91. Another approach quite similar is the one proposed by Chavan [57], where three recommender systems based on content-based, collaborative filtering, and hybrid techniques were analyzed. The best precision and recall results were obtained by using the hybrid technique, 0.20 and 0.55, respectively. Both of the works mentioned above proposed various diet recommendation systems. However, it does not seem appropriate to compare the results of our approach with theirs for several reasons. First, all approaches make diet suggestions, but in different ways. For instance, the former recommends diets for diverse patients with different types of disease. Still, neither the type of disease suffered by patients nor the type of diets utilized for the experiments were specified. On the other hand, the latter describes different systems, but they focus more on
ingredients and recipes than diet recommendations. In addition, there are other interesting approaches also mentioned above (see Section 2.2), such as DIETOS [43] or Nutriflow [58], which propose similar ideas, but do not provide an evaluation.

5. Conclusions and Future Work

In the medical community, obesity is considered as a chronic illness that comes from a multifactorial etiology [59]. This disorder is characterized by the unchecked increase of fat tissue in the body. As a preventable disorder, we need to develop mechanisms that help people overcome it. Therefore, we believe that the knowledge representation techniques are remarkably useful to that end. It provides suitable elements such as concepts, object, and data properties to accurately define the data in this domain. Its usage provides the opportunity to the biomedical community to mitigate this disorder with consistent, reusable, and standardized descriptions about obesity and dietetic food.

In this work, two main contributions can be identified. On the one hand, the recommender system is capable of making suggestions considering certain users’ diet restrictions depending on their medical condition. On the other hand, we implemented the Ontology of Dietary Recommendations (ODR), which concretely integrates information about obesity, nutrition, food, and health. In addition, a couple of algorithms were implemented to convert heterogeneous data into semantic axioms, such as user profiles and recipe ingredients. We expect this conceptual model will help the intelligent systems and the medical profession to exploit this knowledge base in different ways.

This work aimed at providing general diet suggestions without considering users’ requirements such as tastes, food restrictions, diseases, and allergies, which represent a significant constraint. Furthermore, our approach does not provide a dietetic diary for users, which means that there is no dynamic readjustment, but instead, this adaptation must be carried out manually. We plan to address these limitations as part of future work. Likewise, the way recipes are modeled in the ontology will be enriched by including more details concerning the nutritional properties, thus facilitating the development of different added-value features. So far, each recipe is characterized by properties such as has_fat, has_calories, has_sodium, has_cholesterol, or has_protein. Some other attributes such as energy, fiber content, and carbohydrate content can be defined in the “Recipe” ontology concept so as to enable the system to manage new relevant indicators and provide better recommendations. That additional information can be gathered from publicly available databases such as Open Food Facts (https://world.openfoodfacts.org/ (accessed on 12 December 2021)).

For the proof-of-concept implementation of the system presented in this manuscript, the knowledge base was populated relying on some scientific publications and information gathered from the Internet. For it to be deployed to customers, nutrition experts should be consulted to review and update the knowledge base by including only those diets and recipes that are endorsed by the scientific community and defining the most accurate and feasible indicators to determine the (nutritional) health condition of individuals. Indeed, as pointed out above, research has shown that the BMI alone is not the best tool to assess the grade of obesity [60]. A better understanding of the individuals’ overall health status and nutritional requirements can be obtained by combining different markers. Finally, the FoodAllergies concept was included in the ontology, but this contribution does not go far enough. We believe that it would be interesting for end users to axiomatically expand the ontological hierarchy defining more specific user profile diets by considering allergens and, consequently, to provide a better recommendation service.

**Funding:** This work is part of the research project LaTe4PSP (PID2019-107652RB-I00) funded by MCIN/AEI/10.13039/501100011033 and was partially supported by the Seneca Foundation, the Regional Agency for Science and Technology of Murcia (Spain), through Project 20963/PI/18, and the Community of Madrid, through the Young Researchers R+D Project. Ref. M2173 – SGTRS (co-funded by Rey Juan Carlos University).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The Ontology of Dietary Recommendations (ODR) is publicly available at: https://agrisemantics.inf.um.es/ontologies/ODR (accessed on 17 December 2021). Publicly available datasets were analyzed in this study. These data can be found here: https://www.kaggle.com/elisaxxygao/foodrecsysv1 (accessed on 17 December 2021), https://www.kaggle.com/ankurba jaj/obesity-levels (accessed on 17 December 2021).

**Acknowledgments:** We would like to thank the five students who were part of the panel of advisors for their collaboration in the evaluation experiment.

**Conflicts of Interest:** The funders had no role in the design of the study; in the collection, analyzes, or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

**References**


8. Johnson, S. *The Importance of Nutrition and Preventing Malnutrition in Older Adults: A Literature Review and Informational Booklet*; UNL Digital Commons: University of Nebraska-Lincoln: Lincoln, Nebraska, 2021; Volume 294.


13. Rodríguez-García, M.A.; García-Sánchez, F.; Valencia-García, R. Knowledge-Based System for Crop Pests and Diseases Recognition. *Electronics* 2021, 10, 905. [CrossRef]


37. Zhang, X.; Ferreira, P.; de Matos, M.G.; Belo, R. Welfare Properties of Profit Maximizing Recommender Systems: Theory and Results from a Randomized Experiment. MIS Q. 2021, 45. [CrossRef]


42. Rodríguez-García, M.A.; Valencia-García, R.; Colomo-Palacios, R.; Gómez-Berbís, J.M. BlindDate recommender: A context-aware ontology-based dating recommendation platform. J. Inf. Sci. 2019, 45. [CrossRef]


45. Toledo, R.Y.; Alzahrani, A.A.; Martínez, L. A Food Recommender System Considering Nutritional Information and User Preferences. *IEEE Access* 2019, 7, 96695–96711. [CrossRef]


54. Palechor, F.M.; de la Hoz Manotas, A. Dataset for estimation of obesity levels based on eating habits and physical condition in individuals from Colombia, Peru and Mexico. *Data Br.* 2019, 25, 104344. [CrossRef]


