Extending HeLiS: From Chronic Diseases to Behavior Change

Mauro Dragoni and Valentina Tamma

Fondazione Bruno Kessler, Trento, Italy
dragoni@fbk.eu

Abstract. The use of knowledge resources in the digital health domain is a trending activity significantly grown in the last decade. In this paper, we present two extensions of HeLiS, an ontology aiming to provide in tandem a representation of both the food and physical activity domains and the definition of concepts enabling the monitoring of users' actions and of their unhealthy behaviors. The presented extensions focus on two aspects: the modeling of relationships between food categories and the most common chronic diseases and a top-level layer representing the barriers that users may encounter during the self-management of their lifestyles and/or their chronic diseases. We describe the construction process and the main concepts that have been included.

1 Introduction

Chronic diseases, such as heart disease, cancer, and diabetes, are responsible for approximately 70% of deaths among Europe and U.S. each year and they account for about 75% of the health spending[1],[2]. Such chronic diseases can be largely preventable by eating healthy, exercising regularly, avoiding (tobacco) smoking, and receiving preventive services. Prevention at every stage of life would help people stay healthy, avoid or delay the onset of diseases, keep diseases they already have from becoming worse or debilitating; it would also help people lead productive lives and, at the end, reduce the costs of public health.

People can start their own prevention process by simply monitoring their lifestyles, in terms of dietary habits and physical activities they do. In order to support this, structured resources able to combine all information and to support the integration of monitoring facilities have to be developed. Besides the domain knowledge needed for representing with a semantic flavor foods' composition and the effort level of physical activities for measuring the compliance between a user profile and the rules associated with it, there are two other aspects enabling the realization of smart platforms. First,

[1] http://www.who.int/nmh/publications/ncd_report_full_en.pdf

[2] https://www.cdc.gov/media/releases/2014/p0501-preventable-deaths.html

knowledge representing how much a specific food category can affect the onset or the worsening of a specific disease. Second, knowledge about the barriers (physical of psychological) that a user may have concerning the adoption of healthy habits.

In this paper, we presents two extensions of the HeLiS ontology [1] [3] representing the two aspects mentioned above. The first extension provides a conceptual model representing the risk level of each food categories already defined in HeLiS associated with the onset or worsening of the most common five chronic diseases (i.e. diabetes, kidney diseases, cardiovascular diseases, hypertension, and obesity). While, the second extension, provides an abstract layer of a conceptual model representing the barriers that a user may encounter during the self-management of his/her lifestyle or of his/her chronic disease (e.g. knowledge representing why a diabetes patient is not able to check his/her glycemia constantly).

The relevance of these extensions with respect to the state of the art pivots around the integrated model representing (i) a fine-grained representation of the links between food categories and chronic diseases and (ii) a top-level representation of the clinical (from the psychological perspective) barriers that be exploited for developing more fine-grained models supporting the realization of behavior change paths. Around these two extensions, the HeLiS ontology provides a flexible support to rules modeling that can be used for the reasoning on data provided by users. Besides the conceptual model per se, the HeLiS ontology represents a valuable resource for the healthcare domain thanks to the knowledge included into the provided resource.

The remain of the paper is organized as follows. Section 2 provides a brief overview of the main ontologies concerning the healthy lifestyle domain. In Section 3 we describe the methodology we followed for modeling the two extensions, while Section 4 shows the main entities of the conceptual model. Finally, Section 5 concludes the paper.

2 Related Work

We provide in this Section a brief summary of the most relevant work on ontologies describing both the food and the physical activity domains.

In [2] the authors describe food intake patterns identified by applying new food categories, in particular: (i) nutrient composition and energy density, (ii) current scientific evidence of health benefits, and (iii) culinary use of each food. In [3], a process is presented for a rapid prototyping of a food ontology oriented to the nutritional and health care domain that is used for sharing existing knowledge. However, unfortunately, this resource is no longer available.

The contribution presented in [4] discusses the design and development of a food-oriented ontology-driven system (FOODS), used for food or menu planning in a restaurant, clinic/hospital, or at home. FOODS comprises (i) a food ontology, (ii) an expert system using such an ontology and some knowledge about cooking methods and prices, and (iii) a user interface suitable for users with different levels of expertise. Its aim is to support the management of treatment plans for patients affected by type 1 or type 2 diabetes. Instead, the work presented in [5] focuses on the integration of different

[3] http://w3id.org/helis

domain ontologies, like food, health, and nutrition, in order to help personalized information systems to retrieve food and health recommendations based on the user's health conditions and food preferences. Recently, the work presented in [6] describes an ontology modeling the protected names of brands, from the raw materials to the production process.

A set of ontologies have been proposed that collect information about packaged food. Examples are Open Food Facts [4] and Food Product Ontology [7]. However, their focus on categorizing and describing packaged food led to low coverage of concepts describing food compositions.

Recently, the FoodOn ontology [5] has been released. This ontology represents foods from a different perspective with respect to the HeLiS ontology. Instead of focusing on food composition, they aim to realize a food description system that registers food manufacturers. Indeed, the FoodOn ontology includes, for each product, its origin, the physical attributes, processing, packaging, dietary uses and geographical origin.

Finally, in [8] the design steps are described, the working mechanism, and the case of use of the Ontology-Driven Mobile Safe Food Consumption System (FoodWiki) using semantic matching. This resource aims to address problems similar to the HeLiS ontology. However, no information about physical activities and their correlation with food categories are included in the ontology nor the possibility of modeling in a flexible way rules users should follow and the possible associated violations.

Concerning physical activity, we report two ontologies, both available through the BioPortal [6] website.

The first one is the SMASH (Semantic Mining of Activity, Social, and Health data) [7] ontology. The goal of the SMASH ontology is to describe concepts correlating physical activities and social networks. The system developed upon this ontology aims to sustain weight loss with continued intervention with frequent social contacts. The coverage of the SMASH ontology is very limited. Indeed, only 18 activities are defined.

The second one is the Ontology of Physical Exercises (OPE) [8]. Here, physical activities are modeled from the functional perspective. Thus, exercises are described in terms of movements, how the different musculoskeletal parts of the human body are engaged, and which are the expected health outcomes. Also in this case, the coverage of the physical activity domain is limited because only general categories of activities, like *AerobicExercise* or *IsotonicExercise*, are defined.

3 The HeLiS Ontology

The development of the HeLiS ontology followed the need of providing a knowledge artifact able not only to provide a representation of domains concerning healthier lifestyles, but, also, to support further activities like, for example, remote medical monitoring and support to behavior change. The two extensions presented in this paper go

[4] Open Food Facts. Available online: http://world.openfoodfacts.org/who-we-are
[5] http://foodontology.github.io/foodon/
[6] http://bioportal.bioontology.org/
[7] http://bioportal.bioontology.org/ontologies/SMASHPHYSICAL
[8] https://bioportal.bioontology.org/ontologies/OPE

into the same direction by extending the dietary part with information concerning the risk level of each food category with respect to the onset or worsening of a subset of chronic diseases and by presenting a representation of barriers that can obstruct persons to start of behavior change process for improving the quality of their lifestyle.

The process for building these two extensions followed, as happened for the HeLiS ontology, the METHONTOLOGY [9] methodology. This approach is composed by seven stages: Specification, Knowledge Acquisition, Conceptualization, Integration, Implementation, Evaluation, and Documentation. The overall process involved three knowledge engineers and four domain experts from the Trentino Healthcare Department.

The choice of METHONTOLOGY was driven by the necessity of adopting a lifecycle split in well-defined steps. The development of these HeLiS ontology extensions required the involvement of the experts in-situ. Thus, the adoption of a methodology having a clear definition of the tasks to perform was preferred. Other methodologies, like DILIGENT [10] and NeOn [11], were considered before starting the construction of the HeLiS ontology. However, the characteristics of such methodologies, like the emphasis on the decentralized engineering, did not fit our scenario well.

Specification And Knowledge Acquisition. The *Specification* and *Knowledge Acquisition* stages completely overlapped during the building of the two proposed extensions. The purpose of the HeLiS ontology is two-fold. On the one hand, we want to provide a detailed and integrated model of the food and diseases domain. On the other hand, we want to support the definition of behavior change process by taking in account the barriers that can affect users.

As for the HeLiS ontology, also the presented extensions have been modeled with a *high* granularity level. Concerning the representation of associations between food categories and diseases, we defined the *DiseaseRiskLevel* concept that allows to link each food category with their risk level for a specific disease. The risk level is represented by a datatype property representing such a risk level with an integer value. This knowledge was acquired directly from a set of focus groups run with domain experts.

While concerning barriers, we defined which are the main type of barriers that we want to consider for supporting the development of third-party behavior change applications. The conceptualization of barriers and of the different state of change has been created by extracting knowledge from domain specific unstructured resources [12].

Conceptualization. The conceptualization of the HeLiS ontology was split into two steps. The first one was covered by the knowledge acquisition stage, where most of the terminology is collected and directly modeled into the ontology. Examples are the diseases and the barriers. While the second step consisted in deciding how to represent, as classes or as individuals, the information we collected from unstructured resources. Then, we modeled the properties used for supporting all the requirements.

During this stage we relied on several ontology design patterns (ODP) [13]. However, in some cases we renamed some properties upon the request of domain experts. In particular, we exploit the logical patterns *Tree* and *N-Ary Relation*, the alignment pattern *Class Equivalence*, and the content patterns *Parameter*, *Time Interval*, *Action*, *Classification*.

Integration The integration of both extensions has two objectives: (i) to align them with a foundational ontology, and (ii) to link it with the Linked Open Data (LOD) cloud. The first objective was satisfied by aligning the root concepts of both extensions with ones defined within the DOLCE [14] top-level ontology. While, the second objective was satisfied by aligning our ontology with the UMLS Knowledge Base [9] since it has been included within the LOD cloud recently. This way, it may work as a bridge between the latter and the HeLiS ontology.

Implementation As for the HeLiS ontology, both extensions are represented by using the RDF/XML [10] language in order to provide a formal representation enabling the check of inconsistencies, the visualization of ontology structure, and the ease of maintenance. The editing of the ontology is demanded to the MoKi tool [15], while the exposure of the ontology is granted by the services available from the HeLiS ontology website.

Evaluation To evaluate our ontology we adopted the metrics described in [16–20]: *Accuracy*, *Adaptability*, *Clarity*, *Completeness*, *Consistency/Coherence*, and *Organizational fitness*.

The overall *Accuracy* of the extensions has been judged as good. The knowledge of the domain experts was in-line with the complexity of the use axioms. Indeed, within the HeLiS ontology there are not very complex axioms. Then, by considering the representation of the real world, the evaluators agreed on the correctness of the ontology in describing the domain.

Concerning the *Adaptability* of the ontology, the evaluators focused on the possible extension aspects. They verified that the ontology can be extended and specialized monotonically since once the two extensions are populated, the ontology does not react negatively to these changes because its consistency is preserved.

About the *Clarity* of the ontology, the evaluators agreed with the strategy decided by the modeling team about using concept labels communicating the intended meaning of each concept and the use of definitions and descriptions of the main concepts of the ontology, especially for the root concepts of each branch. Moreover, each definition has been well documented within the ontology in order to make the meaning of each concept understandable by who uses the ontology.

The experts agreed about the *Completeness* of the HeLiS ontology. However, they distinguished among the TBox and the ABox. Indeed, concerning the TBox, the evaluators agreed about the completeness of the ontology and the lexical representations of the concepts. While, regarding the ABox, the evaluators highlighted the necessity of including further nutritional diseases in the future. This observation is definitely pertinent, especially, if we consider the possibility of developing end-users applications.

With the introduction of the two extensions, the HeLiS ontology has been judged, also, *Consistent* and *Coherent*. Consistent because no contradictions were found by the evaluators. Coherent because the evaluators observed little bias between the documentation containing the informal description of the concepts and their formalization.

[9] https://www.nlm.nih.gov/research/umls/
[10] https://www.w3.org/RDF/

Finally, concerning the *Organizational fitness*, the HeLiS ontology has been deployed within the organization as a web service in order to make it easily accessible by the community and potential stakeholders. A focus group has been organized with both the modeling team and the evaluators for discussing about the adopted methodology, that was judged appropriate by considering the necessity of working in-situ all together and of synchronizing the commitments of all the people involved.

Documentation The documentation of the presented extensions of the HeLiS ontology has been done from two perspectives. First, during the whole modeling process, a document has been prepared by the people involved in the construction process. This activity was necessary because the development of the extensions and their sustainability are granted by a public funding program [11]. Thus, all performed steps were documented and archived within the funding dossier. Second, in order to ease the readiness of the ontology for users, we provided a different documentation file generated by using the LODE [12] system and available on the ontology website.

4 Inside The HeLiS Ontology

The full description of HeLiS is provided in [1]. Here, we briefly report which are the main concepts of the ontology and then we focus on the description of the two extensions.

4.1 HeLiS Root Concepts

The ontology contains six root concepts: *Activity*, *Food*, *MonitoringEntity*, *Nutrient*, *TemporalEvent*, and *UserEvent*. Beside these, we also defined the *User* concept that does not play the role of superclass of any concept but that is fundamental for associating specific events with the people did them.

Figure 1 shows a general overview of the ontology with the main concepts.

The *Food* root concept subsumes two macro-groups of entities descending from *BasicFood* and *Recipe* concepts. Instances of the *BasicFood* concept describe foods for which micro-information concerning nutrients (carbohydrates, lipids, proteins, minerals, and vitamins) is available, while instances of the *Recipe* concept describe the composition of complex dishes (such as *Lasagna*) by expressing them as a list of instances of the *RecipeFood* concepts. This concept reifies the relationships between each *Recipe* individual, the list of *BasicFood* it contains and the amount of each *BasicFood* The *Nutrient* concept subsumes 81 different type of nutrients properly categorized. Nutrients are instantiated with through individuals describing a specific amount of a nutrient. Then each *BasicFood* is linked to all the necessary nutrients' individuals through the *hasNutrient* object property.

The *Activity* concepts subsumes 21 subclasses representing likewise physical activity categories and a total of 856 individuals each one referring to a different kind

[11] More details about the sustainability plan of HeLiS are reported in [1]

[12] http://www.essepuntato.it/lode

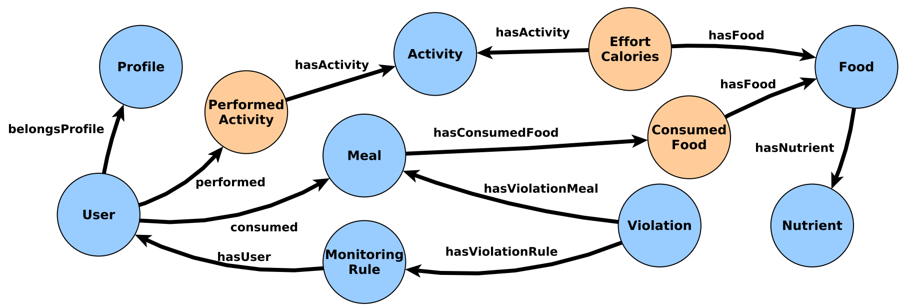

Fig. 1. Overview of the HeLiS ontology.

of activity. For each activity, we defined datatype properties providing the amount of calories consumed in one minute for each kilogram of weight and the MET (Metabolic Equivalent of Task) value expressing the energy cost of the activity.

The *TemporalEvent* concept defines entities used for representing specific moments or delimited timespan which the data to analyze refers to.

The *UserEvent* concept subsumes the conceptualization of information that a user can provide, i.e. food consumption and performed activities, and also links them with the possible violation that can be generated after their analysis. Concerning the representation of users' activities and personalized information, we modeled the *ConsumedFood* and the *PerformedActivity* concepts. Both concepts are used as reification of the fact that a user consumed a specific quantity of a food or performed an activity for a specific amount of time.

Finally, concepts subsumed by the *MonitoringEntity* one are responsible for modeling the knowledge enabling the monitoring of users' behaviors. Here, we can appreciate five concepts exploited at reasoning time for detecting undesired behaviors associated with users: *MonitoringRule*, *Violation*, *Profile*, *Goal*, and *Interval*.

4.2 Food-Disease Extension

The first extension of the HeLiS ontology consisted in adding, to the dietary domain, information concerning the risk level of food categories with respect to specific diseases. Figure 2 shows an excerpt of the HeLiS ontology representing the associations representing the risk level that food categories have with respect to the onset or worsening of some chronic diseases.

We mentioned early that instances of the *BasicFood* concept describe foods for which micro-information of nutrients are available. Moreover, these instances belong also to subclasses of the *BasicFood* concept, such as *Pasta*, *Aged Cheese*, *Eggs*, *Cold Cuts* and *Vegetal Oils*. On the other hand, instances of the *Recipe* concept, describe the composition of complex dishes (such as *Pasta with Carbonara Sauce*) by expressing them as a list of instances of the *RecipeFood* concepts. This concept reifies the relationships between each *Recipe* individual, the list of *BasicFood* it contains and the amount

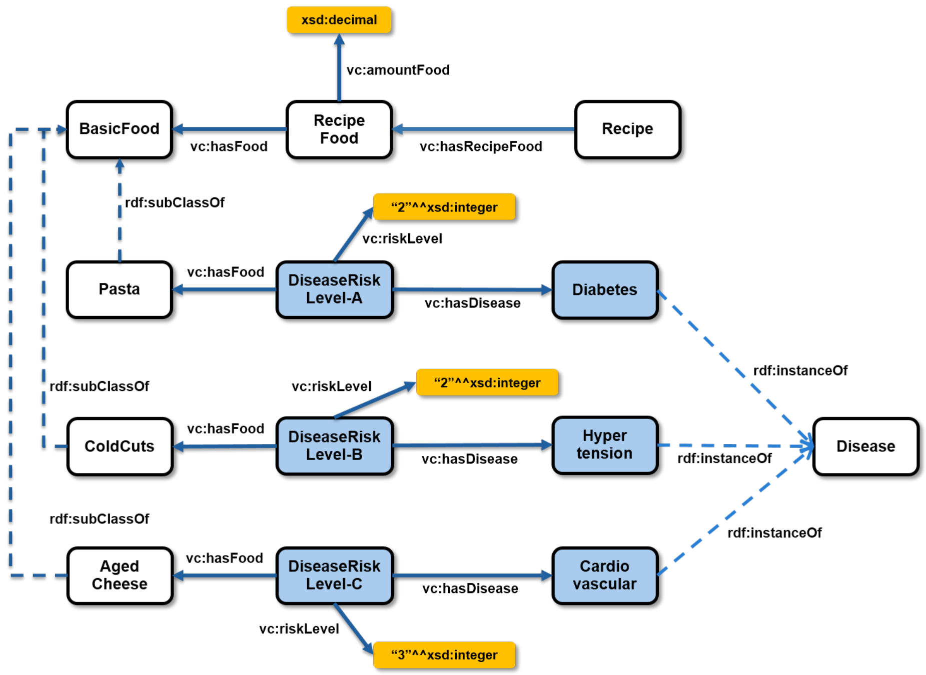

Fig. 2. Excerpt of the HeLiS ontology including the main concepts (white boxes) and instances (blue boxes) exploited by our semantic platform. Solid lines are object properties, dashed lines are RDF core properties

of each *BasicFood*. Besides this dual classification, instances of both *BasicFood* and *Recipe* concepts are categorized under a more fine-grained structure.

The *Disease* concept defines the diseases supported by the system such that information about the risk level relationship with specific *BasicFood* is available. Currently, we instantiate the *Disease* concept for diabetes, kidney diseases, cardiovascular diseases, hypertension and obesity. Diseases are defined as single individuals instead of concepts for avoiding the creation of a new individual for each specific disease for each user. Instances of the *DiseaseRiskLevel* concept reifies the relationships between each *Disease* and *BasicFood* individuals and with the risk level of a *BasicFood* for that *Disease*. The risk level is represented by a value ranging from 0 (no risk) to 3 (high risk). For readability we report in Figure 2 only some instances of the *DiseaseRiskLevel* concept, e.g., *DiseaseRiskLevel-A*, *DiseaseRiskLevel-B*, and *DiseasesRiskLevel-C*.

4.3 Behavior Change Barriers Extension

The second extensions concerns the modeling of the barriers that a user can encounter during the maintenance of healthy habits or the self-management of chronic diseases. This extension is composed by three main parts: (i) the classification of the barriers,

(ii) the representation of the different state of changes, and (iii) a new taxonomy for classifying the list of physical activities defined within HeLiS.

The *Barrier* concept is the root of the barriers classification branch that subsumes six different kind of barriers. The *EnvironmentBarrier* refers to the impossibility of performing an action due to unfavorable climatic conditions, the cost of the equipment need, the lack of safety, etc.. *HealthBarrier* concerns the presence of some disease preventing to complete specific action. This concept enables the possibility of importing external medical knowledge bases (e.g. the UMLS). This way, HeLiS will be connected with medical knowledge that can be exploited at reasoning time. The *PersonalBarrier* concept represents all barriers associated with the real-life situations (e.g. job conditions) that obstruct the performance of specific actions. Then, the *PhysicalBarrier* and *PsycologicalBarrier* concepts are related to hindrances given by physical pains (e.g. knee injury) or emotional status (e.g. fear) that block a person in performing specific actions. Finally, the *SocialBarrier* concept mainly refers to possible lacks of support from people close to patients (parents, friends, etc.).

The second part consists in the abstract representation of the transtheoretical model. Such a model is used in psychology for supporting the behavior change process that a user can perform for changing their lifestyle or habits. Here, we defined the basic concepts which instances can be linked by the *UserStatus* concept already defined in HeLiS and that is used as reification of the status in which a *User* is during a specific *Timespan*. The main concepts we defined are *StateOfChange* that is the root concept of this branch, and then the six phases in which a *User* can be: *PreContemplation*, *Contemplation*, *Preparation*, *Action*, *Maintenance*, and *Termination*.

Finally, this extension provides a new taxonomy of physically activities defined in the core of HeLiS. The taxonomy defined within the core of HeLiS classifies physical activities by type. Differently, this extension provides a classification of physical activities from two different perspectives: the energetic system generally used for performing the action (e.g. aerobic or anaerobic), if the activity required flexibility abilities, and the intensity (or effort) level of each activity. The rationale of this classification is given by the necessity of defining the relationships between barriers and physical activities. For instance, in case a user suffers from a asthma, such a *HealthBarrier* may obstruct the performance of an *OutdoorActivity*.

5 Conclusions And Future Work

In this paper, we presented two extensions of the HeLiS ontology concerning (i) the representation of the risk level of each food categories defined in HeLiS associated with the onset or worsening of a subset of chronic diseases and (ii) the modeling of a knowledge layer representing the barriers that a user may encounter during the self-management of his/her lifestyle or of his/her chronic disease. The knowledge modeled within the HeLiS ontology combines information extracted from unstructured resources with the ones collected from domain experts coming from the medical domain. We described the process we followed to build the ontology and which information we included.

References

1. Dragoni, M., Bailoni, T., Maimone, R., Eccher, C.: Helis: An ontology for supporting healthy lifestyles. In: International Semantic Web Conference (2). Volume 11137 of Lecture Notes in Computer Science., Springer (2018) 53–69
2. Grafenauer, S., Tapsell, L., Beck, E.: Beyond nutrients: classification of foods to identify dietary patterns for weight management. In: 16th International congress of Dietetics
3. Cantais, J., Dominguez, D., Gigante, V., Laera, L., Tamma, V.: An example of food ontology for diabetes control. In: Proceedings of the International Semantic Web Conference 2005 workshop on Ontology Patterns for the Semantic Web. (2005)
4. Snae, C., Bruckner, M.: Foods: A food-oriented ontology-driven system. (2008) 168–176
5. Helmy, T., Al-Nazer, A., Al-Bukhitan, S., Iqbal, A.: Health, food and user's profile ontologies for personalized information retrieval. In Shakshuki, E.M., ed.: Proceedings of the (ANT 2015), London, UK, June 2-5, 2015. 1071–1076
6. Peroni, S., Lodi, G., Asprino, L., Gangemi, A., Presutti, V.: FOOD: food in open data. In Groth, P.T., Simperl, E., Gray, A.J.G., Sabou, M., Krötzsch, M., Lécué, F., Flöck, F., Gil, Y., eds.: The Semantic Web - ISWC 2016 - Kobe, Japan, October 17-21, 2016. Volume 9982 of Lecture Notes in Computer Science. (2016) 168–176
7. Kolchin, M., Zamula, D.: Food product ontology: Initial implementation of a vocabulary for describing food products. In: Proceedings of the 14th Conference of Open Innovations Association FRUCT. FRUCT'14, Helsinki, Finland, Finland, FRUCT Oy (2013) 191–196
8. Çelik Ertuğrul, D.: Foodwiki: A mobile app examines side effects of food additives via semantic web. J. Med. Syst. **40**(2) (February 2016) 1–15
9. Fernández-López, M., Gómez-Pérez, A., Juristo, N.: Methontology: from ontological art towards ontological engineering. In: Proc. Symposium on Ontological Engineering of AAAI. (1997)
10. Pinto, H.S., Staab, S., Tempich, C.: DILIGENT: towards a fine-grained methodology for distributed, loosely-controlled and evolving engineering of ontologies. In de Mántaras, R.L., Saitta, L., eds.: Proceedings of the 16th Eureopean Conference on Artificial Intelligence, ECAI'2004, including Prestigious Applicants of Intelligent Systems, PAIS 2004, Valencia, Spain, August 22-27, 2004, IOS Press (2004) 393–397
11. Suárez-Figueroa, M.C.: NeOn methodology for building ontology networks: specification, scheduling and reuse. PhD thesis, Technical University of Madrid (2012)
12. Michie, S., West, R., Campbell, R., Brown, J., Gainforth, H.: ABC of Behaviour Change Theories. (06 2014)
13. Hitzler, P., Gangemi, A., Janowicz, K., Krisnadhi, A., Presutti, V., eds.: Ontology Engineering with Ontology Design Patterns - Foundations and Applications. Volume 25 of Studies on the Semantic Web. IOS Press (2016)
14. Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.: Sweetening ontologies with dolce. In: Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web, Springer-Verlag (2002) 166–181
15. Dragoni, M., Bosca, A., Casu, M., Rexha, A.: Modeling, managing, exposing, and linking ontologies with a wiki-based tool. In: LREC. (2014) 1668–1675
16. Obrst, L., Ceusters, W., Mani, I., Ray, S., Smith, B. In: The Evaluation of Ontologies. Springer US, Boston, MA (2007) 139–158
17. Gangemi, A., Catenacci, C., Ciaramita, M., Lehmann, J.: Modelling ontology evaluation and validation. In Sure, Y., Domingue, J., eds.: The Semantic Web: Research and Applications, 3rd European Semantic Web Conference, ESWC 2006, Budva, Montenegro, June 11-14, 2006, Proceedings. Volume 4011 of Lecture Notes in Computer Science., Springer (2006) 140–154

18. Gómez-Pérez, A.: Ontology evaluation. In Staab, S., Studer, R., eds.: Handbook on Ontologies. International Handbooks on Information Systems. Springer (2004) 251–274
19. Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing? Int. J. Hum.-Comput. Stud. **43**(5-6) (1995) 907–928
20. Grüninger, M., Fox, M.: Methodology for the Design and Evaluation of Ontologies. In: IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing, April 13, 1995. (1995)