

NAct: The Nutrition & Activity Ontology for Healthy Living

Dorothea TSATSOU^{a,1}, Elena LALAMA^b, Saskia L. WILSON-BARNES^c,
Kathryn HART^c, Véronique CORNELISSEN^d, Roselien BUYS^d,
Ioannis PAGKALOS^e, Sofia BALULA DIAS^f, Kosmas DIMITROPOULOS^a, and
Petros DARAS^a

^aInformation Technologies Institute, Centre for Research & Technology Hellas,
Thessaloniki, Greece

^bDepartment of Endocrinology and Metabolic Diseases, Charité Universitätsmedizin,
Berlin, Germany

^cSchool of Biosciences and Medicine, Faculty of Health and Medical Sciences,
University of Surrey, Guildford, UK

^dDepartment of Rehabilitation Sciences, Katholieke Universiteit Leuven, Leuven,
Belgium

^eDepartment of Nutritional Sciences and Dietetics, International Hellenic University,
Thessaloniki, Greece

^fFaculdade de Motricidade Humana, Universidade de Lisboa, Lisbon, Portugal

Abstract. This paper presents the NAct (Nutrition & Activity) Ontology, designed to drive personalised nutritional and physical activity recommendations and effectively support healthy living, through a reasoning-based AI decision support system. NAct coalesces nutritional, medical, behavioural and lifestyle indicators with potential dietary and physical activity directives. The paper presents the first version of the ontology, including its co-design and engineering methodology, along with usage examples in supporting healthy nutritional and physical activity choices. Lastly, the plan for future improvements and extensions is discussed.

Keywords. ontology, health, healthy living, nutrition, physical activity, user modelling, recommendation, personalisation

1. Introduction

Nutrition research is a fast-moving multidisciplinary field, which combines the expertise of different professionals across different disciplines. A limitation for this research field is that it is not dependent on one variable, but on many, and to analyse this in a practical and ethical way represents a significant issue. Randomized controlled trials are the gold standard on which many dietary recommendations are predominantly based upon

¹Corresponding Author: Dorothea Tsatsou, Information Technologies Institute, Centre for Research & Technology Hellas, 6th km Charilaou-Thermi road, 57001, Thermi, Thessaloniki, Greece; E-mail: dorothea@iti.gr.

(e.g., [1], [2], [3]). However, the principal limitation of these is that they are not personalised to an individual user.

Nowadays, through the use of artificial intelligence (AI) we can support an individual remotely and less invasively, through healthy lifestyle recommendations [4], for the general population, while also potentially improve the self-management of non-communicable diseases such as obesity, diabetes and cardiovascular disease [5]. To this end, knowledge based systems that rely on expert-verified knowledge enable advanced personalization of healthy lifestyle directives to each individual while at the same time adhering to consolidated and ethical guidelines of different fields of nutrition research.

To this end, this paper presents the NAct ontology, engineered based on evidence-based expert knowledge of different professionals in the nutrition, activity and health fields. Previously developed expert systems suggest the alteration of one variable for an individual's lifestyle. Whereas, NAct ontology and the knowledge-based system that employs it as the backbone for intelligent personalized decision making, aims to fill the gap by adopting a holistic approach. This approach pertains to the adoption of semantic entities and rules that connect each subject's implicit and explicit nutritional and well-being goals, and these goals with the situational condition of the subject and standardized European nutritional and well-being directives.

The structure of the document is as follows. Section 2 provides a brief overview of related work, focusing on the distinct lack of relevant ontologies and comparing NAct with the two most relevant approaches. Section 3 describes the core of the work of this paper, detailing the methodological engineering approach and the main ontology constituents, while providing usage examples and evaluation details. Section 4 informs the reader on the documentation and publishing activities for NAct, while Section 5 provides a conclusion and describes already ongoing future work.

2. Related work

Previous research yielded several key European and International food and nutrient databases complete with few pre-existing nutritional ontologies. The list of food and activity databases is non-exhaustive and thus will not be listed. We will however mention the McCance and Widdowson food database [6] and the Compendium of Physical Activities [7], which were deemed by the domain experts as the vastest and most adequate databases adhering to European nutritional, health and well-being standards. These databases subsequently inspired, to an extent (re top level foods and activities), the respective NAct aspects.

However, the purpose of NAct is not to exhaustively model all possible foods/ ingredients and activities in a mere list or even taxonomy, but rather provide a serve as a robust and intelligent backbone for a knowledge-based AI recommendation system. Such a system would only benefit from a well structured, well defined ontology to serve as the TBox² for subsequent logical inference of suitable nutritional and activity directives to users of a smart healthy living platform. The problem and its requirements are detailed in Section 3.1.

Of the few relevant ontologies that were identified within the literature, most lacked rich semantic correlations, or do not model key components that are required for the purposes of the knowledge-based expert AI system that employs NAct, since they

²Terminological Box

serve a different purpose than the scope of NAct. For example, seminal works such as FoodOn [8], ONS [9], FOBI [10] and CDNO [11] are of relatively shallow expressivity, focusing on extensively modelling, structuring and relating food products and their biochemical role for data retrieval, while lacking enough axiomatic interconnections that allow for advanced recommendation of healthy dietary directives. Comparably, several ontologies that deal with biochemical properties of foods such as ONE [12] and FIDEO [13] bear similar expressivity and are focused on the biochemical properties of foods in relation to very particular health issues (respectively, epidemics and drug reactions) that eschew from the general healthy dietary directives domain.

The two most relevant ontologies to the proposed problem, are the Food Ontology (FOKB) [14] and the HeLiS ontology [15]. Both model food types and nutritional information about them, with the FOKB delving into details about properties of food products, including additives and governing agents (e.g. *anticaking*, *antifoaming*), while HeLiS modelling foods and nutrients as well as physical activities.

The main purpose of FOKB is to serve as the background knowledge to determine side effects of compound and manufactured foods to users allergies and some medical conditions, which is relevant to one of NAct's main requirements, i.e. consider allergies and medical conditions in healthy living recommendations. In the context of NAct, this pertains to a core food/ingredient layer, with relevant connections to medical conditions. FOKB on the other hand delves into the specifics of properties after a produce has been processed (e.g. *additives* etc.). Most importantly, FOKB lacks semantics about particular nutrient-to-food relations that may be used to promote nutritional best practices, as well as any connection to physical activities in relation to conditions. Lastly, FOKB does not interrelate produce information with any nutritional and well-being user goals.

As far as HeLiS is concerned, this is the only other ontology known to date that includes both nutritional and physical activities information. It also includes classifications of nutrients, which FOKB lacks. However, there is a distinct lack of axiomatic interconnection between food types and nutrients or physical activities and properties that affect undertaking these activities (e.g., a medical condition). Rather, those facets are merely presented as a hierarchy of concepts under which a plethora of predetermined individuals are instantiated (e.g. particular, non-updateable activities and undefined nutrient specifics, e.g. *Alcohol_000* under alcohol). There is no freedom to instantiate anything else under these classes, whereas the expert system that employs NAct aims to be able to instantiate any foods, activities and any other information under its core set of abstract entities. Most prominently, HeLiS lacks relations or axioms at the schema basis to liaise the aforementioned information (nutrients, foods, activities) with each other either with respect to particular medical conditions, allergies or with dietary/well-being goals.

However, both of these ontologies have inspired technical aspects of the engineering of NAct, as per relevant shared objectives, i.e., the foods, nutrients and activities structure.

3. Methodology

Engineering the NAct ontology followed the Methontology [16] methodology. This pertains to seven stages: specification, knowledge acquisition, conceptualisation, integra-

tion, implementation, evaluation and documentation. Each stage's developments per the NAct ontology are detailed in the following subsections.

Methontology was elected due to empirical affirmation in past ontology engineering endeavors that the method facilitates the process of creating a new ontology in a collaborative manner by multi-disciplined domain experts and ontology engineers. It is also found to enable pragmatic observations and requirements gathering and, consequently, an efficient process to maintain and evolve the ontology.

3.1. Specification

This phase documents the purpose of the ontology, its semantic expressivity and its scope. The objectives behind engineering NAct can be summarised in the following:

- Model in a slim and holistic manner food-specific nutritional information and activity-specific well-being information.
- Model nutritional and well-being user goals and relate them with nutritional and well-being information.
- Model medical conditions, allergies, intolerances, deficiencies and lifestyle dietary choices and related them with nutritional and well-being information.
- Model properties that define specificities of the aforementioned relationships that aid in the selection of appropriate meals and physical activities for a given person.

The core engineering scope behind this objective is to refrain from a non-exhaustive listing of all foods, activities and their respective detailed information as can be found in existing databases, but rather abstract and generalise as much as possible to basic food and activity types and the most prominent of their respective nutritional and well-being impact, in order to ensure tractability and at the same time decidability in the inference process. To this end, expert-provided information has been distilled into a set of well-defined concepts, relations between them and complex rules that connect them.

The expressivity chosen, to align also with the reasoning capacities of the reasoning component that employs NAct, namely the LiFR fuzzy reasoner [17], lies within the OWL 2 RL³ fragment.

3.2. Knowledge Acquisition

The foundations of the NAct ontology is a wealth of evidence-based information gathered from nutrition scientists, medical experts and scientists with a vast expertise in kinesiology and rehabilitation sciences within the PROTEIN EU⁴ project consortium. The concepts of the ontology were based on the information gathered from the health professionals, which was then connected with the expertise of consortium engineers in semantics, AI/expert systems and logic-based inferencing.

Relations and rules have been developed within the project, to enable end users to achieve their nutritional goals and to relate the nutrients and medical conditions of consumers with the nutritional requirements within the different PROTEIN user groups, such as the overall healthy population, as well as patients with obesity, cardiovascular disease (CVD), Type 2 Diabetes and iron deficiency.

³<https://www.w3.org/TR/owl2-profiles/>

⁴<https://protein-h2020.eu/>

As aforementioned, following a review of the current literature no databases or ontologies currently exist that adequately model the correlations between the mode of physical activity (PA) and nutrition with specific dietary and well-being goals or with medical conditions. Furthermore, no ontologies that specify PA or nutritional rules for particular conditions and diets such as the ones pertaining to the PROTEIN project, as discussed previously were identified. Therefore, the problem at hand required the novel conception of a condition and goals-specific ontology in relation to nutritional and PA aspects.

Overall, NAct has been developed through close and immediate collaboration between the ontology engineering experts and the medical/ nutrition/ PA experts within the PROTEIN consortium, following the analysis of various European databases standards and guidelines (mentioned in Section 5). Several case-based workshops were held during the winter and spring of 2020, discussing ontology requisites and trade-offs, in terms of foods, physical activities, medical conditions and user goals and designing the rules that would interrelate these facets. Workshop results were recorded on an online spreadsheet tracker and used to put experts' knowledge in a machine-understandable formalization under the NAct ontology.

After an initial set of five workshops per user group (overall population, obesity/overweight population, athletes, iron deficiency, type 2 diabetes) an iterative process of engineering the ontology and presenting it to experts for revision was followed, which resulted in the first version of NAct.

3.3. Conceptualisation: NAct in Depth

This phase deals with the glossary of terms that comprise of the core ontology vocabulary, identifying all the useful domain knowledge and its semantics, as well as the inference rules that will guide a personalized food and activity recommendation system. This vocabulary and rules were a product from the crystallization of all the information that was gathered from the databases and other ontologies and relevant vocabularies examined, but also from the important intangible knowledge offered by the experts in the dedicated virtual and physical workshops held between the experts and the ontology engineers.

To this end, Figure 1 represents the top level concepts of the NAct ontology.

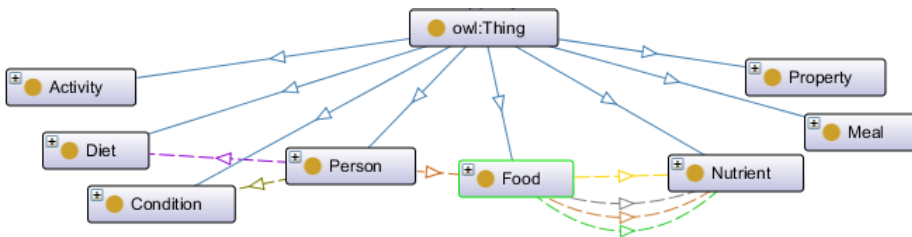


Figure 1. The top level concepts of the NAct ontology.

Activity (Fig. 2) models a hierarchy of physical activities. This hierarchy was inspired by the Compendium of Physical Activities [7], as well as by the activities of the HeLiS ontology, while it was revised by domain experts as per its compliance to European well-being directives and minimised to the optimal granularity through collaboration of ontology engineers with domain experts.

Condition (Fig. 2) covers the main medical conditions pertaining to the specific patient user groups of PROTEIN, i.e. cardiovascular disease, diabetes and obesity. In addition, some other prominent conditions were included for the general population. Most importantly, a complete set of allergies, intolerances and deficiencies were modelled, to cover the most important dietary and exercise restrictions and needs for all users employing a healthy lifestyle directives recommendation system.

Diet (Fig. 2) includes a set of dietary restrictions that may affect the food choices of users. It was decided by the domain experts that the ontology's focus should not align with preferential (e.g. Mediterranean) or commercial/popular (e.g. Atkins) diets, but rather maintain a high individualisation level per each user and their respective needs, i.e. combining preferences and needs in a flexible way rather than relying on diet templates. For this reason, only particular lifestyle or condition-related choices (e.g. vegan, halal) that provide specific dietary restrictions were modeled.

Meal and Person (Fig. 2) consist of basic classifications of meal types (e.g. breakfast) and of users (e.g. overweight adult). The former serves as a filter for the final decisions of the reasoning-based nutrition and activity AI advisor. The latter correlates to specific nutritional and well-being guidelines, as defined in particular axioms (described further on).

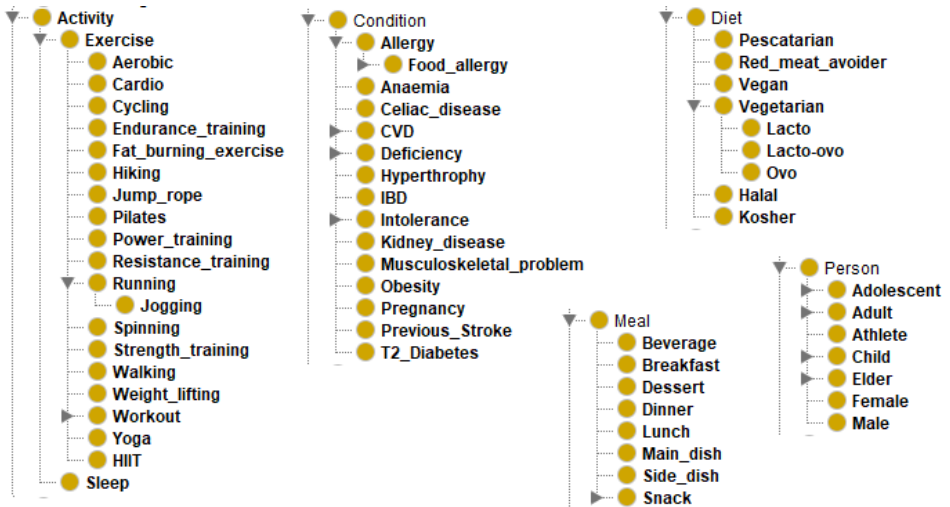


Figure 2. Activity, Condition, Diet, Meal and Person.

Food (Fig. 3) comprises a non-exhaustive hierarchy of principle foods. This is the main point where an important trade-off needed to be made in comparison to the plurality of detailed variations of foods that exist in existing food databases: the granularity must not be too deep, rather the most universally commonly ingredients of meals need to be included in their basic form, and for all those primary components not included, comprehensive food categories need to be available, so that undefined meal ingredients can be classified under the categories. Only a minimal set of *compound foods*⁵, common

⁵By compound foods, we denote foods that pertain a composition of basic ingredients.

in European diets (e.g. pasta, bread) or popular in particular lifestyle diets (e.g. seitan, falafel for vegetarians) were modeled. This vocabulary was primarily inspired by the McCance and Widdowson database [6], while engineers have taken into account the related HeLiS and FOKB classes, while the final sub-hierarchy was supervised and adapted by the experts based on the European Commission's *Food-Based Dietary Guidelines* [18].

Nutrient (Fig. 3) is a crucial sub-hierarchy in NAct. This sub-hierarchy was constructed based on the directives of the European nutritional guidelines [19]. It serves as the means to correlate specific foods and food groups with nutrients and subsequently determine the most and least nutritionally valuable meals per each individual user as per their specific (explicit) preferences and conditions. What drives the personalization system's decisions under NAct's scope is the finite set of nutrients, not an exhaustive list of foods and a voluminous set of instances denoting properties of each individual food, thus boosting both the system's flexibility as well as the recommender's computational efficiency.

A well-structured and meticulous hierarchical structure for both foods and nutrients was imperative in the scope of achieving NAct's purposes. The correlation among foods, food super-groups, nutrients and nutrient super-groups, conditions and goals is the core for determining the suitability of meals for each specific personalized nutrition application user.

Lastly, **Property** (Fig. 3) contains several types of important properties that need to be correlated with relevant nutritional and activity suggestions, like activity properties (e.g. level and intensity of activities, food and nutrient properties, food attributes, cooking/ preparation styles, etc., but most prominently **Goals** and ways to ensure their achievement. Goals influence the core of the nutrition and activity AI advisor and were provided by the domain experts based on multi-disciplinary empirical evidence and observations.

3.3.1. Relations and Rules

Pivotal to the aforementioned correlations between concepts was the definition of a minimal and meaningful set of binary relations (i.e. object properties). These relations were used in rules that drive the reasoning-based advisor's inference process. Rules in NAct comprise GCIs⁶ and non-GCIs axioms.

The defined relations and an example of inference rules are displayed in Figures 4 and 5. Most relations are assigned with a domain and/or range that define the semantic relation they support. For instance, the property "*highIn*" has *Food* as domain and *Nutrient* as range. This means that *Foods* (and only foods) may be *highIn* one or more *Nutrients* (and only nutrients).

One of the most important set of object properties is the food-to-nutrient relations (*highIn*, *lowIn*, *containsNutrient*) and subsequent rules. In order to gather knowledge regarding these rules, the nutritional correlation of all ontology foods was examined following the European Commission's Food Claims⁷. Consequently, relevant inference rules were automatically extracted based on concentration of nutrients in relevant foods and food types.

⁶General Concept Inclusion

⁷https://ec.europa.eu/food/safety/labelling_nutrition/claims/nutrition_claims_en

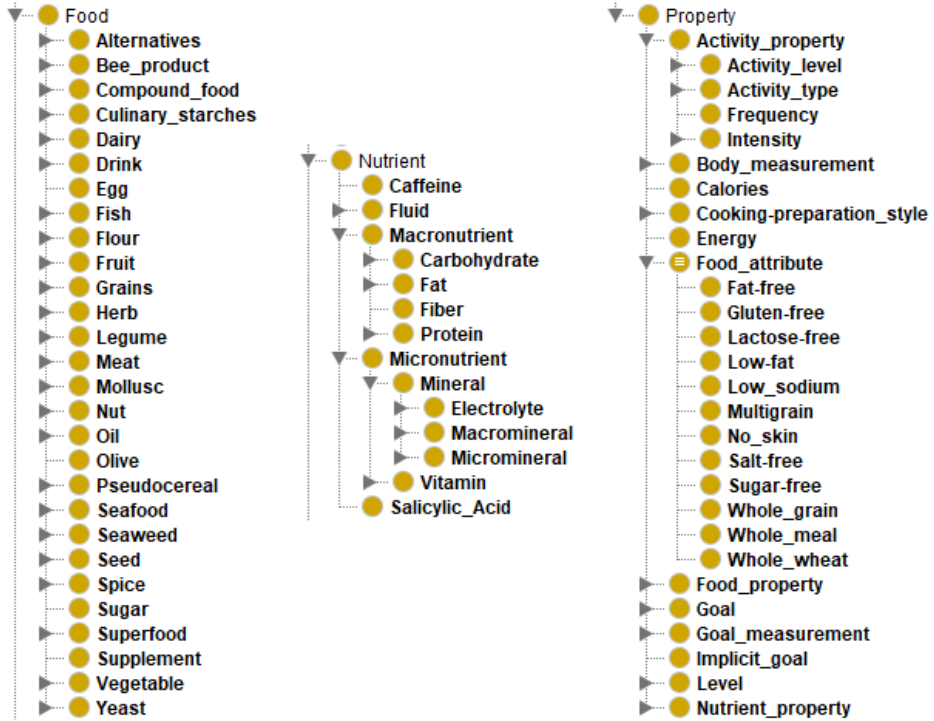


Figure 3. Food, Nutrient and Property.

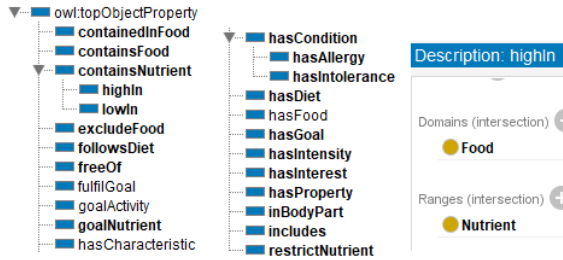


Figure 4. NAct relations.

General class axioms:

- Banana_allergy and (excludeFood some Banana) SubClassOf owl:Nothing
- Caffeine_intolerance and (restrictNutrient some Caffeine) SubClassOf owl:Nothing
- Calcium_deficiency and Magnesium_deficiency and Potassium_deficiency and Sodium_deficiency SubClassOf Electrolytes_deficiency
- Celery_allergy and (excludeFood some Celery) SubClassOf owl:Nothing
- Crustacean_allergy and (excludeFood some Crustacean_shellfish) SubClassOf owl:Nothing
- Egg_allergy and ExEgg SubClassOf owl:Nothing
- ExCod_liver_oil and Vegan SubClassOf owl:Nothing
- ExCod_liver_oil and Vegetarian SubClassOf owl:Nothing
- ExCrustShellfish and Kosher SubClassOf owl:Nothing
- ExDairy and Ovo SubClassOf owl:Nothing

Figure 5. NAct GCIs.

Several rules have been created based on obvious correlations (e.g. foods that cause specific allergies) or more implicit correlations defined by the experts, for each condition in relation to foods, nutrients, activities and their relevant properties, in order to ensure that the user will not be recommended with foods/activities that they need to avoid based on their conditions or that they need to consume more of and activities that they need to undergo less or more of. An example would be $Gluten \sqcap \exists restrictNutrient.Gluten.Intolerance \sqsubseteq \perp$, which restricts from a Gluten intolerant user's diet any foods that contain Gluten.

Another example is the set of axioms pertaining to Goals. Goals may be explicitly declared through the user profile or be implicitly derived from the inference engine, based on relevant rules that have been modelled within the ontology, e.g. $Iron.Deficiency \sqsubseteq \forall hasGoal.IncreaseIronIntake$. This rule defines that every person that has iron deficiency always has an implicit goal to increase their iron intake.

Based on the above, the inference engine will promote Foods that are defined to be high in the nutrient Iron, since another rule included in the ontology is that $IncreaseIronIntake \sqsubseteq \forall highIn.Iron$.

Lastly, goals contain even more complex rules, such as $Adult \sqcap Athlete \sqcap Muscle_gain \sqsubseteq \forall highIn.Protein \sqcap \forall highIn.Carbohydrates$, which denote that if the user is an Adult and an Athlete, and their (explicitly declared in the user profile) goal is Muscle Gain, then they should increase Protein intake and Carbohydrates intake (therefore consume more foods that are rich in these nutrients).

3.4. Integration: NAct in Action

NAct has been integrated with the knowledge-based expert system of the PROTEIN project, namely the AI Advisor, which employs the LiFR fuzzy reasoner for inferring the optimal meals, restaurant menu items and physical activities to recommend to a given user, based on this user's dietary and medical profile.

The recommendation system matches the explicit user-declared profiles (per user) against all possible meal and activity options available in the PROTEIN system, taking into account the nutritional, biomedical and physical activity background knowledge modeled in NAct.

In order to achieve this, semantic profiles of recommendation candidates (meals/restaurant menu, activities) and user profiles are automatically created from the list of ingredients of available meals and activities, as well as from the list of each user's dietary and medical premises. These profiles add candidate- and user-pertinent axioms to the TBox (most of the TBox comprising of NAct), while providing the ABox to complete the matching problem's KB.

These semantic profiles' purpose is dual: (a) transform the meals/activities and user profiles into reasoner-understandable formalizations, but most importantly (b) impose implied concept and relation instances, beyond the ones explicitly available in the profiles, in order to instigate a query process in the inference mechanism based on the ontology's model. An example of a semantic candidate (meal or activity) profile is shown in Table 1. Similarly, a semantic user profile example can be seen in Table 2.⁸

⁸It should be noted that LiFR supports fuzzy concept assertions, therefore it can accept concept instance degrees such that $\langle a : C \bowtie d, d \in [-1.0, 1.0] \rangle$ and preference weights $w \cdot C, w \in [0.0, 1.0]$. In crisp cases, $d \geq 1.0$ and $w = 1.0$ is implied. Such clauses will be used in the ABoxes and the inference examples that will follow further on.

Table 1. Semantic candidate profile example.

$\exists \text{includes} . (\text{Spinach} \sqcap \dots) \sqsubseteq \text{Breakfast_C2_1500}$	Constituents of the candidate. In this case, a conjunction of this meal's ingredients
$\langle \text{sp_1} : \text{Spinach} \rangle$	Instance <i>sp_1</i> of type Spinach
$\langle \text{candidate}, \text{sp_1} : \text{includes} \rangle$	The candidate meal contains <i>sp_1</i> of type Spinach (for preference check)
$\langle \text{sp_1}, \text{nutr} : \text{containsNutrient} \rangle^7$	Look for the nutrients that all classes which <i>sp_1</i> asserts contain
$\langle \text{sp_1}, \text{nutr} : \text{highIn} \rangle^7$	Look for what nutrients all classes that <i>sp_1</i> asserts are high in
$\langle \text{sp_1}, \text{nutr} : \text{lowIn} \rangle^7$	Look for what nutrients all classes that <i>sp_1</i> asserts are low in
$\langle \text{ingr}, \text{sp_1} : \text{containsFood} \rangle$	Look for compound foods that contain, as ingredient(s), all classes that <i>sp_1</i> asserts
$\langle \text{user}, \text{sp_1} : \text{excludeFood} \rangle$	Look for asserted premises for which all classes that <i>sp_1</i> asserts must be excluded

Table 2. Semantic user profile example

$\exists \text{hasInterest} . (\text{Vegetable} \sqcap \dots) \sqsubseteq \text{uid75}$	What the user likes to eat or do (activity-wise)
$\exists \text{hasInterest} . (\text{Yoghurt} \sqcap \dots) \sqsubseteq \text{uid75_dis}$	What the user doesn't like to eat or do (activity-wise)
$\text{uid75} \sqcap \text{uid75_dis} \sqsubseteq \perp$	Disjoint user likes and dislikes
$0.89 \cdot \text{Vegetable}$	Preference weight
$\exists \text{fulfilGoal} . \text{ImplicitGoal} \sqsubseteq \text{uid75}$	Default user profile axiom: search for implicit goals inflicted by user premises
$\langle \text{user} : \text{uid75} \rangle$	User instance
$\langle \text{user} : \text{Iron_Deficiency} \rangle$	User has iron deficiency
$\langle \text{user} : \text{Banana_Allergy} \rangle$	User has banana allergy
$\langle \text{user}, \text{goal} : \text{hasGoal} \rangle^8$	Look for goals that can should fulfilled for this specific user
$\langle \text{nutr}, \text{goal} : \text{goalNutrient} \rangle^8$	Look for which goals an asserted nutrient can fulfil
$\langle \text{act}, \text{goal} : \text{goalActivity} \rangle^8$	Look for which goals an asserted activity can fulfil
$\langle \text{candidate}, \text{goal} : \text{fulfilGoal} \rangle^8$	Look if the candidate (meal, activity) fulfils a goal
$\langle \text{user}, \text{nutr} : \text{restrictNutrient} \rangle$	Look if there is any premise inferred that causes the restriction of a nutrient for this user

3.4.1. Usage Examples

This section details the main test scenario of the inference process, validating the capacity of NAct to yield appropriate recommendations and restrictions based on of user-related information.

⁷It is anticipated that for the next expansion of LiFR, fuzzy relation assertions and weighted relations will be included and *highIn*, *lowIn*, *containsNutrient* will be assigned with different weights and thus assertions for them will result to different entailment degrees in the inferred model. Until then, only the *highIn*, *containsNutrient* instances are actually included in the employed ABox, as they denote a significant impact of a nutrient in particular user goals or deficiencies.

⁸In the premises of the recommendation problem at hand, it was decided that any goal holding true for the candidate is sufficient to produce a match, therefore only one *goal* instance is employed. If one wants to use NAct to discern between fulfilled goals, we encourage using an enumeration of *goal* instances, such as *goal_1*, *goal_2*, etc. for each explicit goal or medical condition that is included in the user profile. To this end, the user profile must include a set of the referenced relation instances, one per *goal_X* instance.

Table 3. Example of candidate fulfilling user preference

Ontology axioms	User premises	Candidate facts
$Spinach \sqsubseteq Vegetable$	(1a) $\exists hasInterest.Vegetable$	$\exists includes.Spinach \sqsubseteq$
$includes^- \equiv hasInterest$	(1b) $\sqsubseteq uid75$	$Breakfast_C2_1500$ (3a)
	$0.89 \cdot Vegetable$	$\langle sp_1 : Spinach \rangle$ (3b)
	$\langle candidate, sp_1 : includes \rangle$ (2c)	
Inference		
(1a) : \Leftrightarrow	$vegetable(x) \leftarrow spinach(x)$	
\therefore (3b)	$vegetable(x) \leftarrow spinach(sp_1)$	
\models	$vegetable(sp_1)$ $\Leftrightarrow vegetable(sp_1) \geq 1.0$	(4) The meal contains a vegetable
(2b) : \Leftrightarrow	$vegetable(x) \geq 1.0 \cdot 0.89$	
\therefore (4)	$vegetable(sp_1) \geq 1.0 \cdot 0.89$	
\models	$vegetable(sp_1) \geq 0.89$	(5) The fact that the inferred model contains a vegetable is important to the user by 0.89
(1b) : \Leftrightarrow	$\begin{cases} includes(x,y) \leftarrow hasinterest(x,y) & (1bi) \\ hasinterest(x,y) \leftarrow includes(x,y) & (1bii) \end{cases}$	
[1bii] : \therefore (2c)	$hasinterest(x,y) \leftarrow includes(candidate, sp_1)$	
\models	$hasinterest(candidate, sp_1)$	(6) The user may be interested in a candidate that includes sp_1 , i.e. a spinach instance
(2a) : \Leftrightarrow	$uid75(x) \leftarrow hasinterest(x,y), vegetable(y)$	
\therefore (5), (6)	$uid75(x) \leftarrow hasinterest(candidate, sp_1), vegetable(sp_1) \geq 0.89$	
\models	$uid75(candidate) \geq 0.89$	The given candidate meal satisfies the user profile, with a suitability degree of 0.89

One aspect pertains to meals and/or activities that *should* be recommended to a given user, because they may satisfy the user's preferences (Table 3) or because they may satisfy a particular user goal (Table 4). In the subsequent examples, the DL (Description Logics) axioms and instances will be translated to propositional logic clauses demonstrating the inference process.

The other major aspect in NAct's usage pertains to *rejections* of foods and/or activities. Rejections are of the most important operations of the recommendation system. They determine whether a candidate must absolutely not be recommended or even presented to the user. They are evoked whenever a logical contradiction (refutation) occurs when reasoning over a candidate. This happens in two cases:

A. When an ingredient in a meal or a type of activity has been explicitly declared by the user as one of their disinterests.

B. When an ingredient in a meal or a type of activity comes with in contrast with one of the user's characteristics (e.g. meat in case of a vegetarian user).

C. When a nutrient of an ingredient in a meal or a property of an activity (e.g. high intensity running) is actively prohibited given the user's medical condition(s).

Table 4. Example of candidate fulfilling goal

Ontology axioms	User premises	Candidate facts
$Spinach \sqsubseteq \forall highIn.Iron$	(1a) $\exists fulfilGoal.ImplicitGoal$ $\sqsubseteq uid7$	(2a) $\exists includes.Spinach$ $\sqsubseteq Breakfast.C2.1500$ (3a)
$Iron_Deficiency \sqsubseteq$ $\forall hasGoal.IncreaseIronIntake$	(1b) $\langle user : Iron_Deficiency \rangle$	(2b) $\langle sp_1 : Spinach \rangle$ (3b)
$IncreaseIronIntake \sqcap$ $\exists goalNutrient.Iron$	$\langle user, goal : hasGoal \rangle$ (2c)	$\langle sp_1, nutr : highIn \rangle$ (3c)
$\sqsubseteq ImplicitGoal$	(1c) $\langle nutr, goal : goalNutrient \rangle$ (2d)	
	$\langle candidate, goal : fulfilGoal \rangle$ (2e)	
Inference		
(1a): \Leftrightarrow	$iron(y) \leftarrow highin(x, y), spinach(x)$	
\therefore (3b)	$iron(y) \leftarrow highin(spi_1, nutr), spinach(spi_1)$	
\models	$iron(nutr)$ (4)	The meal contains the nutrient Iron
(1b): \Leftrightarrow	$increaseironintake(y) \leftarrow$ $hasgoal(x, y), iron_deficiency(x)$	
\therefore (2b), (2c)	$increaseironintake(y) \leftarrow$ $hasgoal(user, goal), iron_deficiency(user)$	
\models	$increaseironintake(goal)$ (5)	The goal <i>increase iron intake</i> is inferred as true for this KB
(1c): \Leftrightarrow	$\left\{ \begin{array}{l} IncreaseIronIntake \sqcap A \sqsubseteq ImplicitGoal \text{ (1ci)} \\ \exists goalNutrient.Iron \sqsubseteq A \text{ (1cii)} \end{array} \right.$	
(1cii): \Leftrightarrow	$a(y) \leftarrow goalNutrient(x, y), Iron(x)$	
\therefore (2d), (4)	$a(y) \leftarrow goalNutrient(nutr, goal), iron(nutr)$	
\models	$a(goal)$ (6)	The nutrient needed to fulfill this goal holds true for this KB
(1ci): \Leftrightarrow	$implicitgoal(x) \leftarrow increaseironintake(x), a(x)$	
\therefore (5), (6)	$implicitgoal(x) \leftarrow$ $increaseironintake(goal), a(goal)$	
\models	$implicitgoal(goal)$ (7)	An implicit goal is satisfied for this user
(2a): \Leftrightarrow	$uid75(x) \leftarrow fulfilgoal(x, y), implicitgoal(y)$	
\therefore (2e), (7)	$uid75(x) \leftarrow fulfilgoal(candidate, goal), implicitgoal(goal)$	
\models	$uid75(candidate)$	The given candidate meal satisfies the user profile; the suitability degree is implied to be 1.0

Due to length restrictions, a complete rejection example will not be detailed. Axioms in the ontology that imply \perp (owl:Nothing), e.g. $Banana_allergy \sqcap \exists excludeFood.Banana \sqsubseteq \perp$ and $Gluten_Intolerance \sqcap \exists restrictNutrient.Gluten \sqsubseteq \perp$ are designed exactly to cause such refutations whenever relevant foods, nutrients, activities or other properties that come in contrast with the user profile are inferred.

In the same mentality, in terms of user preferences, the disjointness axiom $uidX \sqcap uidX_dis \sqsubseteq \perp$ of Table 2 is employed to cause such refutations. Therefore, whenever a meal or activity fulfills a user interest (with one or more ingredients for the meal case),

while at the same time another candidate (e.g. ingredient) fulfills the disinterests, the reasoner will issue a refutation, causing for the said meal to be rejected for this user from the list of candidates.

3.5. Implementation

NAct is an OWL ontology, falling in the OWL 2 RL expressivity fragment, as mentioned before. Thus it leverages rich expressivity and computational efficiency in order to enable robust logic-based inferencing for content recommendation, but at the same time reduce the computational cost. Throughout its lifecycle, it has been engineered using the Protégé¹¹ ontology editor.

3.6. Evaluation

Based on a pre-defined pool of >1400 expert-defined meals and >50 physical activities of different intensity levels available in the PROTEIN platform, experiments were held using the LiFR reasoner with over 70 synthetic user profiles including one or more allergies, deficiencies, intolerances, diet choices and medical conditions, with several combinations thereof, in a pre-release phase of the first version of the PROTEIN system. Meals, activities and user profiles were semantically transcoded as described in Section 3.4.

Through these experiments, NAct has been validated technically in terms of logical Consistency and Completeness, Soundness and Decidability as well as of Computational Efficiency [20], [21]. The tests were held both by technical staff as well as the domain experts, simulating their patients and clients.

NAct has been found to be decidable (sound & complete) - complete in the sense that any expression that is logically implied by the KB¹² that includes NAct and the meal/activity and user profiles as previously described, can be derived. It is also consistent - in the sense that only purposeful contradictions arise during the reasoning process.

In terms of computational efficiency, the results depend on the respective high computational efficiency of the LiFR reasoner, as described in [17] and vary according to the computational capacities of the machine that runs the inference service. In any case, memory consumption is insignificant (re LiFR), while matching a single meal's semantic profile with a given user's semantic profile on top of NAct takes 1-3 seconds on a Intel Core i5 on 3.3GHz, depending on the number of instances in the meal and user profiles.

NAct however pends validation in the ongoing PROTEIN pilots in terms of Consistency, Completeness¹³ and Conciseness, to what it concerns fully covering the well-being recommendation needs of the users of the project.

4. Availability and Documentation

NAct is publicly available under the Creative Commons Attribution-ShareAlike License (version 3.0)¹⁴, under a persistent PURL URI, namely <http://purl.org/nact>. The

¹¹<https://protege.stanford.edu/>

¹²As per the definition of Logical Completeness of [21]

¹³In the sense of recall

¹⁴<http://creativecommons.org/licenses/by-sa/3.0/>

ontology is published on GitHub, in a dedicated project and repository¹⁵, while the ontology specification and documentation (LODE [22] version) web page¹⁶ will be permanently maintained through GitHub pages.

Two PROTEIN project deliverables serve as the means to document the first version and subsequent evolutions of the PROTEIN ontology. Moreover, technical documentation was provided by means of the OWLDoc¹⁷ ontology documentation producing tool.

Furthermore, in NAct's website a public summation of the developments of each release is maintained, accompanied by formal documentation of the ontology's contents produced via the LODE [22] tool. The OWLDoc documentation is also available on the site.

5. Conclusions and Future Work

This paper presented the first version of the novel NAct ontology, which innovatively combines evidence-based and consolidated EU standards-based nutritional, medical and preferential elements for advanced individualization of meal and physical activity recommendations in an intelligent AI-based healthy lifestyle system.

As the work presented comprises the first version of such an expert-based system, only having undergone synthetic trials and expert evaluation, and is yet to be tested in real-world pilots, evolution of the ontology is expected in the near future.

It is for example a known fact to the engineers and experts that not all prominent inference rules that can be modeled for the domain in question are included in this first version of the ontology, however the first pilots aim to reveal redundancies of the first version and pinpoint the most important rules that have not yet been included in the system. One major action point for experts and ontology engineers, taking place before summer 2021, will be to add several more relevant rules to the ontology relating more medical conditions to physical activities and their properties (e.g. intensity).

In addition, later extensions will also delve in formally defining the semantics of the modelled entities, as well as in providing mappings to entities of similar semantics in seminal related ontologies and/or vocabularies.

The ontology engineers and domain experts will continue their collaboration to extend and revise the novel ontology - at least - throughout the PROTEIN project's lifecycle, following own observations while using the system, but most importantly based on end users' evaluation in the first pilots.

Acknowledgments

This work has received funding from the European Union's Horizon 2020 research and innovation programme, grant *n*^o 817732, PROTEIN ("PeRsOnalized nutriTion for hEalthy livINg").

¹⁵<https://github.com/nutritionactivityontology/nact>

¹⁶<https://nutritionactivityontology.github.io/nact/>

¹⁷<https://protegewiki.stanford.edu/wiki/OWLDoc>

References

- [1] (EFSA) EFSA. Tools for critically appraising different study designs, systematic review and literature searches. EFSA supporting publications. 2015;12(7):836E.
- [2] Hariton E, Locascio JJ. Randomised controlled trials—the gold standard for effectiveness research. *BJOG: an international journal of obstetrics and gynaecology*. 2018;125(13):1716.
- [3] Schwingshackl L, Chaimani A, Hoffmann G, Schwedhelm C, Boeing H. A network meta-analysis on the comparative efficacy of different dietary approaches on glycaemic control in patients with type 2 diabetes mellitus. *European journal of epidemiology*. 2018;33(2):157-70.
- [4] Theodoridis T, Solachidis V, Dimitropoulos K, Gymnopoulos L, Daras P. A survey on AI nutrition recommender systems. In: *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*; 2019. p. 540-6.
- [5] Wilson-Barnes S, Gymnopoulos L, Dimitropoulos K, Solachidis V, Rouskas K, Russell D, et al. *PeRsonalised nutriTion for hEalthy living: The PROTEIN project*. *Nutrition Bulletin*. 2021;46(1):77-87.
- [6] McCance RA, Widdowson EM. *McCance and Widdowson's the Composition of Foods*. Royal Society of Chemistry; 2014.
- [7] Shephard R. compendium of physical activities: a second update of codes and MET values. *Yearbook Of Sports Medicine*. 2011;2012:126-7.
- [8] Dooley DM, Griffiths EJ, Gosal GS, Buttigieg PL, Hoehndorf R, Lange MC, et al. FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration. *npj Science of Food*. 2018;2(1):1-10.
- [9] Vitali F, Lombardo R, Rivero D, Mattivi F, Franceschi P, Bordoni A, et al. ONS: an ontology for a standardized description of interventions and observational studies in nutrition. *Genes & nutrition*. 2018;13(1):1-9.
- [10] Castellano-Escuder P, González-Domínguez R, Wishart DS, Andrés-Lacueva C, Sánchez-Pla A. FOBI: An ontology to represent food intake data and associate it with metabolomic data. *Database*. 2020;2020.
- [11] Andrés-Hernández L, Baten A, Azman Halimi R, Walls R, King GJ. Knowledge representation and data sharing to unlock crop variation for nutritional food security. *Crop Science*. 2020;60(2):516-29.
- [12] Yang C, Ambayo H, De Baets B, Kolsteren P, Thanintorn N, Hawwash D, et al. An ontology to standardize research output of nutritional epidemiology: from paper-based standards to linked content. *Nutrients*. 2019;11(6):1300.
- [13] Bordea G, Nikiema J, Griffier R, Hamon T, Mougin F. FIDEO: Food Interactions with Drugs Evidence Ontology. In: *11th International Conference on Biomedical Ontologies*; 2020. .
- [14] Çelik D, Elçi A, Akcicek R, Gökçe B, Hürcan P. A safety food consumption mobile system through semantic web technology. In: *2014 IEEE 38th International Computer Software and Applications Conference Workshops*. IEEE; 2014. p. 348-53.
- [15] Dragoni M, Bailoni T, Maimone R, Eccher C. Helis: An ontology for supporting healthy lifestyles. In: *International Semantic Web Conference*. Springer; 2018. p. 53-69.
- [16] Fernández-López M, Gómez-Pérez A, Juristo N. *Methontology: from ontological art towards ontological engineering*. 1997.
- [17] Tsatsou D, Dasiopoulou S, Kompatsiaris I, Mezaris V. LiFR: A lightweight fuzzy DL reasoner. In: *European semantic web conference*. Springer; 2014. p. 263-7.
- [18] EFSA Panel on Dietetic Products, Nutrition, and Allergies (NDA). Scientific opinion on establishing food-based dietary guidelines. *EFSA Journal*. 2010;8(3):1460.
- [19] (EFSA) EFSA. Dietary reference values for nutrients summary report. Wiley Online Library; 2017.
- [20] Gómez-Pérez A. Evaluation of ontologies. *International Journal of intelligent systems*. 2001;16(3):391-409.
- [21] Obrst L, Ceusters W, Mani I, Ray S, Smith B. The evaluation of ontologies. In: *Semantic web*. Springer; 2007. p. 139-58.
- [22] Peroni S, Shotton DM, Vitali F. Making Ontology Documentation with LOD. In: *I-SEMANTICS (Posters & Demos)*. Citeseer; 2012. p. 63-7.