Semi-Automatic Population of Ontology of Philippine Medicinal Plants from On-line Text

Nathalie Rose Lim-Cheng, Junn Richmond C. Co, Christa Hannah S. Gaudiel, Darah F. Umadac, and Nadine L. Victor

De La Salle University-Manila
* Nathalie Rose Lim-Cheng: nats.lim@delasalle.ph

Abstract: Ontologies are formal representations of knowledge organized as concepts of a domain with their relationships defined. These are used in various fields including biomedical informatics. However, to create a functional ontology, it must be populated with instances of the defined concepts. Since it is a tedious task with constantly updating data, especially in the field of healthcare, information extraction (IE) can be employed to fulfill this task semi-automatically. In this paper, semi-automatic ontology population through information extraction from online articles will be discussed. The design of the ontology is based on available information that can be consistently extracted from the available on-line text. These information focus on the medicinal properties of the plant (i.e., what illness can it be applied to, which body part does it affect/cure, the preparation instruction, and the plant part to be used). Lastly, we present some results from initial usability assessment and from comparison with Gold standard. The test shows we can get an 85.71% accuracy.

Key Words: Ontology population, medicinal plants, information extraction