A two-staged approach to developing and evaluating an ontology for delivering personalised education to diabetic patients

Susan Quinn, Raymond Bond and Chris Nugent

Computer Science Research Institute, School of Computing & Maths, University of Ulster, Newtownabbey, County Antrim, UK

Corresponding author: Quinn-S47@email.ulster.ac.uk

Abstract

Ontologies are often used in biomedical and health domains to provide a concise and consistent means of attributing meaning to medical terminology. Whilst they are novices in terms of ontology engineering, the evaluation of an ontology by domain specialists provides an opportunity to enhance its objectivity, accuracy and coverage of the domain itself. This paper provides an evaluation of the viability of using ontology engineering novices to evaluate and enrich an ontology that can be used for personalised diabetic patient education. We describe a methodology for engaging healthcare and information technology specialists with a range of ontology engineering tasks. We used 87.8% of the data collected to validate the accuracy of our ontological model. The contributions also enabled a 16% increase in the class size and an 18% increase in object properties. Furthermore, we propose that ontology engineering novices can make valuable contributions to ontology development. Application specific evaluation of the ontology using a semantic-web based architecture is also discussed.

Keywords: Ontology, personalisation, diabetes

1. Introduction

Decreasing levels of physical activity and an escalation in unhealthy diets have contributed to an increased international prevalence of diabetes [1]. It is estimated that 382 million people are currently diagnosed as diabetic whilst a further 175 million remain undiagnosed [1]. Diabetes is a chronic condition in which the body may not produce sufficient amounts of the hormone insulin, or may not be able to use the produced insulin effectively [1]. The medical complications of the condition can include cardiovascular disease, kidney disease and neuropathy, moreover diabetes can have a psychological impact on an individual resulting in increased stress or depression [2,3]. Complications such as these may reduce an individual's quality of life, and may result in disability or death [1,4]. Managing diabetes can be a complex and demanding responsibility for the diabetic patient [5] and may involve, for example, daily monitoring routines, managing medication, and lifestyle changes such as healthy nutrition and increased exercise [2]. One fundamental approach to assisting patients with the management of their condition is through the provision of high quality education. There is an increasing emphasis that education should enhance a patient's knowledge and skills regarding the management of their diabetes, and empower them to take an active role in their treatment [4,6,7]. Structured educational programs have been successful in assisting patients with self-monitoring, increasing knowledge and improving their health outcomes [1,7,8]. It has also been suggested that diabetic education will be most effective if it is individualised to the particular medical, educational, social and cultural background of each individual [9,10].

In a previous study we presented an ontology based architecture to provide web based personalised education to patients that have been diagnosed with diabetes or obesity [11]. We proposed that the educational content could be tailored to focus on the particular health objectives and personal characteristics of each patient, and could be transcribed at a readability level suitable for the patient's educational needs. We suggested that the personalised education produced may be more comprehensible, usable and engaging for the patient when compared to a generic approach such as standardised leaflets. A central component of the architecture comprised a Web Ontology Language (OWL) [12] ontology which represented the domain knowledge necessary for the production of the personalised education. The ontology included a user model that captured information about the patient, and a model of the health conditions diabetes and obesity including symptoms, treatments and complications. In order to construct the ontology, we undertook a review of medical literature that focused on diabetes and obesity. However, in order to ensure that the developed ontology was an accurate and concise model of the domain we determined to conduct a two-staged approach to the evaluation of the ontology. In the first stage we engaged healthcare and information technology specialists to complete ontology engineering tasks to define vocabulary and organisational structures for the ontology. These were used to evaluate the accuracy of our ontology. During the second stage of evaluation the refined ontology was included in the architecture in order to assess whether it could be used in the production of personalised patient education. In this paper we describe the development, implementation and results of both phases of ontology evaluation. The remainder of the paper is structured as follows. Section 2 reviews literature related to the evaluation of health domain ontologies. Section 3 describes the ontology based architecture for personalised patient education. Section 4 describes the design and development of the ontology engineering tasks, and section 5 discusses how the knowledge contributions were used. Section 6 discusses the application-based evaluation of the ontology. Section 7 provides a discussion of the findings and section 8 describes the conclusions drawn from this work.

2. Related work

2.1 Ontological models of diabetes

An ontology has been defined as “a shared understanding of some domain of interest” [13]. It reflects a particular world view of a domain by representing domain knowledge as an arrangement of concepts, definitions and relationships, and defining shared agreement on the meaning of these components [13]. An ontology formalises the meaning of knowledge within a domain in a format that is comprehensible for humans and machines [14,15,16]. Ontologies are being increasingly used within the biomedical and health domains to standardised medical vocabularies and provide validated semantics for medical data [16]. Within the domain of diabetes, ontological models have been developed to express a range of clinical knowledge related to aspects of diabetes management such as medication, lifestyle changes, and medical checks. Arwan et al. [17] proposed the use of ontologies to facilitate the production of food recommendations for diabetic patients. The proposed architecture incorporated a Calorie Foods ontology which is used to represent various characteristics of diabetic patients including calorie needs, and a Foods ontology to represent food attributes. SPARQL Protocol and RDF Query Language (SPARQL) queries are used to derive semantic matches between the ontologies and produce the nutritional recommendations. An ontology based recommendation system that supports physicians in prescribing diabetic medication is presented by Chen et al. [18]. An ontology is used to model medication related knowledge such as name, side effects and contraindications, and medical tests such as HbA1c (glycated haemoglobin). Semantic Web Rule Language (SWRL) rules are incorporated to determine suitable medication recommendations for a particular patient including dosage and associated monitoring. Chammas et al. [19] proposed a tool for diabetic patients that provides graduated levels of advice for avoiding diabetic podiatry related complications. Central to this is a computational model that comprises an ontology which captures patient information such as podiatry observations, symptoms, lifestyle factors and medical test results. SWRL rules then determine the category of guidance and advice provided to a patient.

Ontologies have also been developed that capture the knowledge necessary to assist with the identification of diabetes. Rahimi et al. [20] developed the Diabetes Mellitus Ontology (DMO) to assist with the diagnosis and management of individuals with diabetes. An algorithm was developed, based on the DMO, which utilised data derived from Electronic Health Records to identify patients with type 2 diabetes mellitus. Alharbi et al. [21] developed an ontology based clinical decision support system to diagnose diabetes and provide an appropriate treatment plan. Clinical Practice Guidelines (CPGs) were utilised to develop a diabetes domain ontology, a patient ontology and SWRL diagnostic rules. The SWRL rules utilise data captured in the ontology, including lab tests and symptoms, to identify cases of diabetes and pre-diabetes. The reviewed literature highlights how ontologies can be used to provide a concise model of the medical, contextual and patient information associated with diabetes diagnosis and treatments. Semantic reasoning is often used to facilitate decision making capabilities. Thereby ontology based applications can facilitate the provision of individualised healthcare and treatment for diabetic patients, which may prove useful for both physicians and patients.

2.2 Ontology evaluation

Biomedical and healthcare ontologies may be developed through the re-use of established vocabularies or through the construction of an original ontological model. This paper describes the evaluation of a novel ontology that was developed to facilitate the provision of personalised patient education. Evaluation is an important aspect of ontology engineering which focuses on appraising an ontology, using objective criteria, in order to determine whether it reaches a quality standard [15,22,23]. Most methods of evaluation fall into one of the four classifications suggested by Brank et al. [24]; comparing the ontology to a gold standard, application-based evaluation, data-driven evaluation and appraisal by humans. Many ontology engineering projects will use multiple evaluation methods however one of the most common approaches involves appraisal of the ontology using a set of established measures. Delir Haghighi et al. [25] used *criteria-based evaluation* to assess whether an ontology for medical emergency management in mass gatherings achieved its intended objectives. Domain experts validated the concepts, hierarchies and relationships of the ontology which enabled the identification of concepts that were ambiguous, inconsistent, contradictory or superfluous. In the next stage they used an automated approach to evaluating the ontology coverage. Leximancer was used to extract concepts and terms from domain documents and these were compared with the ontology. Bright et al. [26] implemented a two staged approach to *intrinsic evaluation* of an ontological model of the domain of antimicrobial prescribing. Firstly, they appraised the ontology classes for adherence to the design principles of Cimino's Desiderata. In the next phase of evaluation, domain experts used the laddering technique to evaluate ontology correctness, producing hierarchies that were compared to the ontology taxonomies. Compliance with Cimino's Desiderata and the OBO Foundry's ontology design principles were also used to appraise adherence to standard ontology engineering practices for the bacterial clinical infectious diseases ontology (BCIDO) [27]. Moreover, the authors also addressed the issues associated with gold standard evaluation by devising a semi-automated approach that used clinical practice guidelines, electronic health records and expert case studies to create a collection of domain knowledge with which the ontology was manually compared. The authors proposed that this approach provides an efficient means for updating a gold standard compendium and reduces the reliance on domain experts.

Application-based evaluation focuses on appraising the usefulness of the ontology as a component of a medical application [22,23,28]. Attributes of the application such as performance measurements or expected outputs are used to gauge whether the ontology achieves its intended objectives. Valls et al. [29] gathered feedback from the users of an ontology based system and used this to appraise the feasibility of the underlying ontology. They developed the Actor Profile Ontology (APO) to capture knowledge related to home care assistance. This was incorporated into the K4Care Platform which was then tested by medical professionals in a real life scenario. They developed two questionnaires which were based on the Technology Acceptance Model, and used the generated scores to assess whether the ontology adequately supported certain functionalities and characteristics of the system. For example, the feedback regarding the perceived ease of use of the system was linked with system flexibility, and adaptation and personalisation capabilities, characteristics that are directly supported by the ontology. Thereby user feedback was useful in evaluating whether the ontology successfully supported particular functionalities within the system.

Ontologies are widely used within the biomedical and health domains to underpin clinical decision support systems, data exchange and knowledge management [26,27,29]. Therefore, it is essential that the ontological model is an accurate, unambiguous and consistent reflection of domain knowledge. Evaluation methods provide an opportunity to identify erroneous, inconsistent and redundant data within the ontology. Moreover, evaluation can also confirm that an ontology can achieve intended objectives within a medical system. The studies described have utilised a range of automated, semi-automated and manual evaluation procedures. In many cases human expertise was critical to confirming the validity of the ontology. In common with these approaches we also engaged with domain specialists to confirm the accuracy of our ontological model. However, the novelty in our approach is derived from engaging ontology engineering novices from healthcare and information technology backgrounds to provide ontology modelling decisions. The results were collated to develop a knowledge base which was then used to validate and augment our ontology. Furthermore, application-specific evaluation of the ontology confirmed that the refined ontology achieved its functional objectives.

3. An ontology based architecture for personalised patient education

This paper describes the methods used to evaluate an ontology that is a component of an architecture that provides personalised patient education. The architectural model is illustrated in Figure 1. The ontology represents information related to the four main domain entities and includes a patient model, a medical conditions model, an activity model and an educational content model. The architecture incorporates a Pellet reasoning engine that utilises the ontology and a set of SWRL personalisation rules to determine the composition and style of the education. Java and Spring MVC are used to communicate with the educational repository and create the educational content, which is delivered as a JavaServer Page (JSP). The education will include text information and images.

Figure 2 indicates the steps involved in the development and evaluation of the ontology. The primary design stages included a review of the current approaches to diabetic patient education including pamphlets, booklets and educational websites. Some of the information sources used included Diabetes UK [2], the American Diabetes Association [3] and the UK National Health Service [30]. We used information from these sources to develop a knowledge base of the clinical features related to the symptoms, complications and treatments of diabetes and obesity. In order to design a domain model of diabetes and obesity we extracted concepts from the knowledge base and determined the relationships between these concepts. We used a top-down design approach whereby the classes in the top layer of the ontology define models of the four main domain entities and the information captured for each model becomes increasingly specialised with each subsequent layer.

We also conducted a literature review which focused on generic patient education. This enabled us to identify a number of factors that could limit the effectiveness of generic educational approaches. From these reviews we determined that the personalisation mechanisms would focus on tailoring the content and presentation of the education to the health status and personal characteristics of the patient, and designed a user model accordingly.

In the ontology the user model is represented as the Patient Model. The data captured for each patient includes personal characteristics and preferences, and information regarding the patient's health status. The Patient Model is categorised into seven main profile classes which are detailed in Table 1. The Medical Conditions model represents aspects of diabetes and obesity and is categorised into three strands: symptoms, treatments and complications. Each strand comprises subsumption relationships that provide increasingly specific representations of clinical data. For example, the data captured in the treatments model is categorised as pharmacological and non-pharmacological treatments, and in the subsequent layer pharmacological treatments is further classified as diabetic medications and insulin types. Representing the knowledge with increasing granularity enables an extensive range of diabetes related clinical data to be captured. The ontology also includes an activity model which captures information related to physical activity. The ontology was developed using Protégé, an ontology editor, and has been represented using the Web Ontology Language (OWL).

As illustrated in Figure 2, the evaluation of the ontology was conducted in two stages. The first stage focused on constructing and using a collection of domain expertise to validate and enrich the ontology. The second stage evaluated whether the ontology could be utilised in the production of personalised patient education. If any refinements were made to the ontology throughout these evaluations a Pellet reasoner was used to check the correctness and consistency of the enhanced ontological model.

4. Methodology for domain expert evaluation

In the first stage of evaluation we wished to include expert evaluation of the ontology, however we were aware that not all domain specialists would have experience of using an ontology. Therefore, we aimed to design a methodology that would enable ontology engineering novices to contribute their domain knowledge. We developed a set of ontology engineering tasks which were presented to participants through an online survey. The tasks collected knowledge related to ontology concepts, properties and restrictions. This approach to ontology evaluation had three main objectives, (1) to validate our initial ontology design, (2) to enrich the ontology with new knowledge, and (3) to determine the feasibility and efficacy of using ontology engineering novices to evaluate a health ontology.

4.1 Design of the ontology engineering tasks

It was expected that there would be variation in the technical modelling expertise amongst the participants therefore a challenge lay in developing tasks that would render ontological engineering accessible to non-experts. Consequently, we developed an online multimedia presentation to provide participants with the necessary information. The multimedia presentation used illustrations and analogies to cover technical topics such as organising domain information as classes, using subsumption hierarchies to encapsulate specialisation, and using relationships to model associations between unrelated classes. The presentation also introduced the diagrams and terminology that would be used in the survey. A multimedia presentation was chosen as the medium for conveying this information as it enabled a visual presentation and voiceover that would be more engaging in comparison with textual information. The screencast time was limited to three minutes so that a participant would not be overloaded with unnecessary information.

We wished to collect contributions related to three of the domain entities which included the patient, the health conditions diabetes and obesity, and physical activities. In order to reduce the complexity of the survey it was organised into sections that focused on the following subjects; patient information, physical activities, symptoms, treatments and complications. The tasks were designed to reflect five types of ontology modelling practices.

(1) Suggest new data to be captured in the ontology and assign a cardinality.

(2) Enter terms and organise these in a subsumption hierarchy.

(3) Enter terms and classify these within an ontological model.

(4) Analyse relationships in the current ontology and suggest alternative representation.

(5) Enter new properties for a class.

The layout of the survey was tailored to assist with the modelling tasks. Figure 3 illustrates how layout and terminology was used to assist a participant with technical modelling decisions, for example (a) enabled a participant to enter a subclass name and a superclass name to arrange in a subsumption relationship. The survey also contained descriptions, examples and ontological models to assist with modelling decisions. Health information was also included to contextualise the information being requested. For example, definitions of physical activity intensities were provided to help the participant categorise a list of sporting activities. This information was gathered from various sources including the UK National Health Service [30] and Diabetes UK [2]. The survey and multimedia presentation are available at [31].

4.2 Profile of contributors

The online tasks required knowledge of diabetes and obesity, and the ability to identify associations between the entities in these domains. To this end it was decided to invite participants from two main disciplines, healthcare and information technology. Participants from these backgrounds would provide a diversity in the necessary expertise and skills required for the study. Participants were recruited from professional networks via email.

4.3 Ethical considerations

The study had been approved by a University of Ulster Research Ethics Filter Committee. The first section of the online survey provided the participants with information regarding the purpose, procedure and length of the survey, and indicated that the collected data would be anonymised and stored securely. It was also highlighted that participation was voluntary and that participants could withdraw from the study at any stage. Informed consent was obtained from each participant by completion of an online consent form.

5. Results of domain expert evaluation

The survey comprised 15 questions that collected demographic information, 18 ontology engineering tasks and a comments section. Demographic information was completed by 26 participants but only 21 undertook the ontology engineering tasks. The criteria for inclusion in the study required that a participant completed at least one of the ontology based tasks, therefore five participants were excluded from the study. The remaining participants comprised 15 male and six female participants with an age range of 24 to 70. The disciplines of the participants included information technology (16), mathematics and engineering (3) and healthcare and life sciences (2). Only seven participants had previously used or developed an ontology. The main ontology languages used by these participants included OWL and Resource Description Framework (RDF) while Protégé was the only ontology editor used.

In this study we attempted to appraise the feasibility of including ontology engineering novices in ontology evaluation. In order to appraise the viability of our approach we analysed the vocabulary provided by the participants and the completion rate of the technical modelling tasks.

5.1 Informational contributions

Table 2 illustrates the accumulated contributions for each of the survey sections and how these were used to validate and augment the ontology.

5.1.1 Symptoms, treatments and complications

In total 14 of the tasks collected information related to the symptoms, treatments and complications of diabetes and obesity. The investigators aggregated the contributions and similar suggestions were identified and clustered. Following this, the ontology was searched for a corresponding class or property. If a match was found, then this was interpreted as validation of the existing ontology. If a match was not found a literature search was conducted by the investigators to determine if the suggestion could be verified as a feature of diabetes or obesity. The sources reviewed included Diabetes UK [2], the UK National Health Service [30], MedlinePlus [32] and websites that had been accredited with a Health on the Net certification [33]. If the contribution was verified as an aspect of diabetes or obesity it was added to the ontology, otherwise it was categorised as invalid.

The contributions that related to the symptoms of diabetes and obesity included generic conditions and specific indications of diabetic complications. The suggestions included extreme fatigue, high cholesterol, frequent urination, excessive thirst, unexplained weight loss, dry mouth, abdominal pain, disorientation and fainting. All these contributions were matched with symptoms represented in the ontology. A number of new suggestions were also added as classes in the ontology. This included heart racing as a symptom specific to hypoglycemia and high cholesterol as a symptom of obesity. Most significantly there was a high level of overlapping between the symptoms and complications sections as approximately 15% of the suggested symptoms were verified in the literature as diabetes complications. This would suggest that some participants were unclear of the distinction between diabetes symptoms and complications.

The majority of the contributions for diabetes and obesity treatments were focused on non-medication treatments. The most frequently suggested generic treatments included healthy eating and increased exercise. Diabetic health checks and monitoring routines were also recommended including daily blood glucose measurement, HbA1c blood test, blood pressure measurement, eye screening, foot checks and weight checks. All these suggestions were mapped to classes in the current ontology. A number of suggestions related to mental health assistance. A literature search confirmed that this should be added to the ontology as a treatment. Approximately 4% of the contributions related to medication treatment. Insulin was the most common suggestion and it was also proposed that information relating to insulin delivery methods and regimes should be represented in the ontology. These were added as properties of the Insulin classes. Metformin was the only non-insulin medication mentioned which suggests that there could be a lack of knowledge related to non-insulin medications amongst the participants. In summary 5.4% of the contributions were used to validate the current ontology and 0.7% were used to add new knowledge.

The contributions related to diabetes and obesity complications included both physical and psychological problems. The suggestions included breathlessness, depression, blurred vision, foot ulcers, cardiovascular disease, joint and back pain, poor circulation and amputation. Moreover, while some of the contributions used general descriptions of problems a number of complications were expressed using precise medical terms including neuropathy, retinopathy and sleep apnoea. All of these contributions were mapped to classes in the current ontology. Blurred vision had been represented in the ontology as a symptom of diabetes however a literature search verified that it should also be classified as a complication. A number of new complications were proposed such as loss of mobility, foot pain and heart attack. These were added as classes in the ontology. There were also suggestions that recording contextual information or behavior that increased the risk of a complication should be recorded. This was represented in the ontology as properties of patient behavior. Overall 3% of the contributions were used to validate the current ontology and 1.2% were utilised to add new classes and properties.

5.1.2 Physical activities

One of the main approaches to managing diabetes and obesity is through increased physical activity, however it is important that any sporting information provided to a patient is appropriate to their physical activity level. The first task required the participants to assign physical activity intensities to a list of 28 sports. Majority voting was used to determine the most appropriate intensity from the contributions. In most cases there was agreement amongst the participants however there were conflicting assignations for three of the sports. In these cases a reference sporting compendium [34] was utilised to determine the most appropriate intensity. The participants also contributed 10 new sporting activities, and eight activities of daily living (ADLs) which included various gardening and housework activities. All these suggestions were added as classes in the ontology.

5.1.3 Patient health characteristics

In order to gather information that could be used to enhance the patient profile the participants were asked to suggest patient characteristics that could be relevant to personalised education. Moreover, the participants also had to indicate the cardinality of this characteristic. The most common characteristics included age, weight, height, BMI category, ethnicity, gender, preferred language and literacy level. These suggestions were mapped to classes and properties in the ontology. Novel suggestions included information related to the patient's medical team. This was added to the ontology as it could be useful for inclusion in personalised patient education.

5.2 Technical tasks

The success of our novel approach to ontology evaluation was hugely reliant on the participants' motivation to engage with the ontology engineering tasks. The survey comprised five types of tasks that focused on general ontology modelling activities. Therefore, in order to appraise the participants' willingness to undertake technical tasks we considered the completion rate for each task type. Table 3 indicates the number of participants from each discipline that attempted each task type.

The first task type required a participant to assign a cardinality (exactly 1, at least 1) to a data property. This question was completed by 90.5% of the participants including both participants from a health related discipline. This would suggest that participants with a range of technical expertise would be confident with this data modelling practice. The second task type used subsumption relationships to organise data. Approximately 42.8% of the participants did not attempt any of these questions, including participants from health, information technology, and engineering and mathematics. This would suggest that modelling information in subsumption relationships may prove a complex undertaking for ontology engineering novices, even for those with a technical background. The third task type focused on providing and classifying information. The participant was shown a hierarchical model from the ontology and asked to classify information in different classes. Only 1 participant (from a healthcare background) did not attempt any questions of this type. This would suggest that participants from a technical and non-technical background may be able to comprehend how to classify information within ontology classes. The fourth task type required the participant to analyse an ontological modelling decision. Approximately 61.9% of the participants attempted these tasks. In parallel with the second question type (subsumption relationships) this would suggest that participants may find it difficult to assess the suitability of hierarchical structures for representing information in ontologies. The fifth task type focused on adding new properties for a class. This had the lowest completion rate with only 42.9% of participants attempting these tasks. This was surprising since the first task type, which also focused on adding new data properties, had a high completion rate.

Overall there was a disparity amongst the participants in willingness to attempt the different task types. In total eight participants (38.1%) attempted all five task types. This included one participant with a healthcare background, and one from an engineering background, neither of whom had used an ontology before. The remaining six participants had an information technology background, two of these had previous experience of using an ontology and all had expertise in using databases and scripting languages. It is notable that a participant from a healthcare discipline attempted all of the engineering task types. This would suggest that ontology evaluation need not be limited only to domain experts with technical expertise but that valuable technical contributions can be collected from non-technical personnel.

5.3 Motivation and engagement

It can be challenging to determine a participant's motivation for completing online tasks. Typically a participant will have an intrinsic incentive such as a financial reward or an extrinsic motivation such as charity, learning or enjoyment [35]. No compensation was offered for completing the survey therefore the motivations of the participants did not appear to be intrinsic. The average survey completion time was 31 minutes which would suggest that the level of engagement amongst the participants was high. Moreover, one participant commented that lack of time prevented them from providing more contributions, whilst four others indication that a lack of knowledge of diabetes and obesity had the same effect.

Figure 4 indicates the average number of contributions collected from each expertise category. Surprisingly some of the lowest averaged contributions were from participants that had rated their diabetes expertise in the upper quadrant. However, this may be associated with a reluctance from some participants to engage with the technical modelling tasks as three of the four participants with the highest rated diabetes expertise (7-9) did not attempt any of the type 2 or type 4 ontology engineering tasks.

When analysing the motivations of a crowd of participants it is also significant to determine the profile of members that did not engage with the engineering tasks. This study focused on technical modelling tasks therefore it was interesting to determine the technical experience and health knowledge of participants that would not engage with these types of tasks. In total five participants started the survey but did not complete any of the engineering tasks. Within this group three of the participants had a health background and no previous experience of using databases, programming or scripting languages, while two participants had an information technology background and expertise in all three technologies. All five had never used an ontology before. With regard to diabetes and obesity expertise, two of the participants rated their knowledge as high while the other three rated their knowledge as low. However, with such a small group it is unfeasible to deduce any correlations between any of the mentioned characteristics and non-participation. Only one of the group mentioned a reason for non-participation and indicated that the survey was too long.

6. Application-based evaluation

The second stage of evaluation aimed to assess whether the enhanced ontology could be used in the architecture to facilitate the production of web based personalised patient education. A test suite of Patient Model instances were designed to model patients that had been diagnosed with type 2 diabetes. The test profiles included common symptoms, treatments and complications of type 2 diabetes, and were designed to reflect a diversity of gender, age and educational abilities amongst the patients. Section 6.1 presents a vignette related to one of these test profiles. The following criteria were used to evaluate whether the education produced corresponded with a test profile.

- The education includes the patient's name.
- The textual components of the education describe only the particular symptoms, treatments and complications as asserted in the profile or inferred by the personalisation rules.
- The text is at an appropriate readability level as specified by the patient's educational profile.
- The images chosen are matched to the patient's gender and age.
- The education layout is as specified in the patient's preferences profile.

6.1 Test profile

Jane is a 65-year-old woman who has just been diagnosed with type 2 diabetes. She is worried about this diagnosis, however, her physician explains that her condition can be successfully managed through lifestyle changes and regular health checks. The surgery has a facility to provide personalised electronic education. Jane's personal, educational and health characteristics are added to her patient profile and the personalised educational material is generated. As Jane presented with blurred eyesight, the education highlights the importance of regular eye screening. Moreover, the personalisation rules infer that the text information should be presented in an increased font size and with increased line spacing. Her educational profile is used to determine her level of health literacy and to select the most suitable text components. She finds that the information has a clear flow and uses terminology that she can understand. Age and gender appropriate images are also included, which helps Jane to identify and engage with the information. Figure 5 presents an illustration of the personalised education produced for this test profile.

6.2 Results of application-based evaluation

In all of the test cases the education matched the characteristics as specified in the profile, and all the evaluation criteria were fulfilled. Each patient had been diagnosed with type 2 diabetes therefore the personalisation rules inferred that information related to blood glucose measurement should be added. In each case the education focused only on each patient's particular experience of symptoms, treatments and complications. The textual information was provided at a suitable readability level and the images corresponded with the patient's gender and age. The results of the application-based evaluation indicated that the ontology provided a comprehensive knowledge base to assist with the production of personalised patient education.

7. Discussion

Managing diabetes can require a patient to undertake complex daily monitoring routines, lifestyle changes and longer term health checks. Education has been recognised as an essential aspect of care for diabetes which can enhance the self-management skills of patients, and improve health outcomes [1,6,36]. This paper described the development and evaluation of an ontology that is used in the production of personalised education for diabetic patients. Ontologies can provide the expressiveness to capture the semantics of a diverse range of domain data, and thereby ensure that the data is suitable for processing by reasoning technologies [14]. Ontologies have been successfully utilised to represent a variety of clinical information related to diabetes healthcare and treatments [17-19]. Moreover, semantic reasoning is often used to produce individualised treatment recommendations to assist both physicians, and patients. In common with these applications we needed to develop a model of clinical data related to diabetes and obesity. However, we also needed to capture patient characteristics, and information about the educational components that we would use to personalise the diabetic education. During our review of ontology-based medical applications we did not identify an ontology that captured all the entities in the patient education domain and determined to developed an original ontology for our architecture.

Evaluation is an essential phase of ontology engineering through which the accuracy and correctness of an ontological model can be enhanced [14,22,24]. There are many methodologies available for ontology evaluation [22,37] and most projects will use more than one evaluation method to ensure the validity of the ontology. Scientific and medical literature, and clinical documents such as CPGs are often scrutinised in order to derive evidence based domain knowledge from which an ontological model of a domain can be constructed or evaluated [18,21,27,38]. During the development of our ontology we employed a similar methodology, and undertook a review of medical literature and the websites of medical organisations in order to develop a knowledge base related to diabetes and obesity. Domain experts are often employed during ontology evaluation to appraise the correctness and conciseness of an ontology, or to develop a reference gold-standard to be used for evaluation. As we were unable to engage with domain experts during the initial stages of our ontology design we aimed to use domain expertise to appraise the accuracy and correctness of our ontology. We engaged healthcare and information technology specialists to complete ontology engineering tasks, from which we gathered a collection of domain knowledge with which we could validate our ontology. Approaches to ontology evaluation by domain experts have included the development of relationship models of an ontology, for example by use of the laddering technique [26,27]. In a similar manner we utilised some of the contributions to define subsumption relationships, which enabled us to evaluate the accuracy of some of the ontological relationships within our ontology. The majority of contributions matched the constructs within the ontology, and were used to verify the accuracy of the ontology. This outcome is comparable to domain expert evaluations for other medical ontologies in which the experts were largely in agreement with the ontological models developed for the application [26,27]. Moreover, the new knowledge collected from the participants was added to the ontology, thereby expanding the coverage of the domain. The main novelty in our approach is derived from engaging domain experts that had little or no comprehension of ontologies. We provided artifacts that described various attributes of ontologies, and intuitive interfaces that assisted with engineering tasks. The majority of participants attempted a range of modelling activities. This included participants who had no previous knowledge regarding ontologies, and in particular one participant from a health background who had never used a database or programming language. This would suggest that domain specialists with limited modelling experience can successfully complete technical modelling tasks, and can make valuable contributions to ontology evaluation. There are a number of benefits that can be derived from this approach to evaluation. Firstly, the scope for engaging domain expertise is expanded, as involvement in evaluation activities is not limited to those with ontological modelling experience. Furthermore, as the contributions were collected through an online survey this was a time-efficient approach to collecting domain expertise. However, the collation and validation of the contributions was manually conducted by the investigators. Future work could focus on automation of these processes, thereby increasing the efficiency of this evaluation methodology. The second phase of evaluation focused on the usefulness of the enhanced ontology within the educational architecture. The results of this evaluation indicated that the ontology provides a comprehensive knowledge base for the production of personalised patient education for diabetic patients.

There were a number of limitations with this study. The number of participants that engaged with the online survey was relatively small, and most notably participation from the healthcare domains was very limited. If we had been able to engage with a greater number of participants from a healthcare background the range and specificity of the diabetes clinical information collected may have been enhanced. Another limitation of the study was that we were unable to engage with diabetic patients as study participants. This was an unfortunate omission as the contributions of diabetic patients could have provided valuable feedback on the usability of personalised education. Future development of the methodology could focus on adaptation of the online survey to provide more assistance with the modelling tasks. This could include extending the screencast to illustrate how to complete the different tasks within the survey interface. Another future development relates to the re-usability of the ontology. The ontology was developed as an original ontological model for this application, however in order to enhance the potential for re-use in a similar patient education application, a future development could map the ontology classes to a more widely referenced terminology such as SNOMED CT.

8. Conclusion

This paper described a two-staged approach to ontology evaluation. In the first phase we aimed to evaluate the viability of using ontology engineering novices to enhance an ontology. We developed tasks that collected vocabulary related to diabetes and obesity and organised this in ontological structures. We then used these contributions to validate the accuracy of our ontology and to expand the coverage of the domain. The survey collected a total of 936 contributions. Approximately 87.8% of the contributions were used to validate the ontology and 4.5% were used to add new knowledge to the ontology. There was a 16% increase in the number of classes in the ontology and an 18% increase in the number of object properties. An analysis of the contributions also indicated that we were successful in getting participants from technical and non-technical backgrounds to engage with a range of ontology engineering tasks. Furthermore, application-specific evaluation indicated that the ontology provided a useful knowledge base for the creation of personalised patient education. There are no defined benchmarks by which to determine the quality of an ontology therefore it is unfeasible to determine if the contributions collected have improved the quality of the ontology. Nevertheless, the participants did suggest novel approaches to personalising the education that can be incorporated into our architecture. We would conclude that using the expertise of ontology engineering novices is a viable method of enhancing a healthcare related ontology. We used the knowledge collected to verify the objectivity of our ontology and to develop a more detailed model of the diabetes and obesity domains. We would suggest that ontology engineering can be accessible for novices if sufficient information is provided and if the participant is motivated to engage with modelling tasks.

Declaration of Interest

The authors report no conflicts of interest.

References

[1] International Diabetes Federation Diabetes Atlas, 6th ed. [Internet]. International Diabetes Federation; 2013 - [cited 2016 Dec 2]; Available from: http://www.idf.org/diabetesatlas/download-book

[2] Diabetes UK – Care. Connect. Campaign. [Internet]. Diabetes UK; 2016 - [cited 2016 Dec 2]; Available from: http://www.diabetes.org.uk/

[3] American Diabetes Association. [Internet]. American Diabetes Association; 2016 - [cited 2016 Dec 2]; Available from: http://www.diabetes.org

[4] McDowell J, Brown F, Matthews D, editors. Diabetes A handbook for the primary healthcare team. 2nd ed. Edinburgh: Churchill Livingstone; 2007.

[5] Al Sayah F, Williams B. An integrated model of health literacy using diabetes as an exemplar. Canadian Journal of Diabetes 2012;36:27-31.

[6] Clark M. Diabetes self-management education: A review of published studies. Primary Care Diabetes 2008;2:113-120.

[7] Coppola A, Sasso L, Bagnasco A, Giustina A, Gazzaruso C. The role of patient education in the prevention and management of type 2 diabetes: An overview. Endocrine 2016;53:18-27.

[8] Lawton J, Rankin D. How do structured education programmes work? an ethnographic investigation of the dose adjustment for normal eating (DAFNE) programme for type 1 diabetes patients in the UK. Soc Sci Med 2010;71:486-493.

[9] Managing Older People With Type 2 Diabetes [Internet]. International Diabetes Federation; 2013 - [cited 2016 Dec 2]; Available from: http://www.idf.org/guidelines/managing-older-people-type-2-diabetes

[10] Catherine HY, Parsons JA, Hall S, Newton D, Jovicic A, Lottridge D, Shah BR, Straus SE. User-centered design of a web-based self-management site for individuals with type 2 diabetes-providing a sense of control and community. BMC Medical Informatics and Decision Making 2014;14:60.

[11] Quinn S, Bond R, Nugent C. An ontology based approach to the provision of personalized patient education. 6th International Work-Conference on Ambient Assisted Living and Daily Activities 2014. 67 p.

[12] A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools Edition Edition 1.3. [Internet]. Horridge M; 2011 - [cited 2016 Dec 2]; Available from: http://mowl-power.cs.man.ac.uk/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf

[13] Uschold M, Gruninger M. Ontologies: Principles, methods and applications. The Knowledge Engineering Review 1996;11:93-136.

[14] Ye J, Coyle L, Dobson S, Nixon P. Ontology-based models in pervasive computing systems. The Knowledge Engineering Review 2007;22:315-347.

[15] Zhou L. Ontology learning: State of the art and open issues. Information Technology and Management 2007;8:241-252.

[16] Konopka BM. Biomedical ontologies—A review. Biocybernetics and Biomedical Engineering 2015;35:75-86.

[17] Arwan A, Priyambadha B, Sarno R, Sidiq M, Kristianto H. Ontology and semantic matching for diabetic food recommendations. 2013 International Conference on Information Technology and Electrical Engineering (ICITEE); 2013. 170 p.

[18] Chen R, Huang Y, Bau C, Chen S. A recommendation system based on domain ontology and SWRL for anti-diabetic drugs selection. Expert Syst Appl 2012;39:3995-4006.

[19] Chammas N, Juric R, Koay N, Gurupur V, Suh SC. Towards a software tool for raising awareness of diabetic foot in diabetic patients. 2013 46th Hawaii International Conference on System Sciences; 2013. 2646 p.

[20] Rahimi A, Liaw S, Taggart J, Ray P, Yu H. Validating an ontology-based algorithm to identify patients with type 2 diabetes mellitus in electronic health records. Int J Med Inf 2014;83:768-778.

[21] Alharbi RF, Berri J, El-Masri S. Ontology based clinical decision support system for diabetes diagnostic. 2015 Science and Information Conference (SAI); 2015. 597 p.

[22] Pinto HS, Martins JP. Ontologies: How can they be built? Knowledge and Information Systems 2004;6:441-464.

[23] Hoehndorf R, Dumontier M, Gkoutos GV. Evaluation of research in biomedical ontologies. Brief Bioinform 2013;14:696-712.

[24] Brank J, Grobelnik M, Mladenic D. A survey of ontology evaluation techniques. Proceedings of the conference on data mining and data warehouses (SiKDD 2005); 2005. 166 p.

[25] Delir Haghighi P, Burstein F, Zaslavsky A, Arbon P. Development and evaluation of ontology for intelligent decision support in medical emergency management for mass gatherings. Decis Support Syst 2013;54:1192-1204.

[26] Bright TJ, Yoko Furuya E, Kuperman GJ, Cimino JJ, Bakken S. Development and evaluation of an ontology for guiding appropriate antibiotic prescribing. J Biomed Inform 2012;45:120-128.

[27] Gordon CL, Weng C. Combining expert knowledge and knowledge automatically acquired from electronic data sources for continued ontology evaluation and improvement. J Biomed Inform 2015;57:42-52.

[28] Isern D, Sánchez D, Moreno A. Ontology-driven execution of clinical guidelines. Comput Methods Programs Biomed 2012;107:122-139.

[29] Valls A, Gibert K, Sánchez D, Batet M. Using ontologies for structuring organizational knowledge in home care assistance. Int J Med Inf 2010;79:370-387.

[30] NHS Choices. Your health, your choices. [Internet]. Department of Health; 2016 - [cited 2016 Dec 2]; Available from: http://www.nhs.uk/pages/home.aspx

[31] https://www.surveymonkey.com/s/WMZ527D

[32] MedlinePlus. Trusted information for you. [Internet]. U.S. National Library of Medicine; 2016 - [cited 2016 Dec 2]; Available from: http://www.nlm.nih.gov/medlineplus/

[33] HON. [Internet]. Health On the Net Foundation; 2016 - [cited 2016 Dec 2]; Available from: https://www.healthonnet.org/

[34] Compendium of Physical Activities. [Internet]. Arizona State University; 2016 - [cited 2016 Dec 2]; Available from: https://sites.google.com/site/compendiumofphysicalactivities/

[35] Hossain M. Users' motivation to participate in online crowdsourcing platforms. 2012 International Conference on Innovation Management and Technology Research; 2012. 310 p.

[36] Duncan I, Birkmeyer C, Coughlin S, Li QE, Sherr D, Boren S. Assessing the value of diabetes education. Diabetes Educ 2009;35:752-760.

[37] Yu AC. Methods in biomedical ontology. J Biomed Inform 2006;39:252-266.

[38] Dramé K, Diallo G, Delva F, Dartigues JF, Mouillet E, Salamon R, Mougin F. Reuse of termino-ontological resources and text corpora for building a multilingual domain ontology: An application to Alzheimer's disease. J Biomed Inform 2014;48:171-182.

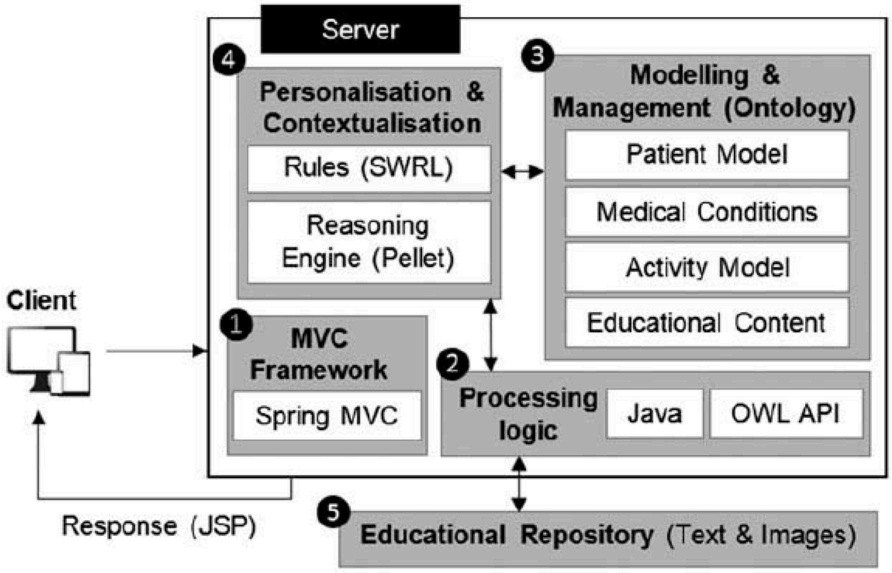

Fig 1.

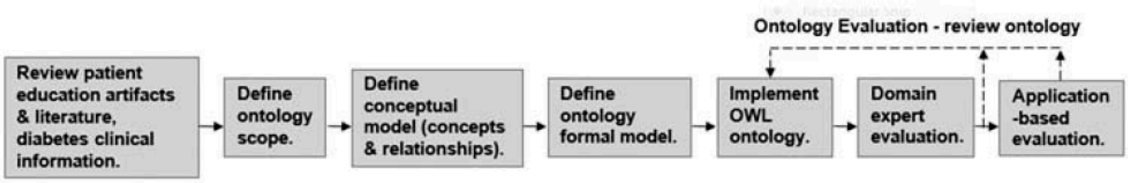

Fig 2.

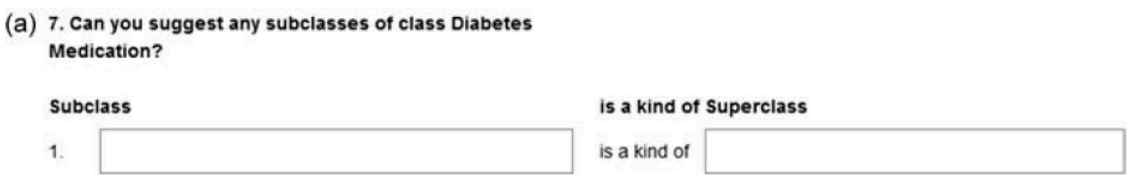

Fig. 3a

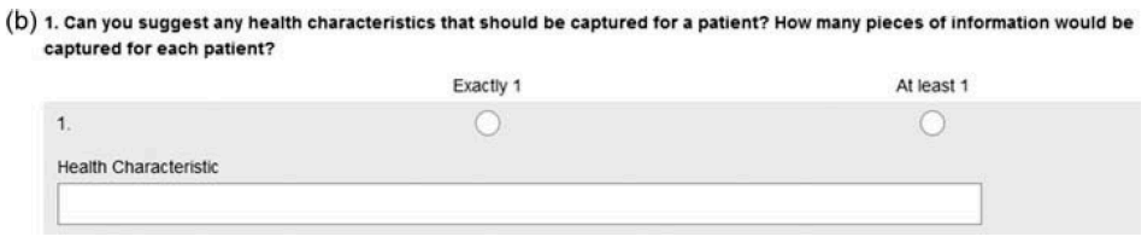
(b) 1. Can you suggest any health characteristics that should be captured for a patient? How many pieces of information would be captured for each patient?
Exactly 1
At least 1
1.
Health Characteristic

Fig 3.b

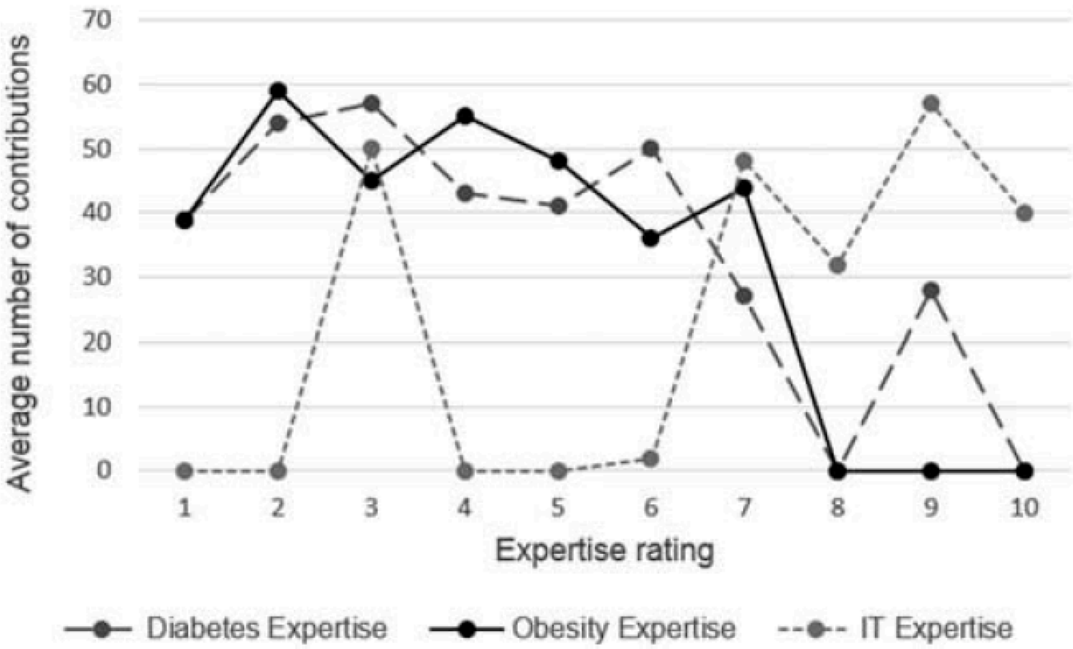
Average number of contributions
Expertise rating
Diabetes Expertise
Obesity Expertise
IT Expertise
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10

Fig. 4

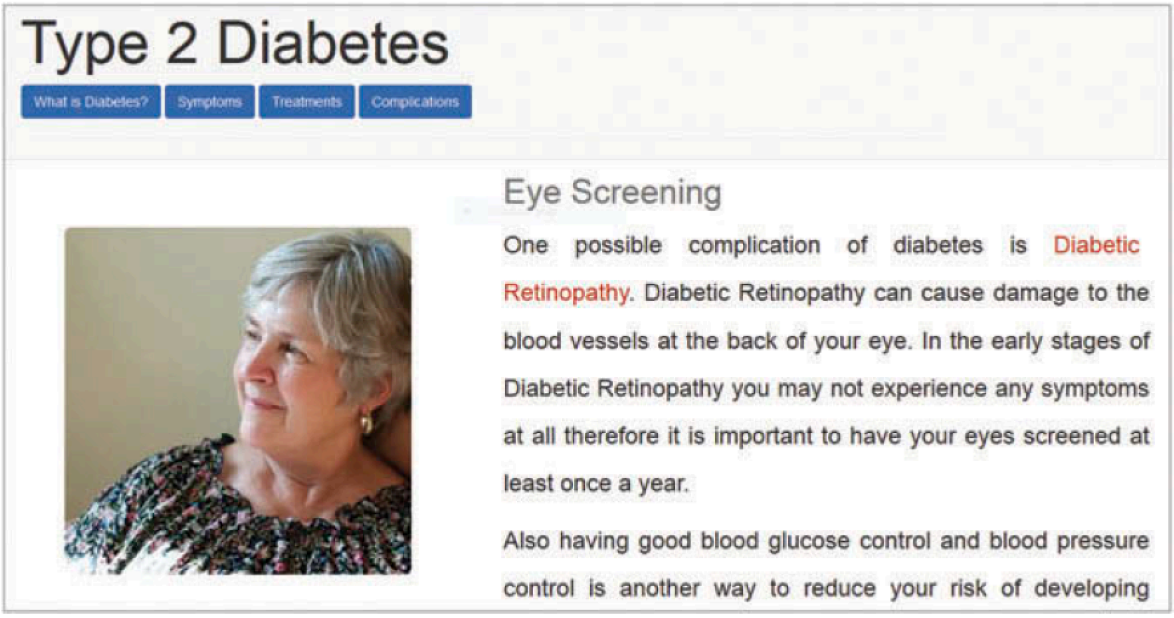
Type 2 Diabetes
What is Diabetes?
Symptoms
Treatments
Complications
Eye Screening
One possible complication of diabetes is Diabetic Retinopathy. Diabetic Retinopathy can cause damage to the blood vessels at the back of your eye. In the early stages of Diabetic Retinopathy you may not experience any symptoms at all therefore it is important to have your eyes screened at least once a year.
Also having good blood glucose control and blood pressure control is another way to reduce your risk of developing

Fig 5.