

# GPT Onto: A New Beginning for Malaysia Gross Pollutant Trap Ontology

Chandrika M.J., Lariyah M.S., and Alicia Y.C. Tang

**Abstract**—Ontology is widely being used as a tool for organizing information, creating the relation between the subjects within the defined knowledge domain area. Various fields such as Civil, Biology, and Management have successful integrated ontology in decision support systems for managing domain knowledge and to assist their decision makers. Gross pollutant traps (GPT) are devices used in trapping and preventing large items or hazardous particles in polluting and entering our waterways. However choosing and determining GPT is a challenge in Malaysia as there are inadequate GPT data repositories being captured and shared. Hence ontology is needed to capture, organize and represent this knowledge into meaningful information which can be contributed to the efficiency of GPT selection in Malaysia urbanization. A GPT Ontology framework is therefore built as the first step to capture GPT knowledge which will then be integrated into the decision support system. This paper will provide several examples of the GPT ontology, and explain how it is constructed by using the Protégé tool.

**Keywords**—Gross pollutant Trap, Ontology, Protégé.

## I. INTRODUCTION

ONTOLOGY has successful been a tool in capturing and representing various fields such as biology, engineering medicine and banking. In this paper, we would like to present GPT Onto where we describe approaches taken in acquiring and representing the GPT domain. The constructed ontology and expert system will be able to assist the decision makers to make a fast decision in selecting the best GPT suited to their chosen area. The objectives of this research are to define GPT knowledge base repository by developing a GPT domain knowledge using ontology and to develop an expert system prototype representing GPT. Collectively, these works will contribute to the Stormwater Ontology construction and maintenance in Malaysia.

Chandrika Mohd Jayothisa is a lecturer in Multimedia University, Malaysia currently a PhD candidate at the College of Engineering, Uniten, Jalan Ikram-Uniten, 43000 Kajang, Selangor, Malaysia (phone: +06-2523221; e-mail: chandrika.mohdjayothisa@mmu.edu.my).

Lariyah Mohd Sidek is a Associate Professor in the College of Engineering, Uniten, Jalan Ikram-Uniten, 43000 Kajang, Selangor, Malaysia (phone : +603 - 89287289; e-mail: Lariyah@uniten.edu.my).

Alicia Y.C. Tang is a Senior Lecturer in the College of Information Technology, Uniten, Jalan Ikram-Uniten, 43000 Kajang, Selangor, Malaysia (phone : +603- 89212336; e-mail: aliciat@uniten.edu.my).

## II. GPT: AN OVERVIEW

Gross pollutant traps (GPT) are devices which are used in trapping and preventing large items or hazardous particles in polluting and entering our waterways. GPT can mainly be found in stormwater drains, industrial plants, urban wetlands, wet markets, beach fronts and other locations. Basically, GPT helps to remove gross pollutants and coarse sediments before it enters the receiving waters.

GPT can be divided into two different categories: a dry load and a wet load. Items that are collected and stored above water levels are known as dry load. Whereas, a wet load are items that are collected and stored below standing water levels and are classified as toxic liquids.

There are basically five different techniques which can be designed using one or any combination techniques. These five techniques are: Screening, Flow separation, stopping the flow water, Sedimentation and Flootation. The number of the GPT should be specifically based on the selected location it is intended to be used. Selecting the best GPT tool helps to weigh in factors such as requirement of water quality, economic, flood management and existing ecosystem environment.

There are few factors an engineer a consultant needs to consider when designing and selecting the suitable GPT. These factors are based on the size of particles to be caught in that location, the frequency of storms or other major water influxes, estimated loading in an area, installation and operating cost, maintenance requirement, safety and many more. There are several type of GPT installed in Malaysia. Among them are Trash Rack, Sediment Trap, Release Net, CleansAll, CDS, SBTR, and Drain Inlet Insert. However, limited information is available on the performance of structural devices in trapping gross pollutants.

## III. CASE STUDY FROM RELATED WORK

### A. Management Syllabus System

Management Syllabus system was created for the Centre for Diploma Programme, Multimedia University. Syllabus Management System is developed based on the combination of both Malaysian Qualification Assurance (MQA) syllabus guidelines with CDP Outcome Based Education format, which focuses on the end-product, and defines what the learner is able to do for a particular subject within a programme [1]. The system was created from scratch since there are no existence ontology shared, build or made available in the internet. By

using Protégé, the syllabus domain has successfully developed. Fig. 1 illustrates the Syllabus Management System taxonomy classes using the Protégé 2000.

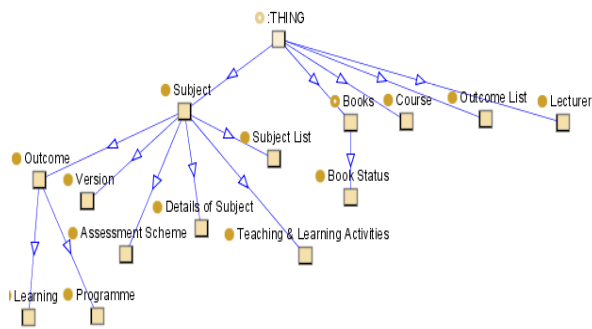


Fig. 1 Taxonomy Classes for the Syllabus Management System

**B. DeGPT System**

Decision Support System for Gross Pollutant Trap (DeGPT) System was developed to assist engineers and authorities to select, design and evaluate both the proprietary and non-proprietary GPT in a particular urban area.

DeGPT, was developed by the Centre for Stormwater and Geohazard Management (CSGM) in University Tenaga Nasional (UNITEN). This standalone system is perceived to be a catalyst for the Department of Irrigation and Drainage (DID) in acquiring, storing and retrieving relevant and resourceful knowledge of GPTs in rehabilitating river system in Malaysia [2]. The system assist the users in determine the design of GPT by inserting relevant data from the users and also data captured from the MSMA manual book.



Fig. 2 Main page of DeGPT System

The system contains GPT selection aid, Design GPT calculation, GP Characteristics, GPT Knowledge bank, Life Cycle Cost calculation and Water Quality Analysis. Fig. 2, illustrate main menu from the DeGPT application and Fig. 3 display the Design GPT application page. Design GPT calculation consist of steps and procedures to capture data and calculate the coefficient, Intensity, rainfall and discharge flow rate before proceeding to the next calculation for the GPT: CDS and SBTR.

However, the data repositories which exist in DeGPT contribute to a driven application system which limits to the

ability to represent the knowledge captured into meaningful information. Hence, ontology is needed to develop and represent the Malaysia GPT Domain ontology.

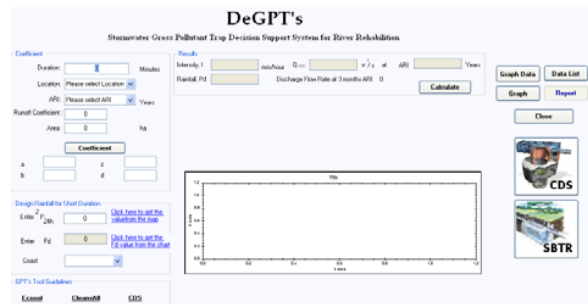


Fig. 3 Design GPT calculation

**C. WaWO Ontology**

Waste Water ontology (WaWO), created to represent and describe the wastewater treatment management with detail studies of the microbiological component [3]. The WaWO has managed to reuse some related parts of the UpperCyc Ontology. The WaWO will then be integrated with the environmental decision support system. Fig. 4 shows the classes in WaWO.

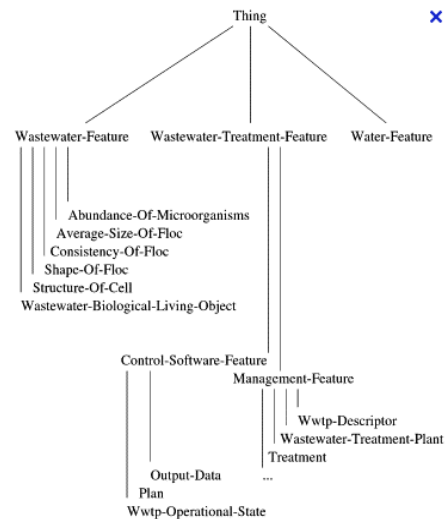


Fig. 4 WaWO Ontology Classes

**IV. GPT ONTOLOGY METHODOLOGY**

**A. Ontology**

Ontologies which signify as a content theory [4] have been successfully being used in various fields in capturing, describing and representing the knowledge domain. Its importance is being recognized in a diverse set of research fields from knowledge engineering, artificial intelligence, knowledge representation, qualitative modeling, language engineering, database design, information modeling, information integration, object- oriented analysis, information

retrieval and extraction, knowledge management and organizations to multimedia and agent-based design [5].

In definition, ontology uses vocabulary which describes a systematic of existence in a particular hierarchy and the relations created between objects which often specialize to some subject matter. The content theories define real world objects of interest (concepts, entities, events, actions or processes), the properties of objects and the relations between them in a certain domain [4]. Each vocabulary used should be closely conceptualized to represent the domain or subject. Therefore, a good ontological analysis helps to clarify the structure of knowledge and defining a good vocabulary representing that knowledge domain.

In order to build a knowledge representation language based on the analysis, one need to associate terms with the concepts and relations in the ontology and devise syntax for encoding knowledge in terms of the concepts [6].

For the case of gross pollutant trap, an ontology of an GPT devices domain, might include vocabulary that describe the design of the gross pollutant devices conceptual elements such as the trap depth, catchment area, water quality, coefficient and the relations between these elements. The catchment area and the coefficient both can be used as a part in predicting the trap efficiency in the SBTR design, where as it can also be used in determine the CDS design.

#### B. Capturing Domain Knowledge of GPT

Ontology domain creation can either be taken from any similar object created, or can also be created from scratch or it can also modified from the existing domain to suit with the current representation on the current knowledge. Yet, there is no one definite right method in developing a knowledge domain and no correct methodologies to be used in creating, analyzing, representing and capturing the knowledge.

There are several tools and method used in developing a new knowledge based domain ontologies which can either be done in a manual method, partonomies, taxonomies or any other organizations levels of ontologies. Basically, there are two methodologies – Toronto Virtual Enterprise (TOVE) and Methontology which will be used in developing the GPT-ontology. The TOVE is a stage-based models and it is appropriate to be used when purposes and requirements of the ontology are clear. Methontology is more involving prototype model and it is more suitable and useful when the environment is dynamic and difficult to understand. There is no one correct way to model a domain, there are always viable alternatives [7]. Mainly, the best solution depends on the application that the developer has in mind, and the tools that she uses to develop the ontology [8]. Fig. 5, illustrate the GPT Onto framework which will belong to the expert system whereas Fig. 6, illustrate the data elements in the GPT Ontology. The GPT Onto will be in the form of RDF and OWL which is then be integrated with SPARQL to express queries and obtain focused results.

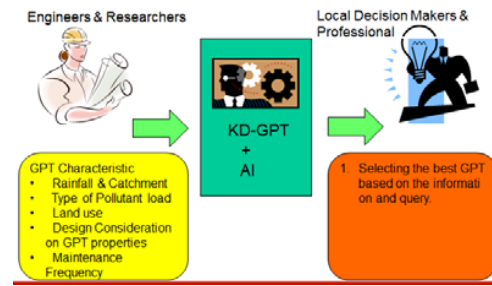


Fig. 5 GPT Ontology Framework

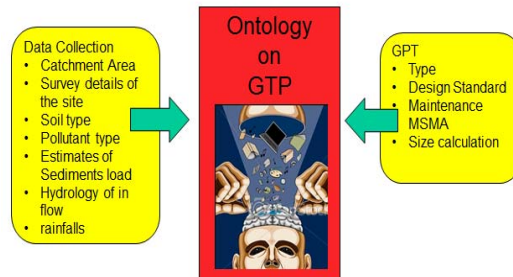


Fig. 6 Data elements in GPT Ontology

#### V. GPT- ONTO

Basically, the design and development of GPT Onto consists of the following research activities:

1. Asses site suitability and catchment analysis
  - a. Site locations, conditions and type of land capability
    - i. Type of area: residential, industrial, market places, parks, commercial/business areas, roads/highways etc
  - b. Pollutant characteristics:
    - i. Helps in determining the type and sizing of GPT
    - ii. Source and type of gross pollutants
    - iii. GPT potential installation location
2. Determine design objective and targets
  - a. Treatment objectives
  - b. GPT Design flows
  - c. Flood capacity
  - d. Trapped pollutant storage
  - e. Maintenance requirements
3. Perform calculation
  - a. Identify the location, get the coefficient, etc.
  - b. Calculate based on MSMA guidelines
  - c. Output: intensity, Qari, Rainfall, etc.
4. Select the type of GPT
5. Determine design flow results and suggestions

GPT criteria, guidelines, design consideration, data and information are collected from Manual Saliran Mesra Alam (MSMA), CSGM, research results, consultants and GPT

companies. MSMA has successful given some brief guidelines in the types and categories of GPT and some general design consideration for the engineers to use. The DeGPT design calculation will be used as guidelines to identify the objects and classes to be created and represented. Protégé 4.2 is used to develop the GPT Ontology. The TOVE methodology will be used for designing GPT Onto. Part of the ontology will be extracted from the DeGPT. Fig. 7 shows the two main classes of GPT domain while Fig. 8 illustrates the different types of GPT.

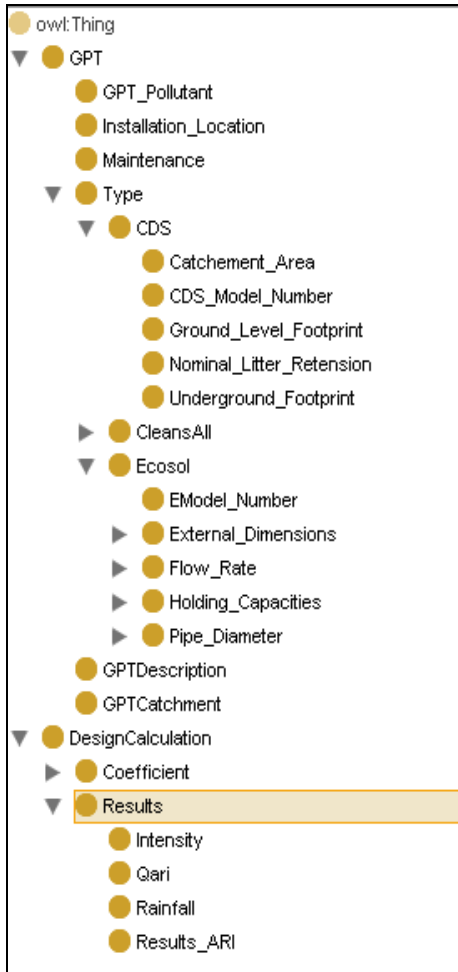


Fig. 7 Classes of GPT Onto

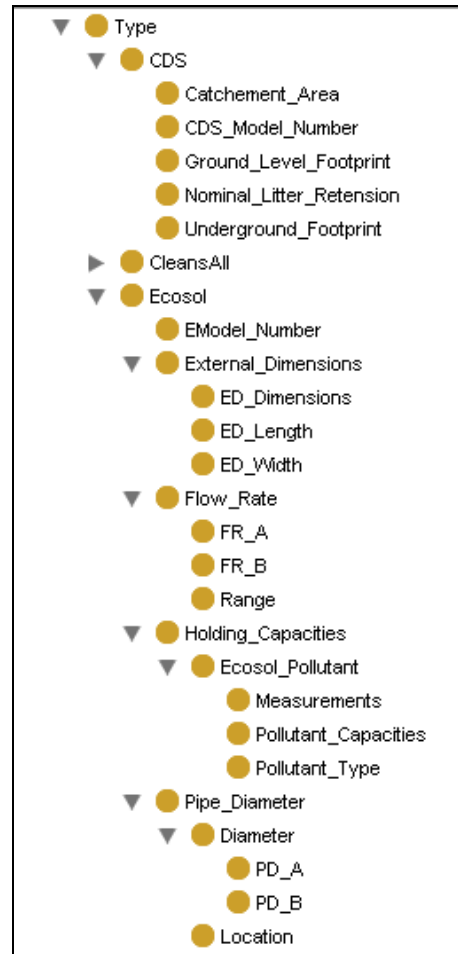


Fig. 8 Type of GPT Classes and Subclasses Onto

## VI. DISCUSSION AND CONCLUSION

To this point the GPT Ontology has produced a total of 2 main classes and 50 sub classes. More classes will be added and the relations between these classes will also be revised accordingly. Progressively and in stages, the relations, functions, axioms and classes will go through several discussions and confirmation with the team of CSGM and researchers. This is why Methontology methodology is used and a few queries will also be devised and tested. After the completion of the basic GPT Onto, we will proceed with the application of SPARQL for data query. Potential users such as researchers and engineers in DID will also be invited to test the results and suggest improvements which need to be made to the prototype and the queries. Finally, GPT Onto will be published and uploaded to the ontology library for public access.

Collecting and storing GTP Onto data will be a challenge as very limited GPT data are not easily made available to the researchers and the public. The GTP Onto is expected to assist engineers and planners in selecting suitable GPT in any Malaysia urbanization area and also to support the Gross Pollutant Management Strategies (GPMS). The development

of the GPT Onto is just at the beginning phase. More iterative stages will be conducted specially towards the construction of the GTP ontology classes, individuals and axioms. A more details representation of the GPT selection and requirements is currently on the way.

#### REFERENCES

- [1] Chandrika M.J. & Mohd S.A, "The Use of Ontology in Developin Syllabus Management System", KMICe2010.
- [2] Chandrika M.J., Lariyah M.S., Liyen L., Sharmila R.M., Huey N.S., Aminah A. & Hidayah B., "An Overview of Knowledge Management in DeGPT System – via Knowledge Audit", RIVERS 2011.
- [3] Luigi C., Ulises C. & Miquel S., "WaWO - An ontology Embedded into an Environmental Decision-support system for Wastewater Treatment Plant Management", 2000, Universitat Politècnica de Catalunya , Spain. Retrieved October 1, 2012, from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.12.486>
- [4] Vasconcelos, J., Kimble, C. and Gouveia F.R., April, "A Design for a group Memory System Using Ontologies", 5th UKAIS Conference, April 2000, University of Wales Institute, Cardiff, McGraw Hill, Forthcoming.
- [5] Nicola Guarin, "Formal Ontology and Information Systems", First International Conference, Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press, pp. 3-15.
- [6] B. Chandrasekaran, John R. Josephson, V. Richard Benjamins, "What are ontologies, and why do we need them?" Intelligent Systems and their Application, IEEE, 1999, pp.20-26.
- [7] M. Bonifacio, P. Bouquet and R. Cuel, "The Role of Classification(s) in Distributed Knowledge Management". In Conference Proceedings KES 2002, Special Session on Classification, Amsterdam, The Netherlands, IOS Press.
- [8] Matteo Cristani, Roberta Cuel, "A Comprehensive Guideline for Building a Domain Ontology from Scratch", Proceedings of I-KNOW '04, Graz, Austria, June 30 - July 2 2004, pg 208.