Towards a teenager tailored ontology
— Supporting inference about the obesity-related health status —
Aleksandra Sojic1, Walter Terkaj1, Giorgia Contini1, and Marco Sacco1
1Istituto Tecnologie Industriali e Automazione (ITIA) CNR, Milano, Italy

ABSTRACT
In this paper we outline the general framework of the ontology that captures the obesity-related features of teenagers. We present our particular choices regarding the ontology-structure, which should be capable of capturing (1) multiple perspectives used to describe an adolescent and (2) reasoning about the individual changes during the time. In the same line, we address several issues related to the modelling of normative concepts related to obesity and depict how the public health concern impacts classification of teenagers according to their phenotypes. In particular, we present a fragment of the ontology that supports inference about individuals and a personalised assessment of the obesity-related health conditions.

1 INTRODUCTION
Overweight and obesity are estimated to result in the deaths of about 320 000 people in western Europe every year [15]. The prevalence rates of obesity among children and adolescents motivated the public health organisations to engage in the promotion of a healthy life style [15]. Nonetheless, the understanding of the causal links between obesity and numerous socio-behavioural aspects of life-style is a complex task that involves multiple domains of knowledge and the analysis of heterogenous kinds of data. This problem can be addressed by adopting Semantic Web technologies, in particular ontologies, that are recognised as a convenient approach to deal with complex and heterogeneous information across various domains, enable knowledge generation via reasoning and support data interoperability [22, 10]. Several studies report on the use of ontology and semantic technologies to target obesity (e.g. [21, 20]). Scala et. al [20] developed an e-Knowledge platform, based on an OWL1 ontology and SWRL [9] rules, classifying individuals according to the obesity level and certain medical conditions (Sarcopenia, Hypertension, Dyslipidemia, Diabete, Insulin resistance, Metabolic syndrome). Arash et. al [21] present the preliminary stage of the ontology designed to support a knowledge-based infrastructure, promoting healthy eating habits and lifestyles. While [20] focuses on adults and [21] on children, in this paper we present the current state of the ontology that captures formally obesity-related knowledge in a teenager tailored model. The main ontological structure is formalised in OWL, specifying certain generic classes that are applicable to any human being and represent explicitly the time-dependent changes in health condition (i.e. the issue that was not addressed by [21, 20]). At a later stage, the ontology should be able to support the information flow and interoperability between the technological tools and platforms employed to monitor the changes of health status, behaviour, and nutritional habits of teenagers. The initial step in the ontology development includes a multi-disciplinary analysis that considers a teenager as a dynamic agent who is constantly changing in the interaction with his environment. The development of the teenager-centered ontology is initiated within the European research project named PEGASO 2, whose main goal is the enhancement of self-awareness and motivation of adolescents towards a healthy lifestyle [6, 2, 3]. The project is driven by the public health concerns aiming at the decrease of the obesity-related risks to health [6, 14]. The target population of the project are the future adults whose behavioural habits at an early age can significantly impact their health status on a life-long scale [2]. The project includes several research initiatives and interventional strategies, most of which go beyond the scope of this paper, e.g. the development of serious games [17, 6] that will promote an healthy life style, the design of a life companion [3], the use of wearable gadgets equipped with sensors to monitor health status [3, 6], the design of mobile applications such as a diary used to record dietary habits, etc. [3, 6].

The multidisciplinary studies of the interactions between a teenager and his environment should provide a comprehensive model, i.e. the so-called Virtual Individual Model (VIM) [2] that will be used as a

*Corresponding author: aleksandra.sojic@itia.cnr.it
1 http://www.w3.org/TR/owl2-overview/
2 http://pegasof4f.eu/
In order to diagnose and modify the health status of an individual it is necessary to take into account his/her current condition through a comprehensive model capable of representing the human being as a whole. Such a model is an abstract representation that aims at integrating the cross-disciplinary knowledge about an individual, his/her characteristics and relationships in a broad context. It aims to comprehensively represent all the components that influence the health status of a teenager that is related to overweight and obesity. The model (i.e. VIM) focuses on three distinct levels in order to characterise the states of an individual: (1) the physical-physiological level, (2) the nutritional level, and (3) the psychological level. All these various perspectives constitute an integral approach to the understanding of an individual, providing the theoretical framework to determine the individual’s health status in a dynamic manner. The theoretical model is considered as the initial ground that is used for the ontology development. The ontology model needs to specify further on a formal level the structure, concepts, and the most relevant relations that hold between them. Designing such an ontology model that integrates all the relevant perspectives is a challenging task. First, it involves collaboration of the experts with different backgrounds (philosophers, biomedical experts, clinicians, psychologists, computer scientists, and engineers), positioning the ontology in an interdisciplinary environment that is dealing with the terminological and conceptual problems in order to merge numerous (domain-specific) perspectives (see e.g. [22]). Second, the ontology design involves the decision on the ontology structure that will be crucial to its successful application. Currently, we consider the most convenient modularisation strategy [1, 4] to organise knowledge. At this point we have decided to divide the main ontology structure according to two criteria: (a) the criterion of disciplinary perspectives [6, 2, 3] specifying modules 1-5) and (b) the criterion of an integrative view (specifying modules 6-7). The resulting structure consists of the following modules:

1. Physical domain describing physical features of a teenager
2. Physical Activity and the related behaviour of a teenager.
3. Physiological domain
4. Psycho-social domain
5. Alimentary behaviour
6. Trends of change of an individual status that is formalised within (1-5). This module will be used to model the decisions about the interventional strategies and a personalised feedback to teenagers
7. Commons module to support the interoperability between the modules (1-6)

The ontology is currently at an early stage of development and has formalised only the segments relevant for the first module. The module distinguishes the relevant body features, linked to measurements and certain classes of health conditions that will be used to define the obesity-related status and potential risk factors. In the following subsection we present the links between the measurements characterising body constitution that are afterwards used to classify the corresponding health conditions.

Capturing normative concepts: assessment of obesity as a health condition

In general terms, a description of a teenager via some structural, functional, and behavioural characteristics is actually capturing aspects considered to be relevant to describe the teenager-specific phenotype. A phenotype is defined as a set of features of an organism that emerges as a result of interactions of his genetic material (specified as genotype) and the environment (see e.g.[11]). Herein, the genotype\(^3\) of a teenager is not considered explicitly. We focus on the phenotypic features describing the class to which a person belongs as determined by the description of his/her physical and behavioural characteristics [11]. Thus, we consider that a person's phenotype belongs to the class \textit{obese} based on his characteristics, description of which (despite of individual variations) fits to the description of an obese phenotype that is typical for every person of a certain age, gender, and with a specific body mass index.

We define typical features of an 	extit{obese} phenotype in terms of a conventional agreement at the current stage of knowledge. The reference system that we use to characterise the physical features of an obese phenotype is provided by the World Health Organisation [14] and it includes the age and gender specific ranges of values, e.g. body mass index of a teenager (see [16]). Thus, we treat the description of body constitution as a specific characterisation of phenotype that is associated with health condition.

Figure 1 presents the hierarchy of the relevant health conditions, specifying the physical constitution that considers adiposity, body fat distribution, body mass, and central obesity. Each of the conditions is associated with a specific classification and linked to the reference values that characterise physical features relative to gender and age [16]. These classifications are distinct as they are using diverse criteria to describe a condition of body constitution.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Fig1.png}
\caption{A fragment of the obesity-related classes used for the assessment of a personal health condition that is directly dependent on gender and age at the time of assessment.}
\end{figure}

For instance, the criterion of body mass (provided as body mass index [16]) in one of the classifications is used to distinguish people as belonging to one of the following groups: 	extit{obese, underweight, overweight} or \textit{normal weight} [14]. According to the classification that considers fat distribution, a person may be classified either as \textit{android} or as \textit{gynoid}. In the following section we provide an example of the personalised assessment of the obese condition according to the measure of body mass index, relative to gender and age at the time of assessment.

A fragment of the ontology

An ontology as an artefact is not intended to cover the world in its entirety, but only chosen aspects of the world, on specific levels of abstraction, and for given purposes. Thus, we present here a ‘simplified view of the world that we wish to represent’[5] while modelling obesity-related knowledge in a declarative formalism. We defined the scope of our universe of discourse that is relevant for the goals of our formal model, i.e. (1) capturing the physical characteristics of a person and then (2) evaluating the health condition specifically for a teenager. The former goal was reached by developing an OWL ontology, whereas the latter by defining SWRL rules. The class hierarchy of the OWL ontology consists of two key classes: 	extit{Person} and 	extit{HealthCondition}. These two classes are linked by a restriction involving the object property 	extit{isInHealthCondition}, so that a 	extit{person} can be associated with one or more health conditions (see Figure 2) while aiming at tracking the evolution of the health condition over time in a dynamic context. The classification of a generic health condition as belonging

\footnote{The class to which a person belongs as determined by the description of his/her physical material made up of DNA passed to the organism by his parents.}
Fig. 2. Depicting the general structure of the classes and relations used to support the reasoning over instances, e.g. the orange arrow (bold dashed) stands for an inferred relation that classifies TomCondition1 as an ObeseCondition.

Fig. 3. The example of the rules used to support reasoning about a personalised assessment of health condition, e.g. TomCondition1 is ObeseCondition (inferred by the Pellet reasoner in Protégé.)

to one the HealthCondition subclasses (see Figure 1) is performed by means of SWRL rules that makes use of (1) physical (structural) and functional (metabolic, etc.) features, (2) gender, and (3) age of a person.

The physical and functional features are directly associated with a health condition, e.g. the body mass index is defined using the data property isCharacterizedByBodyMassIndex (see Figures 2 and 3).

The gender is defined by instantiating a Person as belonging to one of its subclasses, i.e. classes Male and Female.

The exact age of a person when the health condition is evaluated is a crucial information because the reference values for the assessment are particularly varying in the adolescence when body grows and changes [16]. In order to capture this variability that can impact on the assessment, we associate Person with the birth date and HealthCondition with the date of assessment (year and month) by using the following object properties and restrictions:

\[
\text{Class: Person} \\
\text{SubClassOf: } \text{isInHealthCondition only } \text{HealthCondition} \\
\text{isBornInYear only integer}
\]
Having the data related to the date of birth and time of assessment, we can apply a rule modeled in SWRL[9] in order to get an age value associated with a personal condition assessment (see Figure 3), so that all the needed elaborations can be performed by a reasoning tool without needing to interface with other applications.

We tested our ontology by instantiating the above mentioned classes and then running a reasoner (i.e. Pellet reasoner plug-in for Protege®) to properly classify the health conditions. The example in Figure 3 shows an instance of a boy, named Tom, who is assigned two assessments of his health condition: TomCondition1 (assessed in November 2014), and TomCondition2 (assessed in November 2015). Since the reference value for the assessment of an obese condition changes with age (see the reference values in Figure 3), and in our example Tom’s body mass index stays unchanged, only TomCondition1 is inferred to be ObeseCondition. The same reasoning can be performed with the instances of teenagers of different age and gender. We also specify the reference values relevant for assessment of other obesity-related categories of health condition (e.g. provided as measure of waist circumference) by defining a total of 74 SWRL rules. As a comparison, the work Scala et al [20] contains approx. 40 rules.

The facts resulting from the reasoning can be saved into the ontology, thus actually enriching the knowledge. The ongoing work includes the extension of the ontology module with the formal specification of other relevant phenotypic features.

3 DISCUSSION AND FUTURE WORK

We presented the initial steps performed in the ontology design related to the PEGASO project. We depicted how a public health concern such as obesity impacts the decisions about the most relevant classes and relations that are used to represent phenotype of teenager and to integrate various perspectives and domain specific knowledge. Specifically, we presented the preliminary results of the ontology that (1) targets the population of teenagers, (2) formally captures physical aspects of phenotype, (3) classifies health conditions according to the physical constitution (obesity-related classes), (4) associates the condition of physical constitution with the personal assessment of the condition as age and gender dependent, (5) supports reasoning over instances, e.g. individual teenagers that may have assigned diverse conditions at different time points (age dependent assessment).

We are currently working on the design of an interlinked modular structure. Since our target is the phenotype of teenagers, in the future work we will consider to link our ontology with relevant phenotype ontologies5. In addition, we aim at joining the efforts in creating alignments [19, 7, 13] between our ontology and the reference terminologies and ontologies5, as well as foundational ontologies7.

ACKNOWLEDGEMENTS

This work has been partially funded by the EU 7th Framework Programme under the grant agreements No: 610727, “Personalised Guidance Services for Optimising lifestyle in teenagers through awareness, motivation and engagement” (PEGASO). The project is compliant with the European and National legislations concerning the user safety and privacy. We would like to thank all the partners in the project, in particular Claudio Lafortuna, Giovanna Rizzo, and Sarah Tabozzi. We would especially like to thank Laura Cruz for her contribution during the early stages of the project.

REFERENCES


---

5 http://bioportal.bioontology.org/ontologies/MF
http://www.human-phenotype-ontology.org/
6 http://ncit.nci.nih.gov/
http://bioportal.bioontology.org/ontologies/HL7,
http://www.who.int/classifications/icd/en/,
http://bioportal.bioontology.org/ontologies/UO,
http://obi-ontology.org/

---

4 http://clarkparsia.com/pellet/protege/


