NeuMORE: Ontology in Stroke Recovery

Christopher Townsend*, Jingshan Huang*, Dejing Dou†, Haishan Liu†,
Lei He‡, Patrick Hayes§, Robert Rudnick*, Hardik Shah*, Dennis Fell¶, and Wei Liu¶

*School of Computer and Information Sciences, University of South Alabama

†Computer and Information Science Department, University of Oregon

‡Department of Information Technology, Armstrong Atlantic State University

§Florida Institute for Human and Machine Cognition

¶Department of Physical Therapy, University of South Alabama

***Abstract*—Hemiparesis is the most common impairment after stroke, and the initial severity of hemiparesis had been the strongest predictor of neuromotor functional recovery level. However, the intervention response of stroke survivors does not always correlate with their initial level of impairment, which implies the existence of other factors that may significantly affect stroke survivors' recovery process. It is critical to consider these factors in a principled, comprehensive way so that physical rehabilitation (PR) researchers may predict which stroke survivors will respond best to therapy and, as a result, to determine if a particular type of therapy is a more optimal match. Currently, such prediction is primarily a manual process and remains a challenging task to PR researchers and clinicians. Based upon a domain-specific ontology, NeuMORE, we propose a computing framework that aims to facilitate knowledge acquisition from existing sources via semantics-enhanced data mining (SEDM) techniques. It will assist PR researchers and clinicians in better predicting stroke survivors' neuromotor functional recovery level, and will help physical therapists customize most effective intervention therapy plans for individual stroke survivors.**

***Index Terms*—stroke; hemiparesis; extremity dysfunction; neuromotor functional recovery; intervention therapy; ontology; formal semantics; data mining**

I. Introduction

Stroke is the leading cause of adult disability in the U.S., and hemiparesis is the most common impairment after stroke. While the initial severity of hemiparesis had been demonstrated to be the strongest predictor of functional recovery level, recent research has discovered a range of responders who do not always correlate with their initial level of impairment. This implies the existence of factors that play significant roles in the recovery process. To better design targeting intervention therapy strategies, these factors need to be considered comprehensively. It is then possible to predict which stroke survivors will respond best to therapy and, subsequently, to determine if a particular type of therapy is a more optimal match. Currently, such prediction is primarily a manual process, which is not only time-consuming and error-prone, but also subject to humans' limited prior knowledge. In addition, the lack of agreed-upon, formal semantics among distributed physical rehabilitation (PR) groups has witnessed inconsistent terminologies, different schemas, and incompatible data formats, which altogether further aggravate the situation. Therefore, it is challenging for PR researchers and clinicians to obtain information from existing data sources, let alone to share and reuse knowledge acquired.

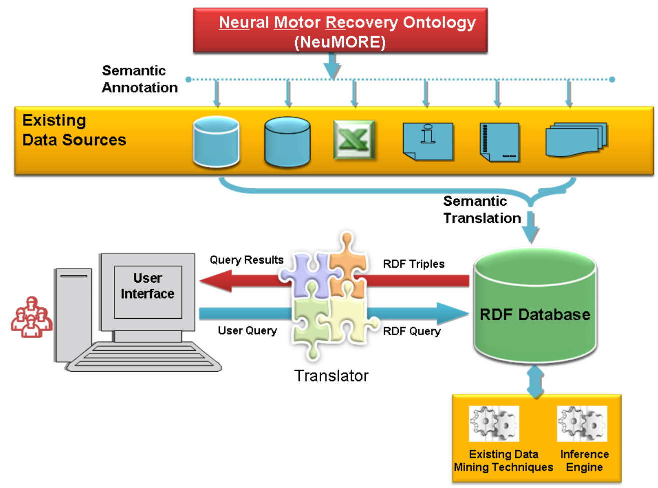

Fig. 1. NeuMORE System Framework

We propose an innovative computing framework to facilitate automated knowledge acquisition and to help PR researchers and clinicians predict stroke survivors' improvements in extremity motor function. Our framework is built upon a domain-specific ontology, the NeuroMotor Recovery (NeuMORE) Ontology, which is *the very first formal knowledge model* in the domain of stroke recovery. With the formal semantics defined in the NeuMORE, and the logic-based inference (a.k.a. ontological reasoning) that follows, it is possible to enhance existing data mining techniques. Resultant techniques from this unique combination are referred to as *semantics-enhanced data mining (SEDM) techniques*, which facilitate effective knowledge discovery, sharing, and reuse from existing sources. Ultimately, our research objective is **to assist PR researchers and clinicians in identifying factors that determine stroke survivors' neuromotor functional recovery process after stroke; thus, more optimal intervention therapy plans can be customized for individual stroke survivors**.

II. Methodologies

A. Phase One: Domain-Specific Ontology

During the iterative, top-down, and knowledge-driven approach to develop the NeuMORE ontology, we have observed the seven practices proposed by the OBO Foundry Initiative [1]. Both ontology engineers and domain experts (PR researchers and clinicians) have been involved, working together to capture domain knowledge, develop a conceptualization, and implement the conceptual model. In addition, we have also reused and extended a subset of concepts from the Basic Formal Ontology (BFO) [2].

B. Phase Two: Annotation of Existing Data Sources

Semantic annotation is the process of tagging source files with predefined metadata, which usually consists of a set of ontological concepts. We adopt a "deep" annotation on both data schemas and data sets. The outcomes are: (1) a set of mapping rules, specified in W3C Rule Interchange Format (RIF) [6], between NeuMORE concepts and elements from existing data schemas; and (2) annotated data sets published in the RDF format.

C. Phase Three: RDF Database in Semantic Translation

Instead of relational databases, a RDF database will be adopted in the semantic translation phase. By making use of a RDF repository to store RDF triples, our system makes it possible to query, manipulate, and reason about existing data through available tools, including, but not limited to, SPARQL Query Language for RDF [7] and OWL reasoners (e.g., Pellet [5] and FaCT++ [3]). In this manner, the logic-based knowledge acquisition can be realized, and existing data mining techniques will be enhanced and have a better performance. With the correspondence between data sources and the global schema (the NeuMORE ontology) established in *Phase Two*, user queries will be translated into RDF-based queries.

D. Phase Four: Complex Query/Search in a Unified Style

We propose to integrate OWL inference engines into traditional data mining techniques, resulted in *semantics-enhanced data mining (SEDM) techniques*, when interacting with the RDF database. Such a unique combination will enable us to take advantage of emerging Semantic Web technologies. In addition, we aim to provide users with a single, *unified* interface that takes their requests in a nonprocedural specification format. While it is possible for existing data sources to have heterogeneous semantics, the NeuMORE system presents users a *uniform* view of information, along with newly obtained (deduced) knowledge via data mining enhanced by the automated inference, whenever possible.

III. The NeuMORE Ontology

A. Ontology Schema

An initial version of the NeuMORE ontology was designed using Protégé 4.0, having 351 concepts in total, along with 29 object properties and 16 data properties. Fig. 2 is a screen shot from the Protégé GUI, demonstrating NeuMORE top-level concepts, which are also listed in Table I in alphabetical order. Also listed in the table are some representative second-level concepts, i.e., direct children of top-level concepts.

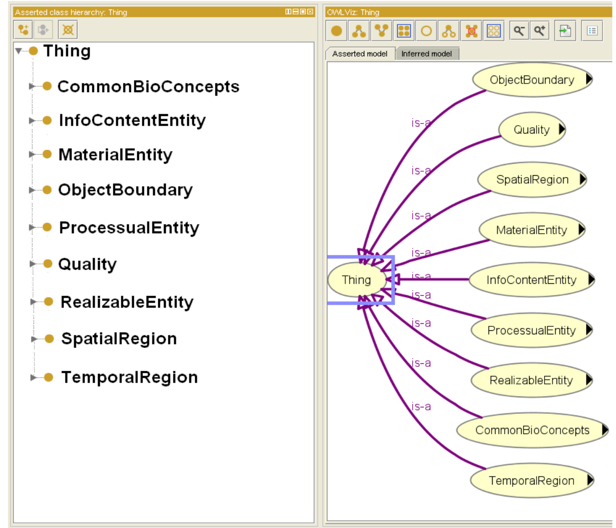

Fig. 2. Top-Level NeuMORE Concepts

TABLE I
NeuMORE Top-Level Concepts

Top-Level Concept	Representative Direct Children
CommonBioConcepts	SignsOrSymptoms, Tissue, Treatment
InfoContentEntity	DataSet, Measurement, PlanSpecification
MaterialEntity	FiatObjectPart, Object, ObjectAggregation
ObjectBoundary	ScalpSurface
ProcessualEntity	Process, ProcessAggregation, ProcessBoundary
Quality	ClinicalCharacteristics, MovementMeasure
RealizableEntity	Disposition, Function, Plan
SpatialRegion	AnatomyPoint, AnatomyAxis, AnatomySurface
TemporalRegion	TemporalInstant, TemporalInterval

B. Class Architecture

The NeuMORE back end is implemented in the C# language, following an object-oriented approach. Relationships such as *superClassOf*, *subClassOf*, *superPropertyOf*, and *subPropertyOf* are placed within an inheritance structure under the common base class *OWLRelation*, which is a generic relationship between two types that specifies both the domain and the range of that relationship. The inheritance structure, together with common subclasses such as *OWLClassRelation* and *OWLPropertyRelation* greatly simplifies the implementation of increasingly complex operations to be performed on the ontology. The NeuMORE ontology has been submitted to the National Center for Biomedical Ontology (NCBO) [4] BioPortal (http://bioportal.bioontology.org/ontologies/44245), and a Web portal has been deployed at http://neumore.cis.usouthal.edu, which contains the up-to-date progress of the project.

References

[1] B. Smith et al. The OBO foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, 25(11):1251–1255, 2007.
[2] Basic Formal Ontology. *http://www.ifomis.org/bfo/*, October 2010.
[3] FaCT++ Reasoner. *http://owl.man.ac.uk/factplusplus/*, October 2010.
[4] NCBO. *http://www.bioontology.org/*, October 2010.
[5] Pellet Reasoner. *http://clarkparsia.com/pellet/*, October 2010.
[6] RIF. *http://www.w3.org/TR/2010/REC-rif-prd-20100622/*, October 2010.
[7] SPARQL. *http://www.w3.org/TR/rdf-sparql-query/*, October 2010.