An ontology for Autism Spectrum Disorder (ASD) to infer ASD phenotypes from Autism Diagnostic Interview–Revised data

Omri Mugzach, BA[1], **Mor Peleg, PhD**[1], **Steven C. Bagley, MD, MS**[2], **Stephen J. Guter, MA**[3], **Edwin H. Cook, MD**[3], and **Russ B. Altman, MD, PhD**[2]

[1]Department of Information Systems, University of Haifa, Israel, 3498838

[2]Department of Genetics, Stanford University, Stanford, CA, 94305

[3]Department of Psychiatry, University of Illinois at Chicago, Chicago, IL 60608

Abstract

Objective—Our goal is to create an ontology that will allow data integration and reasoning with subject data to classify subjects, and based on this classification, to infer new knowledge on autism spectrum disorder (ASD) and related neurodevelopmental disorders (NDD). We take a first step toward this goal by extending an existing autism ontology to allow automatic inference of ASD phenotypes and *Diagnostic & Statistical Manual of Mental Disorders* (DSM) criteria based on subjects' Autism Diagnostic Interview-Revised (ADI-R) assessment data.

Materials and Methods—Knowledge regarding diagnostic instruments, ASD phenotypes and risk factors was added to augment an existing autism ontology via Ontology Web Language class definitions and semantic web rules. We developed a custom Protégé plugin for enumerating combinatorial OWL axioms to support the many-to-many relations of ADI-R items to diagnostic categories in the DSM. We utilized a reasoner to infer whether 2642 subjects, whose data was obtained from the Simons Foundation Autism Research Initiative, meet DSM-IV-TR (DSM-IV) and DSM-5 diagnostic criteria based on their ADI-R data.

Results—We extended the ontology by adding 443 classes and 632 rules that represent phenotypes, along with their synonyms, environmental risk factors, and frequency of comorbidities. Applying the rules on the data set showed that the method produced accurate results: the true positive and true negative rates for inferring autistic disorder diagnosis according to DSM-IV criteria were 1 and 0.065, repectively; the true positive rate for inferring ASD based on DSM-5 criteria was 0.94.

Discussion—The ontology allows automatic inference of subjects' disease phenotypes and diagnosis with high accuracy.

All communication should be with: Omri Mugzach, Department of Information Systems, University of Haifa, 3498838, Israel, Rabin Building, room 7047, omrimug@gmail.com, Phone: 972-52-324-3229, Fax: 972-4-828-8522.

Conclusion—The ontology may benefit future studies by serving as a knowledge base for ASD. In addition, by adding knowledge of related NDDs, commonalities and differences in manifestations and risk factors could be automatically inferred, contributing to the understanding of ASD pathophysiology.

Graphical Abstract
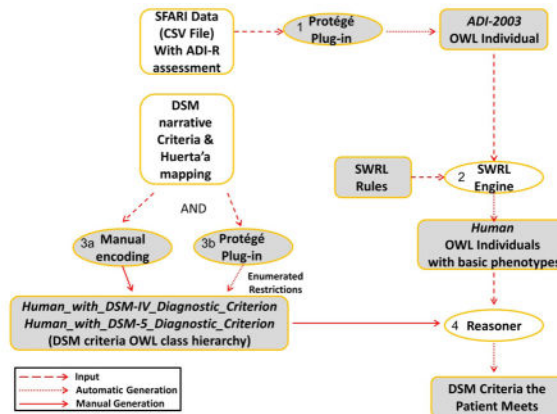
Keywords

ontology; autism; Ontology Web Language; reasoning; diagnosis

1. INTRODUCTION

Understanding the disease processes of complex neurodevelopmental disorders (NDDs), such as Autism Spectrum Disorder (ASD) [1, 2], has been a focus of research for many years. An ability to organize and semantically integrate subject data concerning phenotypic manifestations as well as genetic and environmental risk factors among cohorts of ASD subjects [3, 4] could yield important new knowledge regarding commonalities and differences that characterize subtypes of ASD, and also help elucidate the processes underlying the development of the disorder, whose mechanisms are still unknown. In the long run, comparing manifestations, comorbidities, and risk factors among subjects with related psychiatric disorders (e.g., ASD, depression, bipolar disorder, and schizophrenia) could uncover additional clues regarding the mechanisms of action of ASD and its subtypes. In addition, monitoring the incidence of ASD phenotypes, its subtypes, and the burden of associated comorbidities [5, 6] – along the lines of similar efforts with other diseases, such as diabetes [7] – could help public health efforts to estimate the toll of the disorder on the healthcare system, as well as to evaluate the impact of care on its prevention, progression, and treatment. The presented ontology provides the ability to automatically infer such phenotypes from autism diagnostic instrument data.

When properly designed for such tasks, ontologies aid data integration for cohort-level analysis, as well as reasoning at a single-subject level for the purpose of guiding treatment. They can help standardize data and knowledge about complex diseases and their discourse, and support reasoning tasks for studying them, as has been demonstrated for other

neurogenerative disorders (e.g., Alzheimer's and Parkinson's [8]). Potential relevant data sources for such ontologies include formal databases (e.g., for ASD, the Simons Foundation Autism Research Initiative [SFARI], http://sfari.org/, and the National Database for Autism Research [NDAR] http://ndar.nih.gov/), data from subjects' social networks (e.g., PatientsLikeMe, www.patientslikeme.com), data extracted from hospital and clinic electronic health records [9], and the scientific literature.

Our long-term goal is to elucidate the mechanisms of action of ASD and its subtypes in order that practitioners might better guide and direct patients' treatment. The ontology presented here takes a step toward this goal by enabling formal representation and integration of important data and knowledge about ASD and related NDDs. In the present study, we focus on supporting automatic inference of subjects' ASD manifestations (phenotypes) and diagnosis based on autism assessment data. The diagnostic criteria formally defined in our ontology are taken from the accepted standard as defined by the 4th [10] and 5th editions [11] of the DSM. We have integrated data from SFARI concerning autism assessment results from the Autism Diagnostic Interview–Revised [12, 13].

1.1. RELATED WORK

ASD is a NDD. It was initially described as a disorder comprising repetitive behavior and deficiencies in social interaction and communication capabilities [14, 15].

1.1.1. ASD classification and diagnosis—The DSM [11, 16] is considered the standard classification of mental disorders in the USA [17]. The DSM references corresponding codes from the World Health Organization International Classification of Diseases (ICD, of which the most recent version is the 10th revision, or ICD-10). For each disorder listed, the DSM presents a set of diagnostic criteria which specifies what symptoms must be present and what other conditions must hold for the disorder to be diagnosed. The DSM-IV [10] listed four separate categories of autism spectrum disorder: autistic disorder, childhood disintegrative disorder, pervasive developmental disorder-not otherwise specified (PDD-NOS), and Asperger Syndrome (Asperger's) [10, 19]. The most recent version of the DSM, the 5th edition (DSM-5) [11], treats ASD as a single diagnostic category that may differ in severity and associated features. The DSM-5 also reduced the number of core domains underlying ASD from three domains in the DSM-IV (impaired social interaction, impaired social communication, and restricted behavior patterns) to two, by combining impaired social interaction and communication into a single core deficit [18].

DSM criteria are hierarchical, such that criteria at different levels reflect phenotypes at differing levels of granularity. The DSM-IV criteria for autistic disorder consist of three levels (see http://iancommunity.org/cs/autism/dsm_iv_criteria). The lowest-level (L3) criteria are single specific phenotypes or Boolean combinations of specific phenotypes (e.g., DSM-IV criterion A1(d): "lack of social or emotional reciprocity"). We later refer to L3 criteria as "basic phenotypes". The mid-level (L2) criteria represent the category to which the L3 phenotypes belong. For ASD, Level 2 includes three categories, each manifested by specific L3 phenotypes (e.g., the phenotype A1(d) mentioned above belongs to category A1, "qualitative impairment in social interaction"). Finally, the upper-level (L1) criteria

represent broad standards that incorporate the lower levels as well as diagnostic criteria not captured by the L2 and L3 phenotypes. ASD has three L1 criteria, where the first (criterion A) relates to the L2 and L3 phenotypes, the second (B) relates to the subject's past history (age of symptom onset), and the third (C) serves to exclude alternative diagnoses. With respect to the first upper-level criterion (A), this is defined by a count of the lower-level criteria met: the subject must meet at least six L3 criteria, with at least two from A1 and at least one each from A2 and A3, to meet the upper-level criterion. All three L1 criteria must be met to obtain a diagnosis of autistic disorder.

For its 5th edition, the DSM simplified the criteria for ASD. The DSM-5 criteria are still hierarchical, but they consist of two levels instead of three. The upper level (L1) contains five criteria which must be met to satisfy an ASD diagnosis. The first two criteria (A and B) include a lower hierarchical level, here referred to as L2. The lower-level (L2) criteria (phenotypes) are specific deficits or Boolean combinations of specific deficits (which we later refer to as "basic phenotypes"). The remaining three upper-level (L1) criteria (C, D & E) capture, again, aspects of the disorder not reflected in specific phenotypes – the subject's developmental history (C) and effects of symptoms on functioning (D) – while ruling out alternative diagnoses (E). For example, DSM-5 ASD criterion C states that "Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities, or may be masked by learned strategies in later life)".

The most widely-used instruments for diagnosing ASD are the Autism Diagnostic Interview–Revised (ADI-R) [12, 13] and the Autism Diagnostic Observation Schedule (ADOS) [21, 22]. The ADI-R is a structured interview conducted with the subject's parent or caregiver. It consists of 93 items covering the subject's full developmental history, divided into seven domains: early development (7 items), acquisition and loss of language/ other skills (20), language and communication functioning (21), social development and play (17), interests and behaviors (13), general behaviors (14), and any other current concerns (1). Items are scored using an algorithm provided with the instrument [12]. The ADOS is an observation instrument based on a series of structured and semi-structured tasks involving social interaction between the examiner and the subject. The examiner observes the subject's behavior and uses an algorithm to score behaviors in pre-defined categories, including social reciprocity (the ability to respond to another's actions), restricted and repetitive behaviors, and communication, as well as behavior difficulties not specific to ASD. The ADOS was originally developed to accompany the ADI-R. Both instruments are based on DSM-IV criteria [12, 21] and offer the capability of quantifying severity within certain domains. This approach to measuring social deficits reduces the likelihood that an individual will receive an ASD diagnosis based on severe deficits in only one or two domains. Hence, these instruments are closer, conceptually, to the DSM-IV criteria than to those of the DSM-5 [21]. While the ADI-R score alone is usually sufficient for correctly diagnosing ASD [22], the combination of both instruments is deemed the gold standard for diagnosis [25, 26].

Previous studies have suggested that ASD might be diagnosed using methods designed to measure activity within the brain, such as electroencephalography (EEG) [25] and magnetic resonance imaging (MRI) [26]. Eldridge et al. suggested that statistical analysis of EEG

recordings for neural sensory reactivity is a potential approach to the automatic classification of ASD [25]. They report that their method accurately identified ASD in 79% of cases. Zhou et al. employed graph theory and machine learning analysis of MRI data to characterize and predict ASD with 70% accuracy [26]. The authors found this method more suitable for providing biomarkers for prognosis or monitoring disease progression than for diagnosis. Both methods have relatively low diagnostic accuracy rates and so cannot substitute for the ADI-R.

1.1.2. Ontologies for ASD—Tu et al. [27] created an ASD ontology that follows the principles of ontology development established by the Open Biomedical Ontologies Foundry (www.obofoundry.org) and that conforms to the Basic Formal Ontology (BFO) [28]. BFO is an upper-level ontology designed to support information retrieval, analysis and integration. It promotes a realism-based approach to ontology modeling, which holds that classes in an ontology are universal categories of objects that represent things and processes in reality. Tu et al.'s ontology supports the annotation and integration of scientific data for the purpose of enabling user queries and inferences about ASD-related phenotypes from the NDAR repository. This ontology holds 34 classes representing phenotypes and four classes representing ASD diagnostic instruments (the ADI-R [12], two modules from the ADOS [19], [20], and the Vineland Survey Interview [29]). Additionally, the ontology includes a set of 15 SWRL rules [30] which allow the inference of certain phenotypes for a given patient based on data from the represented ASD diagnostic instruments. Unlike our ontology, this ontology does not include DSM criteria, and so it does not support diagnosis of ASD and inference of ASD phenotypes.

Another ASD-related ontology, created by McCray et al. [31], is an ASD-phenotype ontology developed for the purpose of assessing and comparing the characteristics of ASD diagnostic instruments, calculating diagnostic instrument coverage for the purpose of yielding more accurate diagnosis, and querying data sources based on the ontology terms rather than individual diagnostic instrument terms. The authors grouped questions with similar content from over 24 ASD diagnostic instruments to create a hierarchy of derived phenotypes along with their mapping (codes) to standard ontologies. This work relates to our work as it involves the creation of a new ASD-related phenotypic hierarchy. However, McCray et al. do not include DSM-IV or DSM-5 criteria phenotypes in their hierarchy, and their ontology does not support reasoning over DSM criteria. We have integrated McCray's phenotypic hierarchy into our ontology to yield a more complete phenotypic hierarchy of ASD and to map the phenotypes to standard vocabulary codes.

Additional ontologies related to ASD have been developed to support other use-cases, such as automated ontology construction by performing text-mining [32, 33] or ontology-based information retrieval [34, 35]. The aforementioned ontologies differ from the work described in this paper as they were created in an automatic process to support their ultimate goal of text-mining or information retrieval. The ontology described in this paper was created manually, and it focuses on automatic diagnosis and inference of disease phenotypes and risk factor categories.

1.1.3. Studies analyzing risk factors, comorbidities, and overlap in manifestations between ASD and other NDDs

—Our work is distinguished by its emphasis on using ontologies for explicit, declarative representation of the domain elements and relationships. Most previous work on inference over risk factors, comorbidities, and phenotypes has used probabilistic methods. For example, Rzhetsky et al. [36] used statistical models applied to Electronic Health Record (EHR) data to infer genetic overlap between complex disorders, including autism, bipolar disorder, and schizophrenia. Kohane et al. [5] identified the conditional probability of comorbidities related to ASD using ICD-9 codes from EHR data. Similarly, Peacock et al. [6] queried medical multi-state databases to detect co-occurring conditions among ASD subject records. Finally, Lyalina et al. [37] used Fisher's Exact Test to find enriched associations between all pairs of enriched phenotypes (comorbid diagnoses and symptoms) for autism, bipolar disorder, and schizophrenia.

2. MATERIALS AND METHODS

This section describes the data sources and methods used for the construction and validation of the presented ontology. This ontology extends Tu's [27] ontology by adding an ability to infer ASD phenotypes based on DSM criteria. In addition, it integrates McCray's phenotypic hierarchy [31] into the BFO hierarchy used in our ontology.

2.1. Data sources

Data were obtained for 2642 subjects who completed the ADI-R questionnaire and each had a *CPEA*[1]*-dx* value, which combines the clinicians' best estimate with the diagnostic instruments' score. According the *CPEA-dx* value, 2394 subjects had autistic disorder, 196 had ASD and 52 had Asperger's. The data were obtained from SFARI, a scientific research program within the Simons Foundation (http://www.simonsfoundation.org) that aims to increase scientific understanding of autism spectrum disorders and improve their diagnosis and treatment. SFARI granted access to the data after we obtained ethics approval for this research from the University of Haifa. The obtained data included complete ADI-R item-level scores for all subjects.

Information regarding frequencies and prevalence of ASD-related phenotypes and comorbidities [5,31,37], along with information regarding risk factors, were obtained from the literature (the latter by co-author SCB) [38–44]. Additionally, relevant synonyms corresponding to concepts within the ontology were obtained from the Unified Medical Language System (UMLS), and all standard codes for terms (phenotypes and diseases) included in the ontology were obtained from the ontology by McCray et al. [31].

2.2. Ontology development

There already exists an ontology for autism [27], represented in the Web Ontology Language (OWL) formalism [45, 46]. However, this ontology does not support reasoning over the DSM criteria. In this study, we have augmented that ontology with DSM OWL class definitions, with basic phenotype classes corresponding to ADI-R items, and with a

---

[1]Collaborative Programs of Excellence in Autism (CPEA) – a research network operated by the National Institutes of Health (now superseded by the Autism Centers of Excellence)

complete set of Semantic Web Rule Language (SWRL) rules [27] to infer ASD phenotypes. We chose to represent the DSM criteria explicitly as OWL class definitions rather than use a machine learning algorithm to infer ASD diagnoses because we wanted to generate an explicit knowledge representation that could be comprehended by humans, and that could be used not only to classify patients as having ASD or not, but to infer partial phenotypes according to DSM criteria.

We used the Protégé ontology editor (http://protege.stanford.edu) to develop and extend the ontology developed by Tu et al. [27] to allow inference of DSM autism-related criteria (phenotypes) based on ADI-R data. The ontology is available at BioPortal for Protégé version 4.3 (http://bioportal.bioontology.org/ontologies/ADAR/). Specifically, we added the following:

1. Diagnostic instrument-based basic phenotypes corresponding to all ADI-R items. To ensure compatibility with current standards, we arranged these in a hierarchy corresponding to that created by McCray et al. [31] and adopted their controlled vocabulary codes;

2. SWRL rules deducing these basic phenotypes from coded ADI-R results;

3. OWL classes containing definitions of diagnostic criteria for autistic disorder according to the DSM-IV [10] and for ASD according to the DSM-5 [11]. These formal definitions relate to the basic phenotypes and were created based on expert mapping [9] of the DSM-IV and DSM-5 diagnostic criteria to their corresponding questions (items) in the ADI-R [12]. Using OWL to represent the DSM criteria in terms of ADI-R items, rather than simply using the ADI-R algorithm [12] to diagnose autistic disorder or ASD based on ADI-R results, allows us to utilize OWL's already existing capabilities of Description Logic (DL) [47] for the following:

    a. Using SWRL rules to deduce phenotypic manifestations from ADI-R data. This allows us to display the complete set of ADI-R-based basic phenotypes manifested by a given subject;

    b. Inferring which specific DSM criteria a given subject meets, and the proportion of all subjects who meet any given diagnostic criterion. This is done by executing reasoners [45] to infer which subject instances meet OWL class restrictions;

    c. Comparing DSM criteria between different versions (in this case DSM versions IV and 5). For example, we could automatically infer whether there are subjects who, based on their ADI-R data, meet the DSM-IV autistic disorder diagnostic criteria but not the DSM-5 ASD criteria. Enabling such automatic inference could help researchers understand the underlying phenotypes that are most affected by this shift between diagnostic criteria;

    d. Drawing conclusions about particular correlations concerning ADI-R and DSM. The shift between DSM editions (in this case, IV and 5) might yield differences in how different ASD phenotypes are considered.

To support future data and knowledge integration tasks, we also defined the class structure based on the literature and populated it with instances concerning:

1. Synonyms of the concepts stored in the ontology, so as to allow links to data sources using different vocabularies;

2. Comorbidities of autism and their related prevalence as well as conditional probabilities (frequencies) [5, 37, 6]; and

3. Environmental risk factors for ASD [38–44].

Analysis of comorbidities and environmental risk factors for ASD as compared to other NDDs could help expose potential trends and improve our understanding of the mechanisms underlying ASD.

Figure 1 shows a class diagram depicting the main classes in the ontology.

2.2.1. Supporting DSM definitions and related ASD phenotypes using OWL

2.2.1.1 Basic phenotype representation: Autism-related basic phenotypes are arranged in hierarchies as a sub-tree whose root class is *ASD_related_phenotype*, which is a subclass of BFO's *disposition* class. This is consistent with the original autism ontology of Tu et al., which we have extended. According to BFO, a disposition is a realizable_entity that causes a specific process or transformation in the object in which it inheres, under specific circumstances and in conjunction with the laws of nature (e.g., the disposition of a patient with a weakened immune system to contract disease). This representation is different from that used in the Ontology for General Medical Science (OGMS) [48], where a Phenotype is defined as a combination of one or more Bodily Features (Bodily Component, Bodily Quality, or Bodily Process) of an Organism determined by the interaction of its genetic make-up and environment. This definition is somewhat limiting as it ties phenotypes to the body of an organism, in contrast to the autism ontology of McCray [31] (which we have adopted), where the autism-related phenotypes are personal traits or aspects of social competence. True, the OGMS has in some respects a broader scope than our ontology, as it allows representation of phenotypes that are processes (occurrent) and material objects (independent continuant), and not only dependent continuants such as quality or disposition. However, this broad scope is not necessary for the autism phenotypes associated with ADI-R.

A basic phenotype can be considered a leaf in the autistic phenotype tree. For example, the path from *the ASD_related_phenotype* class to the *ImaginativePlay_NotAvailable* phenotype consists of the following path: *ASD_Related_Phenotype* → ***Personal_Traits*** → ***Cognitive_Ability*** → ***Abstract_Thinking*** → ***Imagination*** → *Imaginative_Play* → *ImaginativePlay_NotAvailable*. The path is shown in Figure 2.

Note that some of the classes above are shown in bold. These classes and their controlled vocabulary codes are taken from McCray's ASD phenotypic hierarchy [31]. We expanded McCray's hierarchy further by adding classes for each ADI-R item that is included in the ADI-R to DSM mapping by Huerta [9] (e.g., *Imaginative_Play*, Figure 2-(2)), along with subclasses which represent each possible response to that ADI-R item (e.g.,

ImaginativePlay_NotAvailable) Additionally, a *Vocabulary Term* instance was added to each subclass as an annotation, when available.

2.2.1.2 DSM criteria hierarchy: As in the DSM, the representations of diagnostic criteria in our ontology are hierarchical, corresponding to different levels of abstraction of DSM-IV autistic-disorder phenotypes and DSM-5 ASD phenotypes. We represented all DSM criteria as OWL classes arranged in a hierarchy stemming from the *Human_with_DSM_Diagnostic_Criterion* class (Figure 3).

The Human class has a property called has_most_abnormal_finding (see Figure 1) that relates each Human individual to the basic phenotypes he exhibits. These basic phenotypes populate this property of the Human individual using an inference made by SWRL rules based on data from the ADI-R items, as explained in Section 2.2.1.3. Then, based on the OWL axioms that define the *Human_with_DSM_Diagnostic_Criterion* classes (Section 2.2.1.4), individuals are classified by a Reasoner according to the DSM criteria that hold for each individual based on his abnormal findings (basic phenotypes corresponding to ADI-R items).

The original narrative DSM hierarchy includes is-a and part-of relationships, which we captured in the *Human_with_DSM_Diagnostic_Criterion* hierarchy (Figure 3):

1. ***IS-A*** relationships were defined between DSM-IV's mid-level (DSM-IV L2) and lower-level (DSM-IV L3) criteria. The same structure was implemented for DSM-5's upper-level (DSM-5 L1) and lower-level (DSM-5 L2) criteria, which relate to DSM-5 criteria A and B. This type of hierarchical relationship states that each of the lower-level criteria extend more abstract criteria. For example, DSM-IV mid-level diagnostic criterion A1 relates to "qualitative impairment in social interaction." DSM-IV lower-level criterion A1(b) is "failure to develop peer relationships appropriate to developmental level," which is a specific impairment in social interaction.

2. In order to meet DSM-IV mid-level (L2) criterion A1, a subject has to meet at least two lower-level (L3) criteria from (A1(a) – A1(d)). Similarly, in order to meet DSM-5 upper-level (L1) criterion B, a subject has to meet at least two lower-level (L2) criteria from (B(1) – B(4)). Therefore, these lower-level criteria are not specializations of their respective mid- or upper- level criteria DSM-IV A1 or DSM-5 B. Instead, we defined a ***Part-Of*** relationship between them. This type of hierarchical relationship states that several lower-level criteria are the parts which create the higher-level criteria.

We did not use the part-of relation for grouping other L3 criteria because their L2 criteria require only one L3 criterion to hold. Hence, these L3 criteria have an is-a relationship to L2 criteria classes. While this is a non-uniform organization, it is correct and corresponds well to the textual representation of the narrative DSM criteria, facilitating comprehension by domain experts.

2.2.1.3 Populating _Human_ instances with basic phenotypes: For each subject who has completed an ADI-R assessment, we create a corresponding _Human_ instance. In order to populate these _Human_ instances with their relevant basic phenotypes, we utilize SWRL rules. A SWRL rule is created for each possible answer to each relevant ADI-R item. The SWRL rules add basic phenotypes to already existing _Human_ instances by mapping each possible answer of the ADI-R (named _ADI-2003_ in Tu's ontology) instance to the corresponding basic phenotype in the ontology. The execution of all SWRL rules results in one or more _Human_ instances populated by all their relevant basic phenotypes (Figure 4).

All SWRL rules consist of two parts: a criterion component (labels A–D of Figure 4) and an action component (section (E) of Figure 4). The criterion component is defined as follows:

A. Retrieve the subject identifying key from the _ADI-2003_ instance;

B. Retrieve the coded ADI-R answer value from the relevant property representing that ADI-R item in the _ADI-2003_ instance;

C. Check whether the ADI-R answer value in the property equals the value required by the specific SWRL rule;

D. Identify the _Human_ instance intended to be populated according to the subject coded key taken from the _ADI-2003_ instance (label D and the first part of label E of Figure 4).

The action component defines the target _Human_ instance property to which the basic phenotype will be added by specifying the relevant property (_has_general_finding, has_most_abnormal_finding_, _has_current_finding_, _has_ever_finding_, and all possible temporal sections corresponding to those of the ADI-R), the _Human_ instance to which it will be added, and the basic phenotype itself, as shown in label E of Figure 4.

Several diagnostic concepts have temporal connotations. According to the mapping by Huerta, the meaning of some ADI-R items (and hence their mapping to basic phenotypes) differs for three age groups: patients <4 years old, 4–10 years old, and >10 years. A simple representation of time was adopted where the property has_general_finding can hold the age category (e.g., Human_with_DSM-IV_definition_A_1_a has_general_finding has SubjectAge_Under_4_Years). In addition, some mappings relate to the most abnormal findings manifested at age 4–5 years. This is represented by the property has_most_abnormal_finding of Human (Figure 1).

2.2.1.4. Defining DSM diagnostic criteria that refer to basic phenotypes: We based our mapping of the DSM criteria to basic phenotypes corresponding to ADI-R items on Huerta et al. [9]. Huerta et al. aimed to evaluate DSM-5 criteria for ASD among children with a DSM-IV diagnosis of pervasive developmental disorder (PDD) using ADI-R and ADOS results to compare DSM-5 and DSM-IV diagnostic criteria. This work yielded a mapping of ADI-R and ADOS items to their corresponding criteria in the two versions.

Huerta's mapping specifies the relevant ADI-R items for each DSM lower-level (DSM-IV L3 and DSM-5 L2) diagnostic criterion (Figure 5). The described ADI-R items are translated into logical OWL expressions (section (2) of Figure 5 and Figure 6) referring to

basic phenotypes, according to the narrative DSM definition. For example, the class *Human_with_DSM-IV_definition_A_2_a* (Figure 6) provides the definition for a lower-level (L3) DSM criterion (phenotype) which involves a Boolean combination of basic phenotypes referring to ADI-R items 30 (overall level of language), 43 (nodding), 44 (head shaking), 45 (conventional/instrumental gestures), and 50 (direct gaze).

OWL classes corresponding to DSM diagnostic criteria L1–L3 are defined as necessary and sufficient OWL class restrictions. The lower level (L3) combines basic phenotypes with logical operators, while the upper level (L1) and middle level (L2) refer to the lower-level (L1) criteria.

The L1 criteria of DSM-IV involve counting the number of L3 criteria from specific L2 criteria. For example, as described above, the DSM-IV Level-1 criterion A states that the subject must meet at least six L3 criteria from A1, A2, and A3, with at least two from A1 and one each from A2 and A3. This requires the support of *k-of-N* counting. Since OWL reasoners cannot perform *k-of-N* counting, we developed a Protégé plugin (section (3) of Figure 5) that produces appropriate class restrictions for different *k-of-N* combinations. The plugin utilizes the Protégé API to access relevant OWL classes in order to insert enumerated combinations of *k-of-N* classes as necessary and sufficient axioms into these classes. We used our developed plugin to add the relevant class restrictions to the represented L2 and L1 DSM criteria.

The plugin utilizes the capabilities already present in description logics reasoners in order to infer which subject instances meet DSM criteria. This method makes the ontology more maintainable and general, and does not necessitate development of reasoning capabilities. The Protégé plugin that we have developed can be reused for other OWL ontologies, since it is not specific to ASD. The plugin enables the enumeration of a set of axioms that captures the combination of k of N classes based on a selection of the number k and the N superclass, retrieved from the user, which will then be added to the definition of the relevant OWL class.

2.2.1.5 A summary of the inference method of ASD-related phenotypes from SFARI data: Figure 7 describes the execution flow of the methods used for inferring ASD-related phenotypes from SFARI data, using the following steps:

1. The subject records in the SFARI dataset were obtained as a comma-separated file. Our Plugin for the Protégé ontology editor converted all data from the SFARI comma-separated file into appropriate OWL individuals of the *ADI-2003* class from Tu's original ontology, representing the ADI-R assessment results for each subject. During this conversion, a corresponding individual of the *Human* class was created for each *ADI-2003* assessment_result, initially with no phenotypes;

2. SWRL rules were executed to infer the relevant basic phenotypes for the *Human* individual based on their relevant ADI-R score values (Figure 4);

3. DSM criteria were represented using the classes *Human_With_DSM-IV_Diagnostic_Criteria* and *Human_With_DSM-5_Diagnostic_Criteria* (DSM criteria class hierarchy) corresponding to the hierarchy of both DSM-IV and

DSM-5, along with relevant definitions (OWL class restrictions) corresponding to logical combinations of the relevant basic phenotypes. Additionally, the Protégé *k-of-N* plugin was used to enumerate restrictions for the middle (L2) and upper-level (L1) classes.

4. We utilized the capabilities of the Pellet OWL reasoner (http://www.clarkparsia.com/pellet/), software able to infer logical consequences from a set of asserted facts or axioms, to deduce which subject instances fulfill which DSM criteria (OWL class restrictions). These results were tabulated by the plugin.

2.2.2 Representing autism related concepts and their synonyms—Each vocabulary concept is represented as an individual of the *VocabularyTerm* class, which holds pointers to the preferred concept and to its synonyms (see Figure 8). The preferred concept and its synonyms are represented as individuals of a class called Concept, which includes the concept's preferred name along with its controlled vocabulary code and the preferred vocabulary name.

Vocabulary Term and *Frequencies* (see below). Individuals were added as annotation properties to their corresponding phenotypes. Using annotation properties allows us to populate classes with knowledge that describes them, rather than knowledge which defines them. This way, we can both define the rules for qualifying as a member of a certain class, and describe the class itself.

2.2.3. Populating the ontology with knowledge regarding comorbidities and risk factors—In addition to the basic phenotype hierarchy and the DSM definitions hierarchy, we included in the ontology information about the frequency of comorbidities and risk factors for ASD. The *Autism High Level Visualizer* individual (Figure 9) displays the added knowledge about autism comorbidities and risk factors.

Two types of frequencies were represented in the ontology:

1. Conditional probability frequencies. These describe the probability of having a certain comorbidity (e.g., epilepsy) given an autism diagnosis P(co | autism), and the probability of being diagnosed with autism given a diagnosis of a certain autism comorbidity P(autism | co). Figure 10 shows an example of a conditional probability individual;

2. The prevalence of a certain condition in a given population – for example, the percentage of subjects diagnosed with autism out of all subjects in a given medical institution [15].

All data related to probabilities were gathered from the literature [5, 37] and inserted into the relevant instances of the frequency classes. Note that all concepts in the ontology, including the comorbidity Autoimmune_Disease and Autism, are part of the ontology. (Figure 8 presents the autism concept along with its UMLS code and synonyms.)

Environmental risk factors for ASD are represented in the ontology using the *Risk_Factor* class (see Figure 11). This class represents: (1) The exposure which is believed to have

influenced the manifestation of the disorder; (2) the time period of the exposure (e.g., during pregnancy, delivery); (3) the subject who was exposed (i.e., mother or child); (4) the exposure class (type), whose possible values are subclasses of BFO's *process* class (*Disease or Syndrome, Natural Process or Phenomenon NOS, or Injury or Poisoning, Obstetric Complications*); and (5) related citations.

2.4. Validation

In order to validate our representation of both DSM-IV criteria for autistic disorder and DSM-5 criteria for ASD, we used subject data from the SFARI dataset. The dataset holds coded data for 2642 subjects, including responses for all items in the ADI-R. All subjects in the dataset were diagnosed with autism, ASD, or Asperger's by expert clinicians. Each subject's final diagnosis is given by the SFARI *CPEA_dx* variable. This variable provides a diagnostic classification based primarily upon (1) the ADI-R, (2) the ADOS, and (3) the Clinician's Best Estimate diagnosis (classification of Asperger's also takes into consideration other values, as specified below). To receive a diagnosis of "autism" in the *CPEA dx* variable, the ADI-R classification must be Autism, the ADOS classification must be Autism or ASD, and the Clinician's Best Estimate diagnosis must be Autism, ASD, or Asperger's. To receive a diagnosis of "Asperger's," an individual must not meet criteria for Autism as specified above, and must meet the following: (1) Chronological Age 60 months; (2) Verbal IQ 80; (3) Age of First Words 24 months; (4) Age of First Phrases 33 months; (5) ADI-R classification is NOT Autism; (6) ADI-R RRB Total 2; (7) ADI-R Social Total 10; (8) ADOS classification is autism or ASD OR ADOS Social-Communication Total 4; and, (9) the Clinician's Best Estimate diagnosis must be autism, ASD, or Asperger's. To receive a diagnosis of "ASD," the individual must not meet the above criteria for autism or Asperger's, the ADI-R classification must be ASD [24], the ADOS classification must be ASD, and the Clinician's Best Estimate diagnosis must be autism, ASD, or Asperger's. If none of the above criteria are satisfied, the diagnosis of "NonSpectrum" is assigned.

All data were processed and automatically inserted into the ontology by creating a new instance of the *ADI-2003* class containing all data for each subject. Following this, we executed all SWRL rules, creating instances of the *Human* class for each subject from the dataset, and used the Pellet reasoner to infer which diagnostic criteria were met by each subject. We then tabulated the results using our Protégé plugin.

For each subject we compared the top-level DSM diagnosis inferred by the reasoner to the *CPEA_dx* variable, which served as our gold standard. For DSM-IV, the inferred top-level diagnosis could be "autistic disorder" or "not autistic disorder", while the *CPEA_dx* variable could have the values "autistic disorder", "ASD" or "Asperger's". Subjects who were inferred by the reasoner as having "autistic disorder" and who had a diagnosis of "autistic disorder" according to the gold standard were considered as true positives. Subjects who were inferred by the reasoner as NOT having "autistic disorder" and who had a diagnosis of "ASD" or "Asperger's" according to the gold standard were considered as true negatives. Subjects who were inferred by the reasoner as having "autistic disorder" but who had a diagnosis of "ASD" or "Asperger's" according to the gold standard were considered as false

positives. Similarly, subjects who were inferred by the reasoner as NOT having "autistic disorder" but who had a diagnosis of "autistic disorder" according to the gold standard were considered as false negatives.

For DSM-5, the top-level diagnosis could be "ASD" or "not ASD", while all possible values of the *CPEA_dx* variable were considered "ASD" according to the new diagnostic criteria introduced in the DSM-5. Subjects who were inferred by the reasoner as having "ASD" were considered as true positives. Subjects who were inferred by the reasoner as NOT having "ASD" were considered as false negatives. We did not have negative examples in the data set, and so true negatives and false positives could not be calculated.

All cases of false positive and false negative inferences were thoroughly examined by the medical experts to elucidate why some subjects did not meet the ASD diagnosis that was inferred by the reasoner. These findings are explained in the Discussion section below.

2.5. Analysis of the spectrum of DSM-IV and DSM-5 criteria met by subjects

As an initial characterization of the spectrum of DSM sub-criteria exhibited by subjects, we calculated the percentage of subject records that were inferred as satisfying each of the DSM-IV criteria/sub-criteria and the DSM-5 criteria/sub-criteria, and plotted these percentages. From such plots we can learn if there are criteria that are present in almost all subjects, some that are very rare, and some that are exhibited at intermediate level.

3. RESULTS

We created 632 SWRL rules deducing 632 basic phenotypes from ADI-R data (about 5–7 SWRL rules for each of the 93 ADI-R items). All basic phenotypes are represented in a hierarchy of phenotypes along with their controlled vocabulary codes taken from McCray et al. [31]. The extended ontology holds 36 classes representing DSM diagnostic criteria as restrictions defining specific Human subclasses (21 subclasses for DSM-IV and 15 for DSM-5, corresponding to lower-level (L3), middle-level (L2), and upper-level (L1) DSM criteria). DSM-IV Criterion *C* is not included in our mapping. Likewise, DSM-5 criteria *C*, *D* and *E* were not included in our mapping (see Discussion section). Additionally, we added 13 disease (comorbidity) classes, 35 frequencies (24 conditional probabilities and 9 prevalence instances), 110 concepts, 35 vocabulary terms (preferred concepts along with their synonyms), 44 environmental risk factors, and 170 classes representing the phenotypic hierarchy that were retrieved from [31].

As explained in Section 2.1, 2394 of the 2642 SFARI subjects (90.61%) are expected to have a diagnosis of "autistic disorder" according to the DSM-IV criteria, and all are expected to have a diagnosis of ASD according to the DSM-5 criteria (since there were no examples of subjects not having ASD according to the DSM-5 criteria in the database). Following our validation procedure, the results in Table 1 and Table 2 show that for the DSM-IV criteria, the true positive rate was 1 and the true negative rate was 0.0645. For the DSM-5 criteria, the true positive rate was 0.94.

As reported in Table 1, 232 records were falsely inferred as having autistic disorder, with the SFARI *CPEA dx* variable classifying 183 of these with ASD and 49 with Asperger's. This represents a low true negative rate of 0.065.

For the DSM-5 diagnostic criteria, of all 2642 subject records with a DSM-5-based ASD diagnosis, which include DSM-IV's autistic disorder, ASD and Asperger's subjects, 157 records were falsely inferred as not having ASD (false negatives).

Figure 12 and Figure 13 show the percentage of records inferred for each of the DSM-IV criteria/sub-criteria and the DSM-5 criteria/sub-criteria, respectively.

4. DISCUSSION

The existing ontologies representing the domain of ASD [27, 31, 32, 39, 41 focus on the displayed phenotypes and in some cases refer to their diagnostic instruments [27, 31], but do not relate to DSM criteria. In this study we have shown that it is possible to infer, with a few notable exceptions, the set of DSM-IV and DSM-5 ASD phenotypes (criteria) that subjects exhibit by using raw ADI-R data. To the best of our knowledge, there is no automatic tool that can relate specific DSM criteria to specific patients based on ADI-R data.

4.1. Inference of ASD-related phenotypes and its importance

Our ontology enables inference of the sub-criteria of DSM, corresponding to phenotypes related to ASD. Figures 12 and 13 present the initial characterization of the spectrum of DSM sub-criteria exhibited by subjects in the present research. As shown in the figures, a high percentage of subjects met the L2 criteria in both DSM-IV (mid-level criteria) and DSM-5 (criterion A and criterion B). Figure 12 clearly shows that almost none of the subjects met DSM-IV criterion A2(a). In other words, almost no subjects in the SFARI dataset had delays in spoken language, and those who did were able to communicate using alternative methods. Since this criterion is usually met by those with lower mental ages, it would be interesting to examine in future work the relationship between different criteria and mental age. In addition, a relatively low percentage of subjects met DSM-IV criterion A2(d), which assesses the existence of social play (65.4%), and DSM-IV criterion A3(c), which deals with stereotyped and repetitive motor mannerisms (64.4%). The results for DSM-5 (Figure 12) reveal a different picture. It seems that all DSM-5 criteria included in the ontology's DSM-5 criteria representation were met by most subjects.

In future research, subjects could be clustered according to their manifestation of this set of basic phenotypes, partitioning ASD into sub-groups. An even more interesting analysis could match subjects' reported risk factors with these and other ASD-related phenotypes, including comorbidities. It is possible that different risk factor exposures and genetic factors manifested during neonatal development could cause different, yet overlapping, phenotypes. This, in turn, implies that studying the relationships between risk factors and manifestations of ASD subgroups could point to different mechanisms of disorder development. Previous studies have detected sub-groups of ASD [49] which differ in their manifestations [50] and in genetic [51] and environmental risk factors [38–44]. The ontology presented here enables examination of phenotypes and environmental risk factors. Future work could extend the

represented hierarchy for risk factors to also include genetic risk factors. With subject data that includes risk factor information along with manifestations of ASD (phenotypes), future studies might use the information contained in our ontology to find correlations between genetic and non-genetic risk factors, subject's geographical area (locality), and manifestations of ASD phenotypes in an attempt to reveal more clues regarding the disorder's mechanisms of action, as well as its relationship to risk factors and locality.

4.2. Inferring autistic disorder and ASD diagnosis

Automatic inference of a subject's disease phenotypes based on that subject's ADI-R data could enable automatic diagnosis of autistic disorder according to DSM-IV criteria and ASD according to DSM-5 criteria. Future work should consider adding severity to the ontology by defining the proper classes and restrictions, thus allowing more accurate inference of the subject's state and the needed treatment, prognosis, and costs, based not only on the diagnosis per se but also on the severity of the displayed phenotypes.

An important feature of our ontology is its ability to support inference of autistic disorder and ASD-related diagnosis according to DSM criteria, based on the ADI-R interview. We evaluated this feature using the SFARI dataset, and specifically, by comparing the ontology's DSM-based inference to the SFARI *CPEA dx* variable. We used this variable as our gold standard as it draws from three sources: the ADI-R, the ADOS, and expert opinion. We did not expect that the automatic inference based on the DSM-IV or the DSM-5 diagnostic criteria (as reflected in responses to the ADI-R) would fully agree with the diagnosis provided by the SFARI *CPEA dx* variable because the ADI-R algorithm which underlies the *CPEA dx* variable sums up individual items across the social, communication, and restricted and repetitive behavior domains. In other words, ADI-R criteria for these domains (as incorporated in the *CPEA dx*) were met if the sum of items reached a given threshold. In contrast, our ontology followed the rules given in the DSM-IV or DSM-5, whereby scores are keyed to number of sub-domains met (e.g., at least two of four social sub-domains).

Looking at the results of our validation, we found that the DSM-IV diagnostic inference had a true positive rate of 1 but a true negative rate of only 0.065. Subjects who were inferred by the reasoner as having autistic disorder but who had a diagnosis of ASD or Asperger's according to the *CPEA dx* variable were considered as false positives. The DSM-5 ASD diagnostic inference had a true positive rate of 0.94 (the DSM-5 true negative rate could not be computed due to a lack of negative examples, hence even the two true positive rates are not comparable because there is a tradeoff between true positive rate and true negative rate). Our clinical expert coauthors (EHC, SCB and SJG) carefully examined the data, and after a thorough consultation, we concluded that subjects who were false positives may have met DSM-IV "autistic disorder" criteria without meeting the ADI-R algorithm criteria for autism. That is, in order to meet ADI-R algorithm criteria for autism an individual must meet or exceed a cutoff score in each domain area (social interactions, communication, and restricted or repetitive behavior), as well as onset criteria; failure to meet the cutoff in any area precludes meeting the ADI-R algorithm criteria for autism. Thus, it is possible for patients to satisfy enough ADI-R algorithm sub-domain items to warrant a DSM-IV autism

diagnosis without actually meeting ADI-R algorithm domain criteria for autism. Therefore, false inference of ASD or Asperger's as autistic disorder is expected. Risi et al. [24] suggested modifications to the ADI-R algorithm in order to capture those individuals who otherwise would fall within a broader autism phenotype (ASD or Asperger's).

The high rate of false negative DSM-5 inferences, as compared to DSM-IV inferences, can be explained by the differences between how the classification is calculated in the *CPEA dx* variable vs. the DSM-5. The number of subjects who were inferred as having ASD according to the DSM-5 was higher than the number of patients having autistic disorder under the DSM-IV criteria. These results are aligned with previous research which shows an increase in ASD prevalence when using DSM-5 instead of DSM-IV criteria [20, 51], since the DSM-5 relates to a single category of ASD while the DSM-IV relates to several diagnoses, where we have focused on autistic disorder. The reported true positive rate for DSM-5 conforms with previous research which showed that 93% of subjects diagnosed according to the *CPEA dx* variable met DSM-5 criteria [52].

4.3. Integration of different types of subject data

In this research, we used SFARI data related to results of ADI-R assessments. However, our ontology includes additional knowledge, such as synonyms for ASD-related basic phenotypes, comorbidities, and environmental risk factors, which may aid in integrating data from other resources. The importance of risk factors for identification of ASD-related subtypes which may shed light on disease mechanisms was discussed above.

Representing phenotypes along with their controlled vocabulary concepts and synonyms could support natural language processing (NLP) of text-based subject records (such as hospital EHRs or online medical forums such as PatientsLikeMe (www.patientslikeme.com)) in order to extract relevant phenotypes. However, this is not a trivial task. DSM concepts, being complex and abstract, can be expressed in many ways in natural language, and cannot be captured by single terms. For example, consider DSM-IV diagnostic criterion A2(a): "Delay in, or total lack of, the development of spoken lanaguage (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime)." This criterion contains a number of composed terms, including "development of spoken language" and "alternative modes of communication," each of which is open to a large number of variations in free text.

The schema of the proposed ontology, with its standardized hierarchy that builds upon the BFO and that includes the phenotypic hierarchy of McCray et al. [31] and standard vocabulary codes, means it can be extended in the future to hold knowledge regarding other related NDDs such as schizophrenia. Based on such knowledge, the ontology could be used in a variety of studies comparing ASD and other NDDs. For instance, in a study unrelated to ASD, Tu et al. [53] utilized OWL's reasoning mechanism to compare eligibility criteria for different clinical trials. Likewise, we propose to compare diagnostic criteria of different diagnostic instruments in order to identify overlaps and other relationships between ASD and other NDDs.

4.4. Use of the k-of-N plugin in other ontologies

As discussed in the previous sections, since OWL does not support counting of type *k-of-N* and since DSM diagnostic criteria are hierarchical and involve counting, we developed a new Protégé plugin which implements this capability by enumerating the *k-of-N* combination as OWL class restrictions. This plugin can be used for other OWL ontologies which require this kind of counting, and which may use large *k* and *N* values. Currently, the developed plugin enables *k-of-N* enumeration of up to three levels of hierarchy. However, it can still be used with hierarchical schemas of more than three levels when performed in different executions (of up to three levels each time).

4.5. Limitations

Not all DSM-IV and DSM-5 criteria were implemented in the ontology. After thorough consultation with our clinical expert co-authors, we decided not to implement the following diagnostic criteria:

1. DSM-IV's upper-level (L1) criterion C, namely, "The disturbance is not better accounted for by Rett's Disorder or Childhood Disintegrative Disorder." This criterion relates to relatively rare conditions and is not supported by any explicit diagnostic instrument other than professional clinical judgment.

2. DSM-5's criterion C, namely, "Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities, or may be masked by learned strategies in later life)." This condition is incorporated in the ADI-R questionnaire. Hence, by definition, DSM-5's criterion C holds for subjects to whom the ADI-R was administered, and therefore does not need to be represented in the ontology.

3. DSM-5's criterion D, namely, "Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning." This criterion is implicit in the way we use the ADI-R items in our mapping. Following Huerta et al. [9], we only used item scores which relate to actual impairments displayed by the subject.

4. DSM-5's criterion E, namely, "These disturbances are not better explained by intellectual disability (intellectual developmental disorder) or global developmental delay. Intellectual disability and autism spectrum disorder frequently co-occur; to make comorbid diagnoses of autism spectrum disorder and intellectual disability, social communication should be below that expected for general developmental level." This criterion is not directly supported by any explicit diagnostic instrument other than clinical judgment.

Our ontology is a first step toward the goal of data integration from varied sources, as explained above. Currently, the ontology is incomplete and requires the addition of many concepts and synonyms to support integration of such data sources. This task can be facilitated by distributing the procedure for adding new concepts in the ontology using tools such as Web Protégé (protégé.stanford.edu).

Our ontology was developed by hand, and suffers from issues of scalability common to all manually developed ontologies. Future work could combine our ontology with other ontologies and with automatic ontology extension mechanisms that were developed for autism or that could be applied in this domain. For example, the text-mining approach for discovering implicit knowledge in biomedical literature suggested by Petric et al. [33] could be used to populate the ontology with rare terms from the ASD domain, and the RajoLink literature-mining method [32] could be used to identify relationships between biomedical concepts in separate and disconnected sets of articles. Other NLP approaches could be used to extend existing concepts in the ontology with synonyms [54]. In addition, the semantic-based text-mining approach by Hassanpour et al. [34] could be used to facilitate knowledge acquisition of rule-based definitions of ASD phenotypes from textual sources. These NLP methods might also be guided by the existing structure of our ontology, and could be used to extract detailed information such as the locality of subjects who participated in studies from which conditional probabilities were drawn. It is true that text-mining algorithms are not as effective when driven by ontologies that have complex structures. This was the case in Tu et al. [53], which used an ontology known as ERGO annotation to drive an NLP algorithm that parsed clinical trial eligibility criteria and created OWL class definitions (axioms) from the parsed text. Nevertheless, once the regularity of the input data source is understood, part of this data-entry process can be automated. For example, SWRL-rule creation for ADI-R items could be automated to enable constant update of ADI-R diagnostic instruments.

Another approach for automatically discovering information that could be added to the ontology uses methods that are not ontology-based. For example, Kohane et al. [5], Lyalina et al. [37], and Peacock et al. [6] developed methods for querying health-care databases for symptoms and comorbidities, while Rzhetsky et al. [36] used statistical models from which genetic overlaps between complex phenotypes of NDDs such as autism, bipolar disorder, and schizophrenia could be inferred.

Apart from the limitations related to completeness and scalability, the ontology's correctness for supporting inference based on ADI-R data has been evaluated with a data set that contains few negative examples. In the data set based on DSM-IV criteria, only 9.39% of the subjects did not have autistic disorder but had other forms of autism spectrum disorder. In the data set based on DSM-5 criteria there were no available negative examples, since in the DSM-5, all forms of ASD are collapsed into a single category. Note that data sets with negative examples are rare because usually only individuals with suspected autism complete the ADI-R questionnaire. Though it would have been interesting to compare the performance over the patients of DSM IV to DSM-5, because we do not have negative examples it is impossible to make this comparison as validated by our consulting statistician.

5. CONCLUSION

We have created an ontology which enables automatic inference, via ADI-R data, of DSM-IV autistic disorder and DSM-5 ASD-related phenotypes and diagnostic criteria. As reported in this paper, we have also successfully validated the ontology by showing that it supports accurate inference of autistic disorder and ASD diagnosis based on ADI-R data using real subject data from SFARI. This work offers a number of contributions to research and

practice. First, from the research perspective, we carried out an initial characterization of the DSM sub-criteria defining ASD-related phenotypes met by different subjects with ASD. This analysis could be extended to clustering in order to characterize subtypes of ASD according to its common combined manifestations. Adding subject data relating to risk factors could reveal relationships between risk factors and manifestations, with implications for prognosis and treatment. Moreover, adding knowledge about related NDDs into the ontology would allow it to automatically infer commonalities and differences (in terms of manifestations and risk factors) between ASD and related NDDs, contributing to our understanding of ASD's mechanisms of action. Finally, automatic inference of autistic disorder and ASD phenotypes from ADI-R data could serve a useful public health function by facilitating efforts to track the relationship between specific DSM criteria and treatment protocols, thus helping experts estimate future expected burdens on the healthcare system.

Acknowledgments

References

1. Rossignol DA, Frye RE. A review of research trends in physiological abnormalities in autism spectrum disorders: immune dysregulation, inflammation, oxidative stress, mitochondrial dysfunction and environmental toxicant exposures. Mol Psychiatry Nature Publishing Group. 2011; 17(4):389–401.

2. Huguet G, Ey E, Bourgeron T. The genetic landscapes of autism spectrum disorders. Annu Rev Genomics Hum Genet. 2013 Jan.14:191–213. [PubMed: 23875794]

3. Stessman HA, Bernier R, Eichler EE. A genotype-first approach to defining the subtypes of a complex disease. Cell Elsevier Inc. 2014 Feb; 156(5):872–7.

4. Mathur S, Dinakarpandian D. Finding disease similarity based on implicit semantic similarity. J Biomed Inform Elsevier Inc. 2012 Apr; 45(2):363–71.

5. Kohane, IS.; McMurry, A.; Weber, G.; MacFadden, D.; Rappaport, L.; Kunkel, L., et al. The Co-Morbidity Burden of Children and Young Adults with Autism Spectrum Disorders. In: Smalheiser, NR., editor. PLoS One. Vol. 7. 2012 Apr. p. e33224

6. Peacock G, Amendah D, Ouyang L, Grosse SD. Autism spectrum disorders and health care expenditures: the effects of co-occurring conditions. J Dev Behav Pediatr. 2012 Jan; 33(1):2–8. [PubMed: 22157409]

7. Miller DR, Safford MM, Pogach LM. Who Has Diabetes ? Best Estimates of Diabetes Prevalence in the Department. Health Care (Don Mills). 2004:27.

8. Ciccarese P, Wu E, Wong G, Ocana M, Kinoshita J, Ruttenberg A, et al. The SWAN biomedical discourse ontology. J Biomed Inform. 2008 Oct; 41(5):739–51. [PubMed: 18583197]

9. Huerta M, Bishop SL, Duncan A, Hus V, Lord C. Application of DSM-5 criteria for autism spectrum disorder to three samples of children with DSM-IV diagnoses of pervasive developmental disorders. Am J Psychiatry. 2012 Oct; 169(10):1056–64. [PubMed: 23032385]

10. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington DC: 2000. text revision

11. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5. Arlington, VA: American Psychiatric Publishing; 2013.

12. Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord. 1994 Oct; 24(5):659–85. [PubMed: 7814313]

13. Rutter, M.; Le Couteur, A.; Lord, C. ADI-R: The Autism Diagnostic Interview-Revised. Los Angeles, CA: West Psychol Serv; 2003.

14. Eisenberg L. Images in psychiatry. Am J Psychiatry. 1993; 151:751. [PubMed: 8166319]

15. Lord C, Petkova E, Hus V, Gan W, Lu F, Martin DM, et al. A multisite study of the clinical diagnosis of different autism spectrum disorders. Arch Gen Psychiatry. 2012 Mar; 69(3):306–13. [PubMed: 22065253]

16. American Psychiatric Association. Diagnostic and statistical manual of mental health disorders. 4. Washington DC: 1994.

17. Andrews G, Slade T, Peters L. Classification in psychiatry: ICD-10 versus DSM-IV. Br J Psychiatry. 1999 Jan; 174(1):3–5. [PubMed: 10211143]

18. Wing L, Gould J, Gillberg C. Autism spectrum disorders in the DSM-V: Better or worse than the DSM-IV? Res inm Dev Disabil. 2011; 32:768–73.

19. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, Dilavore PC, et al. The Autism Diagnostic Observation Schedule – Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. 2000; 30(3)

20. Lord, C.; Rutter, MC.; DPSR. Autism Diagnostic Observation Schedule. Los Angeles, CA: West Psychol Serv; 1999.

21. Lord C, Risi S. Frameworks and methods in diagnosing. Ment Retard Dev Disabil Res Rev. 1998; 4:90–6.

22. Tsuchiya KJ, Matsumoto K, Yagi A, Inada N, Kuroda M, Inokuchi E, et al. Reliability and validity of autism diagnostic interview-revised, Japanese version. J Autism Dev Disord. 2013; 43(3):643–62. [PubMed: 22806002]

23. Filipek PA, Accardo PJ, Ashwal S, Baranek GT, Cook EH, Dawson G, et al. Practice parameter: screening and diagnosis of autism: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. Neurology. 2000; 55:468–79. [PubMed: 10953176]

24. Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P, et al. Combining information from multiple sources in the diagnosis of autism spectrum disorders. J Am Acad Child Adolesc Psychiatry The American Academy of Child and Adolescent Psychiatry. 2006 Sep; 45(9):1094–103.

25. Eldridge J, Lane AE, Belkin M, Dennis S. Robust features for the automatic identification of autism spectrum disorder in children. J Neurodev Disord. 2014 Jan.6(1):12. [PubMed: 24936212]

26. Zhou Y, Yu F, Duong T. Multiparametric MRI characterization and prediction in autism spectrum disorder using graph theory and machine learning. PLoS One. 2014 Jan.9(6):e90405. [PubMed: 24922325]

27. Tu SW, Tennakoon L, Connor MO, Das A. Using an Integrated Ontology and Information Model for Querying and Reasoning about Phenotypes: The Case of Autism. AMIA Annu Symp Proc. 2008:727–31. [PubMed: 18999231]

28. Grenon P, Smith B. SNAP and SPAN: Towards Dynamic Spatial Ontology. 2004; 1(March):69–103.

29. Sparrow S, Cicchetti DDA, Balla-Livonia M. Vineland adaptive behavior scales: (Vineland II), survey interview form/caregiver rating form. Pearson Assessments. 2005

30. Horrocks I, Patel-Schneider P. SWRL: A semantic web rule language combining OWL and RuleML. W3C Memb Submiss. 2004 May.

31. McCray AT, Trevvett P, Frost HR. Modeling the Autism Spectrum Disorder Phenotype. Neuroinformatics. 2013 Oct; 12(2):291–305. [PubMed: 24163114]

32. Macedoni-Lukši M, Petri I, Cestnik B, Urban i T. Developing a Deeper Understanding of Autism: Connecting Knowledge through Literature Mining. Autism Res Treat. 2011 Jan. 2011:307152. [PubMed: 22937244]

33. Petric I, Urbancic T, Cestnik B. Discovering Hidden Knowledge from Biomedical Literature. Informatica. 2007; 31:15–20.

34. Hassanpour S, O'Connor MJ, Das AK. A semantic-based method for extracting concept definitions from scientific publications: evaluation in the autism phenotype domain. J Biomed Semantics Journal of Biomedical Semantics. 2013 Jan.4(1):14.

35. Müller HM, Kenny EE, Sternberg PW. Textpresso: an ontology-based information retrieval and extraction system for biological literature. PLoS Biol. 2004 Nov.2(11):e309. [PubMed: 15383839]

36. Rzhetsky A, Wajngurt D, Park N, Zheng T. Probing genetic overlap among complex human phenotypes. Proc Natl Acad Sci U S A. 2007:11694–9. [PubMed: 17609372]

37. Lyalina S, Percha B, Lependu P, Altman RB, Shah NH. Identifying phenotypic signature of neuropsychiatric disorders from electronic medical records. J Am Med Inf Assoc. 2013; 20(e2):e297–305.

38. Chaste P, Leboyer M. Autism risk factors: genes, environment, and gene-environment interactions. Dialogues Clin Neurosci. 2012; 14(3):281–92. [PubMed: 23226953]

39. Gardener H, Spiegelman D, Buka SL. Perinatal and neonatal risk factors for autism: a comprehensive meta-analysis. Pediatrics. 2011; 128(2):344–55. [PubMed: 21746727]

40. Gardener H, Spiegelman D, Buka SL. Prenatal risk factors for autism: comprehensive meta-analysis. Br J Psychiatry. 2009; 195(1):7–14. [PubMed: 19567888]

41. Grabrucker AM. Environmental factors in autism. Front Psychiatry. 2012; 3:118. [PubMed: 23346059]

42. Landrigan PJ. What causes autism? Exploring the environmental contribution. Curr Opin Pediatr. 2010; 22(2):219–25. [PubMed: 20087185]

43. Newschaffer CJ, Croen LA, Daniels J, Giarelli E, Grether JK, Levy SE, et al. The epidemiology of autism spectrum disorders. Annu Rev Public Heal. 2007; 28:235–58.

44. Waterhouse, L. Rethinking Autism: Variation and Complexity. Acad Press; 2013.

45. Grosof BN, Horrocks I, Volz R, Decker S. Description Logic Programs: Combining Logic Programs with Description Logic Categories and Subject Descriptors. WWW2003. 2003 May.:20–4.

46. Horrocks I, Patel-Schneider PF, Van Harmelen F. From SHIQ and RDF to OWL: the making of Web Ontology Language. J Web Semant Sci Serv Agents World Wide Web. 2003; 1:7–26.

47. Baader, F.; Calvanese, D.; McGuinness, DL.; Nardi, D.; Patel-Schneider, PF. The description logic handbook: theory, implementation, and applications. Cambridge University Press; New York NY, USA: 2003.

48. Scheuermann, RH.; Ceusters, W.; Smith, B. Toward an Ontological Treatment of Disease and Diagnosis. Proceedings of the 2009 AMIA Summit on Translational Bioinformatics; 2009; p. 116-20.

49. Ausderay KK, Furlong M, Sideris J, Bulluck J, Little LM, Watson LR, et al. Sensory subtypes in children with autism spectrum disorder: latent profile transition analysis using a national survey of sensory features. J Child Psychol Psychiatry. 2014; 55(8):935–44. [PubMed: 25039572]

50. Bowers K, Wink LK, Pottenger A, McDougle CJ, Erickson C. Phenotypic differences in individuals with autism spectrum disorder born preterm and at term gestation. Autism. 2014

51. Bernier R, Golzio C, Xiong B, Stessman HA, Coe BP, Penn O, et al. Disruptive CHD8 mutations define a subtype of autism early in development. Cell. 2014; 158(2):263–76. [PubMed: 24998929]

52. Mazefsky CA, McPartland JC, Gastgeb HZ, Minshew NJ. Comparability of DSM-IV and DSM-5 ASD Research Samples. J Autism Dev Disord. 2013; 43(5):1236–42. [PubMed: 23011251]

53. Tu S, Peleg M, Carini S, Bobak M, Ross J, Rubin D, et al. A Practical Method for Transforming Free-Text Eligibility Criteria into Computable Criteria. J Biomed Inform. 2011; 44(2):239–50. [PubMed: 20851207]

54. Percha B, Altman RB. Inferring the semantic relationships of words within an ontology using random indexing: applications to pharmacogenomics. AMIA Annu Symp Proc. 2013 Jan. 2013:1123–32. [PubMed: 24551397]

Highlights

We augmented an autism ontology with SWRL rules to infer phenotypes from ADI-R items

We represented DSM diagnostic criteria for Autism Spectrum Disorder in OWL

We developed a custom Protégé plugin for enumerating combinatorial OWL axioms

OWL Reasoner thus infers autism-related phenotypes from ADI-R questionnaire results

We evaluated the classification results with data from Simons Foundation
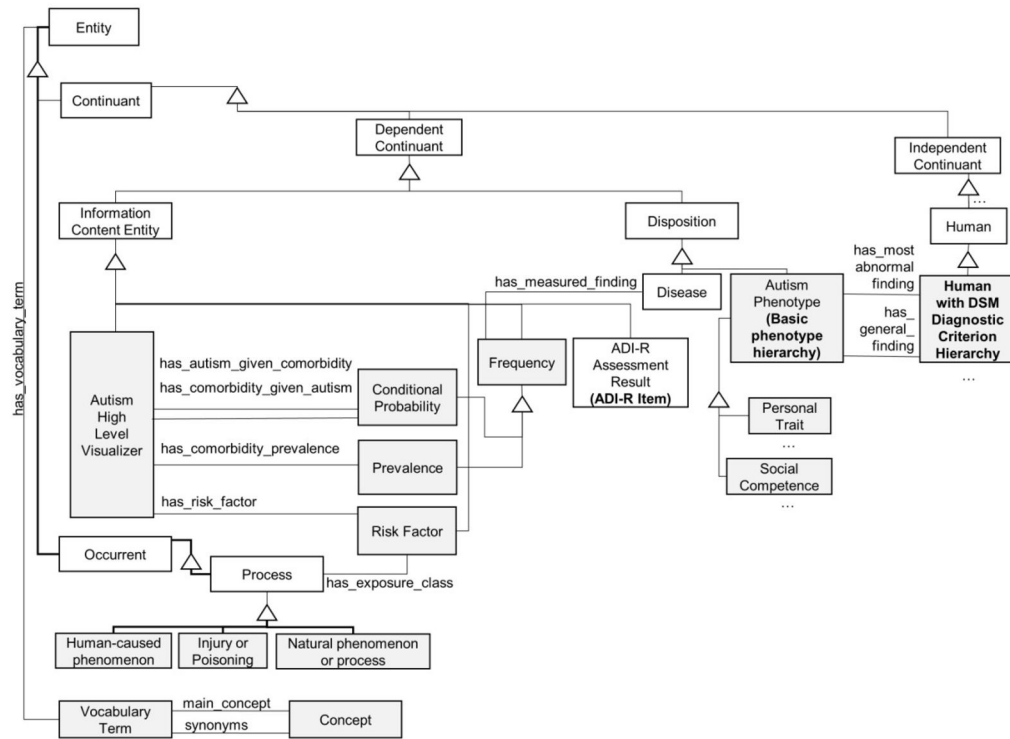
Figure 1.
Top-level class diagram of the autism ontology, showing key classes and relationships. Gray squares represent the classes that extend the ontology by Tu et al. White classes are taken from the BFO.
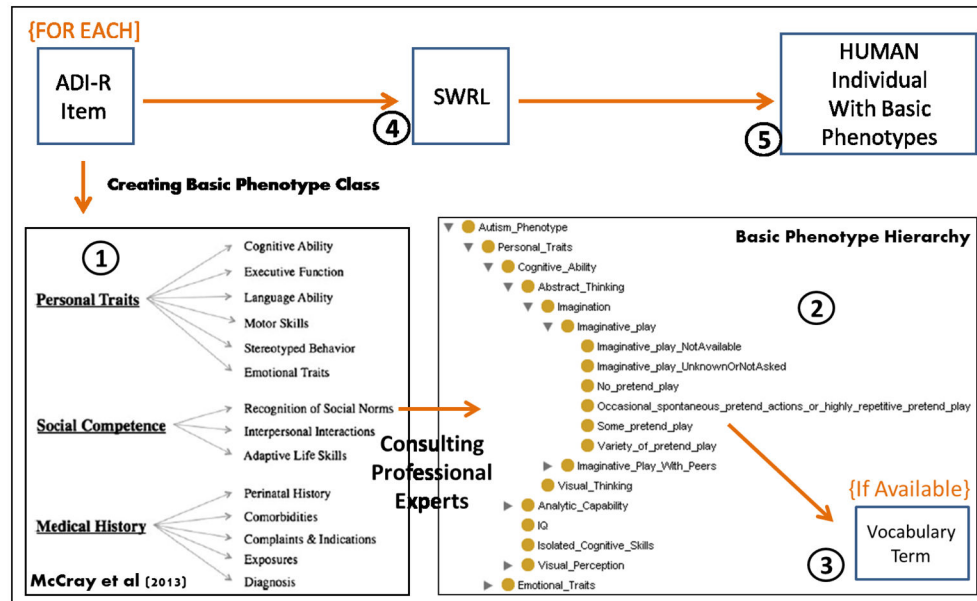
Figure 2.
Ontology population process overview – Basic phenotypes representation. (1) The top-level classes of the basic phenotype hierarchy were taken from McCray et al. (2) The Personal_Traits class from (1) was integrated as a child of the Autism_Phenotype (ASD_Related_Phenotype) class, which is a child of the BFO disposition class. The ADI-R items and their range of values (e.g., ImaginativePlay_NotAvailable) were integrated as children of the concepts in McCray's hierarchy. (3) Vocabulary terms (where available) were added to the concepts in the hierarchy as annotations. (4) SWRL rules were then used to (5) associate with a human subject a basic phenotype from the hierarchy corresponding to an ADI-R item in this human's ADI-R data.

Figure 3.
DSM-IV and DSM-5 class hierarchies

Figure 4. Infering the "Head_Shaking_Never" basic phenotype of a Human from ADI-R data
(1) An individual of the ADI-R assessment result belonging to a patient whose subjectKey is 11000. The item functional communication head shaking (funcon_chshake) has a value of 2. (2) A SWRL rule infers the "Head_Shaking_Never" phenotype for subjects who scored 2 for item 44 in the ADI-R. (3) A specific individual of the Human class (in this example, the individual with ID 11000) with his set of inferred phenotypes, including the one inferred by this SWRL rule.

Figure 5.
Ontology population process overview – DSM diagnostic criteria representation as OWL class hierarchy. (1) To define a DSM criterion in OWL, we obtain from Huerta's mapping a list of ADI-R items (see second row in the table shown in the figure). (2) The basic phenotypes corresponding to the ADI-R items are logically combined into an OWL class expression (see Figure 6). (3) For higher-level (L2, L1) criteria, the k-of-N Protégé plugin is used to create class expressions. (4) The resulting L1, L2, and L3 classes are arranged in a hierarchy. Note that the second part of the DSM criterion in Figure 6 (gesture or mime) was represented using additional ADI-R items related to gesture or mime as provided by the professional experts with whom we consulted.

------------------------------- *Necessary & Sufficient* -------------------------

has_current_finding has *Fewer_Than_Five_Words_Total_or_Speech_Not_Used_on_a_Daily_Basis*
and
has_mostAbnormal4-5_finding has *Never_Nods_Spontaneously*
and
has_mostAbnormal4-5_finding has *Head_Shaking_Never*
and
has_mostAbnormal4-5_finding has *Signs_Gesture_No_Use*
and
has_mostAbnormal4-5_finding has *Gaze_Unusual_or_Odd_Gaze*
------------------------------- *Necessary* -------------------------------

Human_with_DSM-IV_Diagnosis_definition_A2(a)

Figure 6. Combining basic subject phenotypes with logical operators

This example shows the OWL class definition corresponding to DSM-IV's diagnostic criterion A2(a): "delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as **gesture or mime**)". This is a union of five basic phenotypes related to the "most abnormal 4–5" (the most severe phenotype the subject exhibited at age 4–5) or the "current finding" (the phenotype that is currently exhibited). The phenotypes described here are related to the following ADI-R items: (1) overall level of language; (2) nodding; (3) head shaking; (4) conventional or instrumental gestures; (5) direct gaze.
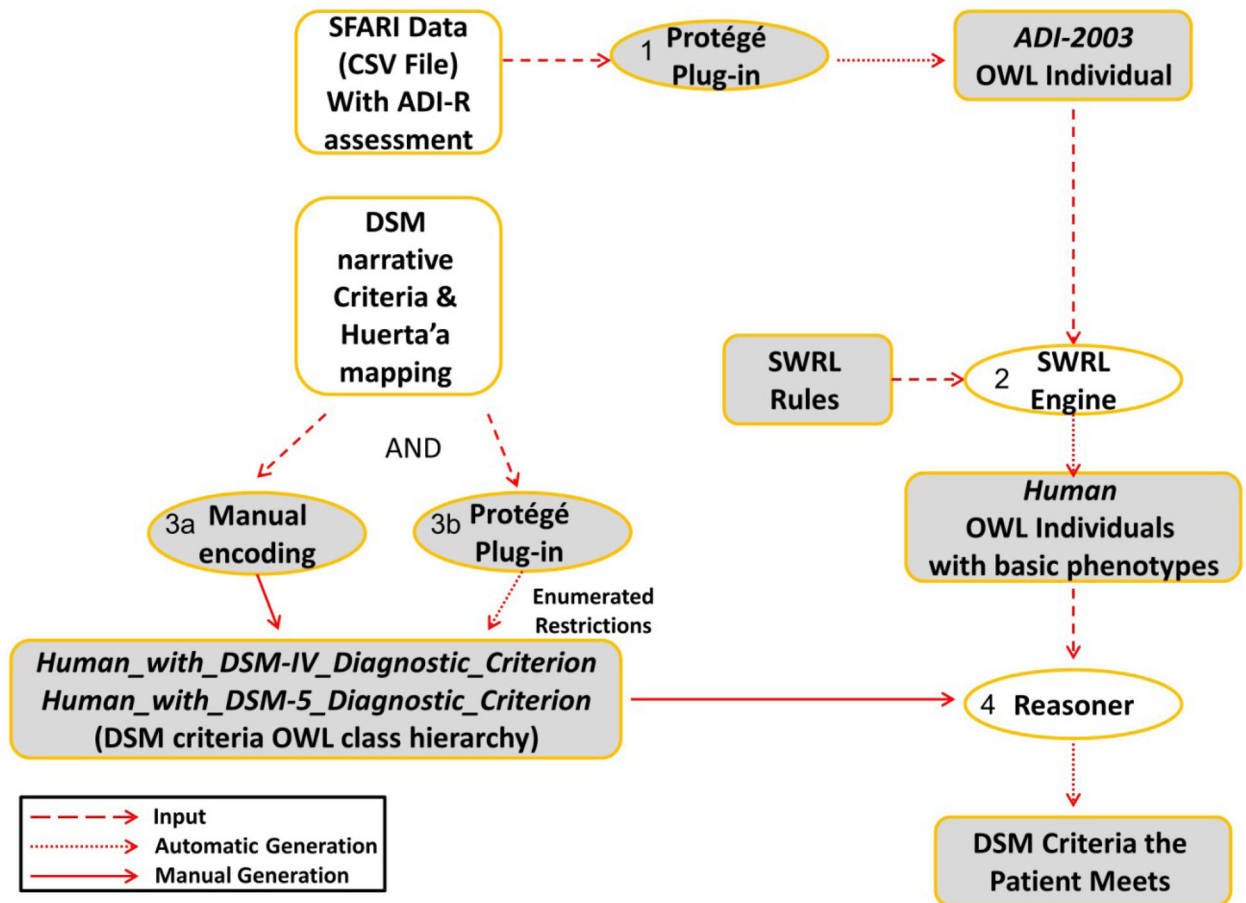
Figure 7.
An overview of the inference of ASD-related phenotypes from SFARI data. Shapes in white show sources and software that were available to us; shapes in gray show our own development. (1) A Protégé plugin was used to generate ADI-R OWL individuals corresponding to ADI-R questionnaire results of patients from the SFARI data set. (2) Each ADI-R result item was translated via a SWRL rule which was executed by the SWRL engine to populate for each OWL Human individual a set of basic phenotypes corresponding to the ADI-R items for that patient. (3) Based on DSM criteria, OWL classes of Human_with_DSM_Diagnostic_Criterion were defined. Combinatorial class expressions were created automatically via a Protégé plugin for enumeration of combinatorial k-of-N expressions. (4) A reasoner was used to infer for each Human patient which DSM diagnostic criteria he meets based on his SWRL-inferred basic phenotypes.
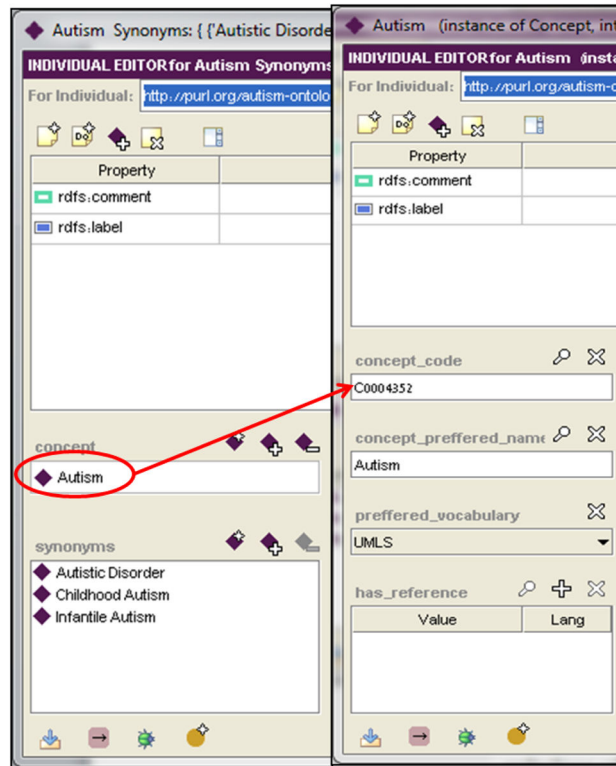
Figure 8.
An Individual representing the concept Autism along with its synonyms Autistic Disorder, Childhood Autism and Infantile Autism. All concepts are instances of the Concept class. The synonyms in this figure are type-of Autism but are still considered as synonyms of the same concept.
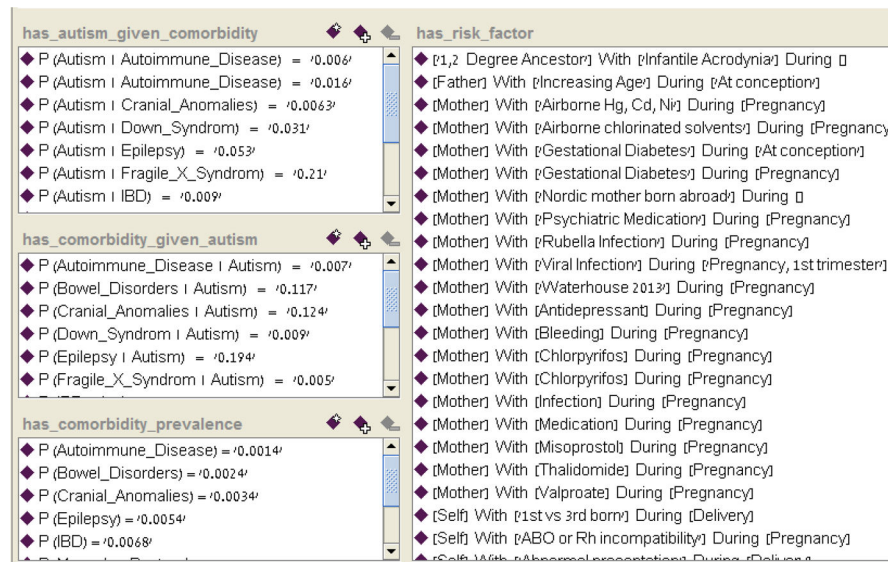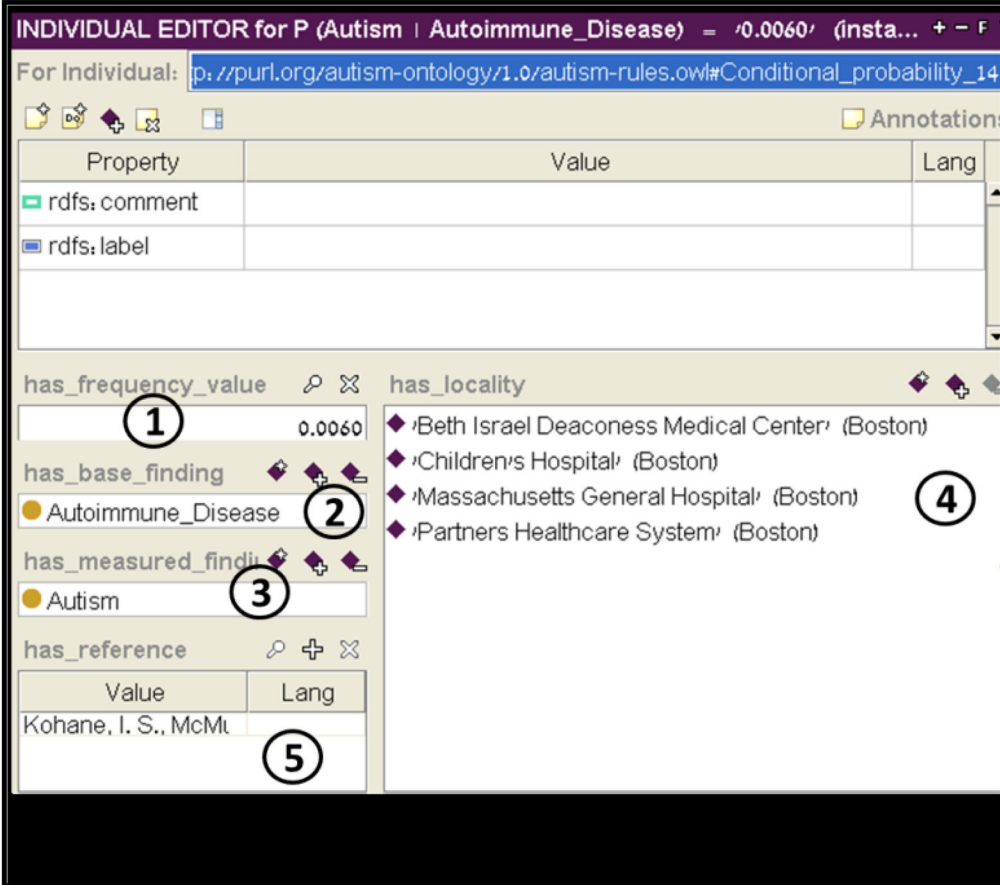
Figure 9.
The Autism_High_Level_Visualizer class enables a high-level visualization of autism risk factors and comorbidities knowledge.

Figure 10.
Conditional Probability individual. The probability (1) that a subject will be diagnosed with autism (3) given that he was diagnosed with autoimmune disease (2) is 0.006. These data were gathered from healthcare systems in the Boston area (4) as reported by Kohane et al. (5). Possible types of healthcare systems are: hospital_outpatient, hospital_inpatient, community_clinic, private_clinic.

Figure 11.
An individual of the Risk Factor class. (1) Gestational diabetes is an environmental risk factor for autism, occurring (2) during pregnancy to (3) the mother of a child who develops ASD. The exposure is of class (4) obstetric complications and cited by Gardener (6).
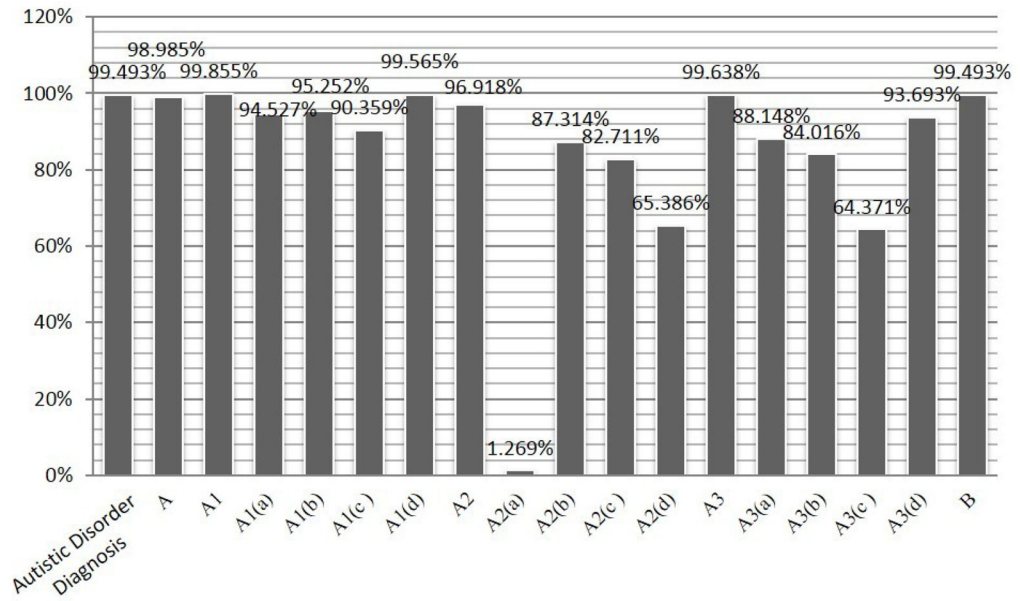
Figure 12.
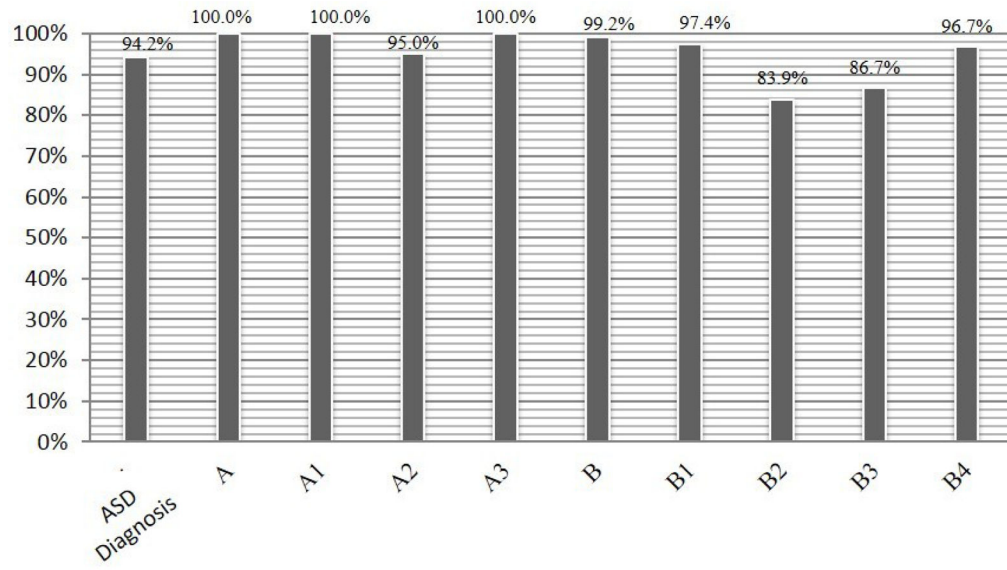Percentage of subject records that fit the represented DSM-IV criteria

Figure 13.
Percentage of subject records that fit the represented DSM-5 criteria

Table 1

Inference results for DSM-IV criteria

	Inferred as having autistic disorder	Inferred as not having autistic disorder
Had diagnosis of autistic disorder (2394)	2394 (true positive)	0 (false negative)
Had diagnosis of ASD (196)	183 (false positive)	13 (true negative)
Had diagnosis of Asperger's (52)	49 (false positive)	3 (true negative)

Table 2

Inference results for DSM-5 criteria

	Inferred as having ASD	Inferred as not having ASD
Had diagnosis of ASD (2642)	2485 (true positive)	157 (false negative)