# Supporting Decision Making in Maritime Environmental Protection with a Knowledge-based Education and Awareness Approach

**Konstantinos Kotis**

University of the Aegean, Dept. of Information and Communication Systems Eng., AI-Lab, Karlovassi, Greece, kotis@aegean.gr

**Andreas Papasalouros**

University of the Aegean, Dept. of Mathematics, Karlovassi, Greece, andpapas@aegean.gr

**Nikitas Nikitakos**

University of the Aegean, Dept. of Department of Shipping Trade and Transport, Chios, Greece, nnik@aegean.gr

**Abstract** In this paper we present an approach which aims to support learners in general and environmental decision makers in particular, towards effective decision making in maritime environmental pollution via education and awareness of specific maritime environmental pollution policies. We build on previous work concerning the automatic construction of multiple-choice questions from ontologies (automatic assessment) and extend it by integrating if-then rules towards building an environmental knowledge base for maritime pollution. Preliminary evaluation of this work is conducted with a prototype environmental pollution (focused on maritime pollution with oil) ontology in OWL and example rules in SWRL for capturing knowledge related to diagnosis, response and environmental-change events of oil spill pollution.

## Introduction

Expert systems for environmental pollution have been around some time (e.g. Meech and Veiga 1997; Ceccaroni et al 2004; Harzikos et al 2008; Karatzas and

Kaltsatos 2007). AI researchers have been working on this topic integrating also new technologies coming from the Semantic Web e.g. (Ceccaroni et al 2004). Work has been done on SWRL to support decision making in knowledge bases for other domains such as Transportation (Gang et al 2008) or Dental domain (Seon and Hong-Gee 2006). Although decision making seems to be well supported on this area, to the best of our knowledge there isn't much that have been done to support environmental decision makers via education and awareness. Intelligent Tutoring Systems (ITS) provide direct customized instruction or feedback to learners whilst performing a task implementing "learning by doing". ITS have been recently proved proper candidates for tackling such issues, using technological advances of Artificial Intelligence techniques in the service of environmental awareness/education and decision making support.

ITS's consist of four different subsystems or modules: the interface module, the expert module, the student module, and the tutor module[1]. The interface module provides the means for the student (learner more generally) to interact with the ITS, usually through a graphical user interface and sometimes through a rich simulation of the task domain the student is learning (e.g., controlling a power plant or performing a medical operation). The expert module references an expert or domain model containing a description of the knowledge or behaviors that represent expertise in the subject-matter domain the ITS is teaching -- often an expert system or cognitive model. An example would be the kind of diagnostic and subsequent corrective actions an expert engineer takes when confronted with an oil pollution alarm at sea. The student module uses a student model containing descriptions of student knowledge or behaviors, including his *misconceptions* and *knowledge gaps*. An apprentice technician might, for instance, not know that an oil spill of 200 tones in a small area of sea surface is not a major oil spill event (knowledge gap) or he may believe that the designated area of oil spoil is small and no action is needed (misconception). A mismatch between a student's behavior or knowledge and the expert's presumed behavior or knowledge is signaled to the tutor module, which subsequently takes corrective action, such as providing feedback or remedial instruction. To be able to do this, it needs information about what a human tutor in such situations would do i.e. the tutor model (Koedinger and Corbett 2006).

An ITS is only as effective as the various models it relies on to adequately model expert, student and tutor knowledge and behavior[1]. Thus, building an ITS needs careful preparation in terms of describing the knowledge and possible behaviors of experts, students and tutors. This description needs to be done in a formal language in order that the ITS may process the information and draw inferences, automatically generating new knowledge as feedback or instructions. Therefore the knowledge contained in the models should be organized and linked to an inference engine. It is through the latter's interaction with the descriptive data that tu-

[1] http://en.wikipedia.org/wiki/Intelligent_tutoring_system

torial feedback is generated in order to support environmental decision making for diagnosis of environmental damage and selection of appropriate responses/actions.

In this paper we propose a built-up on our previous work concerning the proposal of an e-learning approach towards the development of an ITS which automatically constructs multiple-choice questions from any domain ontology. Such built-up is considered as an extension of the OWL knowledge base by integrating SWRL rules. SWRL (W3C 2004b) is a Rule based ontology language, allowing users to take advantage of inferencing new knowledge from existing OWL knowledge bases, towards an OWL/SWRL-based process. We use the maritime environmental pollution as an evaluation domain by representing knowledge needed to capture diagnosis, response and environmental-change events of oil pollution. Such domain is encoded in a prototype OWL ontology and is used in combination to SWRL rules to represent policies and decision making of environmental protection.

SWRL has been developed in order to extend OWL language expressivity, based on a combination of the OWL-DL and OWL Lite sublanguages of the OWL Web Ontology Language (W3C 2004a) with the Unary/Binary Datalog RuleML sublanguages of the Rule Markup Language. SWRL describes the knowledge of OWL ontology by highly abstract syntax expression, which realized the combination between the Horn-like rules and OWL Knowledge Base ($SHOIN(D)=\Sigma$). We use SWRL to formally express productive and deductive rules for diagnosis and response (diagnose and react) policies, in cases where OWL itself is not enough (we refer to the generic example of "parent(?x,?y) $\wedge$ brother(?y,?z) $\Rightarrow$ uncle(?x,?z)" rule) (W3C 2004b) and the additional expressivity power of SWRL is preferred (closer to human way of representing knowledge and easy way of deducing conclusions). The resulted combined knowledge base ($\Sigma$, $P$) is an integration of $SHOIN(D) = \Sigma$ and a finite set of rules $P$.

To the best of our knowledge, although some work has been done towards using SWRL in teaching strategies e.g. (Wang et al 2005), there is not any previous work that seamlessly, and in an automatic fashion, integrates an OWL-DL/SWRL knowledge base with an learning approach to support environmental decision making via education and awareness. In this paper we present a work-in-progress approach which utilizes an environmental ontology and rules (ITS expert model), a set of strategies for identifying the semantics of evaluation material in the form of multiple choice questionnaires (ITS teaching module) and a set of simple techniques for natural language generation (ITS interface module).

In the current version of the proposed approach, no student module is available, thus personalization or complex interaction with students (decision makers) is not supported. We conjecture that the approach can be used by beginners in the environmental pollution decision making domain. Such users do not need to be familiar with the underlining technology of ontologies and knowledge bases, and more important, they do not need to be experts in the domain of environmental pollution. Users must have obtained basic knowledge from text documents or oral presentations related to the domain prior to their questionnaire-based assessment. Such basic knowledge is asserted in the knowledge base manually (currently by knowledge engineers in collaboration with domain experts). Automated population of the ontology with facts is out of the scope of this work.

## The "EnvOPol" Knowledge Base

A knowledge base is a collection of models, stored facts and rules that can be used for problem solving. The "EnvOPol" knowledge base (built for experimentation reasons) integrates a prototype ontology concerning environmental pollution, focusing on maritime pollution by oil. The knowledge has been acquired from Web resources related to sea pollution Factsheets[2], consulting also the hierarchical description of environmental entities provided by the Eionet GEMET thesaurus[3]. Furthermore, domain experts and ontology engineers that have been participating in the experiment contributed their knowledge either informally or formally using ontology engineering tool Protégé[4] ver. 3.4, partially following the ontology engineering methodology HCOME (Kotis and Vouros 2006). An OWL-DL version of the prototype ontology may be viewed at http://www.icsd.aegean.gr/kotis/Ontologies/oilPollution.owl. OWL-DL language was selected due to the maximum expressiveness possible while retaining computational completeness (all conclusions are guaranteed to be computed), decidability (all computations will finish in finite time), and the availability of practical reasoning. Also, OWL-DL is a W3C standard language for Web Documents and applications. Due to space limitations we provide only semantics for a subset of the conceptualizations, in order to be able for readers to follow the examples (model, facts and rules) presented in this paper. A simple hierarchical caption of the ontology is presented in Figure 1.

A main concept is the oil pollution event *(oil_pollution_event ⊑ Event),* which may be of any type, based mainly on its severity importance (currently we have conceptualize *disastrous*, *significant* and *minor* events). Disastrous oil pollution events (*pollution_event_Disastrous_oil_spill ⊑ oil_pollution_event*) are defined as events that concern a large region of oil spill, and the severity of their oil spill and the severity of their spill volume is characterized as disastrous ((*oil_spill_region_size_on_photo ∋ "large") ⊓ (has_oil_spill_volume_severity ∋ oil_spill_volume_severity_disastrous) ⊓ (has_recovery_time_severity ∋ recovery_time_severity_disastrous)).*

2 http://www.ypte.org.uk/environmental-facts.php
3 http://eionet.eu.int/GEMET
4 http://protege.stanford.edu/

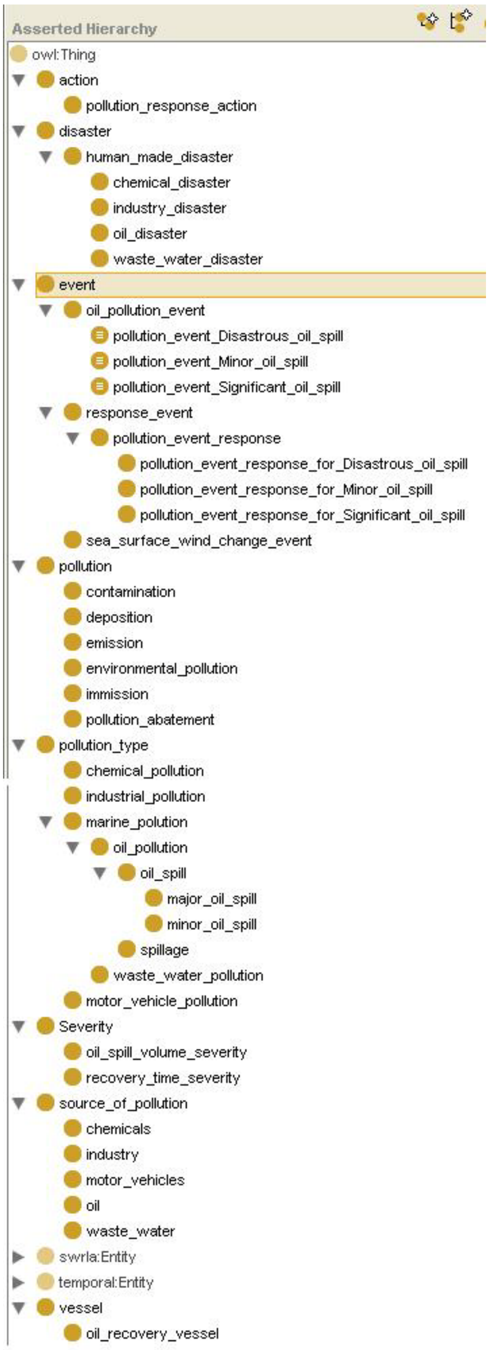
Asserted Hierarchy
owl:Thing
action
pollution_response_action
disaster
human_made_disaster
chemical_disaster
industry_disaster
oil_disaster
waste_water_disaster
event
oil_pollution_event
pollution_event_Disastrous_oil_spill
pollution_event_Minor_oil_spill
pollution_event_Significant_oil_spill
response_event
pollution_event_response
pollution_event_response_for_Disastrous_oil_spill
pollution_event_response_for_Minor_oil_spill
pollution_event_response_for_Significant_oil_spill
sea_surface_wind_change_event
pollution
contamination
deposition
emission
environmental_pollution
immission
pollution_abatement
pollution_type
chemical_pollution
industrial_pollution
marine_polution
oil_pollution
oil_spill
major_oil_spill
minor_oil_spill
spillage
waste_water_pollution
motor_vehicle_pollution
Severity
oil_spill_volume_severity
recovery_time_severity
source_of_pollution
chemicals
industry
motor_vehicles
oil
waste_water
swrla:Entity
temporal:Entity
vessel
oil_recovery_vessel

**Fig. 1.** A hierarchical caption of the ontology taken from Protégé tool

Similarly we define minor and significant oil spill pollution events. The severity of oil spill volume and of the recovery time are primitive classes that classify severity individual objects created for different measurements of recovery time (measured in years) or oil volume (measured in tonnes) respectively. For instance, the *recovery_time_severity_disastrous* individual object describes (with its properties inherited by the related class) the time needed to recover from an event with a disastrous severity i.e. *min_severity_value* property with a value of 100.

A response to an oil pollution event is described as another type of event *(pollution_event_response* $\sqsubseteq$ *Event).* Based on the severity of a pollution event, we distinguish different types of responses, each one initiating different actions for recovery *(* $\forall$*initiate_action. pollution_response_action).* Each response event is related to pollution event e.g. a pollution event response for a disastrous oil spill concerns a pollution event of a disastrous oil spill *(* $\forall$*concerns_event.pollution_event_Disastrous_oil_spill).* An inverse build-in OWL property (*inverseOf(concerns_event, concerns_response))* ensure that events and responses are related in both directions.

Finally, in order to be able to experiment with reasoning related to environmental change knowledge, another type of event is represented, the event of a wind change on the sea surface (*sea_surface_wind_change_event* $\sqsubseteq$ *Event*). Such an event is recorded by sensor input however in our case only simulation data is used for experimentation. Individuals of this event type are different recordings of sea surface wind speed (datatype property with allowed values of "low", "medium", "high") at specific time and date of a specific location.

Using the OWL-DL axioms specified in the KB, we are able not only to assert specific oil pollution events that are fully identified (and assign a specific pollution event response) but also to infer new events by computing inferred types. The inference of such knowledge is achieved via a reasoning mechanism (Pellet 1.5[5]) and the proper design of defined classes (necessary and sufficient conditions). However, as already stated, the "EnvOPol" knowledge base was extended with deductive and production rules in order to represent knowledge for diagnosis and response (diagnose and react) using the SWRL formalism. Some example rules are provided below:

*Example rule set A:* ("*discover which oil pollution events are disastrous based on their severity and oil spill size region on a satellite photo*" and "*retrieve the responses available for such a disastrous event*"):

1. oil_pollution_event(?e)
   $\wedge$ has_oil_spill_volume_severity(?e, oil_spill_volume_severity_disastrous)
   $\wedge$ has_recovery_time_severity(?e, recovery_time_severity_disastrous)

[5] http://clarkparsia.com/pellet/

∧ oil_spill_region_size_on_photo(?e, "large")
→ pollution_event_Disastrous_oil_spill(?e)

2. pollution_event_Disastrous_oil_spill(?e)
∧ pollution_event_response_for_Disastrous_oil_spill(?r)
→ concerns_response(?e, ?r)

3. concerns_response(?e, ?r) → sqwrl:selectDistinct(?r)

*Example rule set B: ("discover which oil pollution events are minor based on their severity and oil spill size region on a satellite photo"* and *"select those which need to be upgraded to disastrous because of a sea surface wind change event with specific characteristics")*:

1. oil_pollution_event(?e)
∧ has_oil_spill_volume_severity(?e, oil_spill_volume_severity_minor)
∧ has_recovery_time_severity(?e, recovery_time_severity_minor)
∧ oil_spill_region_size_on_photo(?e, "small")
→ pollution_event_Minor_oil_spill(?e)

2. sea_surface_wind_change_event(?w) ∧ time(?w, ?wTime) ∧ date(?w, ?wDate)
∧ location(?w, ?wLocation) ∧ pollution_event_Minor_oil_spill(?e)
∧ time(?e, ?eTime) ∧ date(?e, ?eDate) ∧ location(?e, ?eLocation)
∧ windSpeed(?w, ?sNew) ∧ windSpeed(?e, ?sOld)
∧ swrlb:notEqual(?sNew, ?sOld) ∧ swrlb:matches(?sNew, "high")
∧ swrlb:lessThanOrEqual(?eDate, ?wDate)
∧ swrlb:lessThanOrEqual(?eTime, ?wTime)
∧ swrlb:matches(?eLocation, ?wLocation)
→ sqwrl:selectDistinct(?e) ∧ upgrade_to_disastrous(?e, "true")

In this human-readable syntax, a rule has the form: *antecedent → consequent,* where both *antecedent* and *consequent* are conjunctions of atoms written $a_1 \wedge ... \wedge a_n$.

## The "QuGAR-OWL" ITS approach

QuGAR-OWL (*A*utomatic *G*eneration of *Qu*estion items from *R*ules and *OWL* ontologies) is an e-learning approach towards an ITS that generates multiple choice questionnaires from populated OWL ontologies in an automatic fashion (Papasalouros et al 2008). The approach utilizes ontologies that represent both domain and multimedia knowledge. Multimedia questionnaires are currently restricted to items with images. For evaluation and experimental purposes we have produced results with a number of domain ontologies for text-based questionnaires. The approach is open to any source of knowledge that can be mapped to OWL semantics and of course to any source that already uses OWL semantics to represent its knowledge. Heterogeneous and distributed domain-specific knowledge can also be automatically transformed in a QuGAR-OWL-generated questionnaire, given that there is an OWL model that these resources can be mapped to (and aligned).

Certain strategies have been identified and used for selecting the correct answers in question items, as well for selecting distractors (Kehoe 1995). The selected strategies are analytically presented in (Papasalouros et al 2008). Below we provide a simple strategy and a related example question automatically generated for the maritime environmental pollution ontology.

- Strategy A (text-based):

Choose individuals which are not members of a given class, provided that they are members of one of its superclasses. More specifically, if *A(a)* for some *a,* then correct answer is: *A(a)*. For the distractors selection, we assume that *B* is a superclass of *A*. Then, if *B(b)*, *b≠a* and *b* is not an individual of *A*, then *A(b)* is a distractor.

- Generated Question A:

Which of the following sentences is true?

A. PERM01 is a pollution event response for Minor oil spill. (C)
B. PERS01 is a pollution event response for Minor oil spill. (D)
C. PERD01 is a pollution event response for Minor oil spill. (D)
D. PERD02 a pollution event response for Minor oil spill. (D)

In the above, only choice A is a correct answer, indicated with (C), since PERM01 is an individual of ontology class *pollution_event_response_for_Minor_oil_spill.* The other choices, indicated with a (D), are distractors, containing individuals which belong to disjoint sibling classes of the above class (OWL disjointWith axiom has been utilized).

Preliminary work on extending QuGAR-OWL approach to handle rules also (specifically SWRL rules) used with problem solving related domains such as the environmental protection/pollution domain, proves that it can be used as a support tool for improving the effectiveness of decision making via education and awareness of diagnosis/response policies. More specifically, we identify a number of new strategies that extend our previous work with text-based and multimedia-based strategies. In this paper we present the first two rule-based strategies (Strategy B and Strategy C).

- Strategy B (rule-based):

Given that $d1 \wedge d2 \wedge ... \wedge dm \rightarrow v1 \wedge v2 \wedge ... \wedge vk$ is a rule in the knowledge base, where x is a variable and C is a class, and one of the atoms $v1, v2, ..., vk$ in the head of the rule is in the form C(x), then a multiple choice question item can be formed as follows: The rule provides the semantics for the correct answer and distractors are selected among *disjoint siblings* of or among *subclasses* of C. As an example we assume that the following rule exists in the knowledge base:

1. oil_pollution_event(?e)
   ∧ has_oil_spill_volume_severity(?e, oil_spill_volume_severity_disastrous)
   ∧ has_recovery_time_severity(?e, recovery_time_severity_disastrous)
   ∧ oil_spill_region_size_on_photo(?e, "large")
   → pollution_event_Disastrous_oil_spill(?e)

Based on concept *pollution_event_Disastrous_oil_spill,* which appears in the head of the above rule, this strategy generates question items as in the following example.

- Generated Question B:

If an oil pollution event has disastrous oil spill volume severity and disastrous recovery time and large region size on photo, then the pollution event is a(n):

A. Disastrous oil spill pollution event (C)
B. Oil spill pollution event
C. Minor oil spill pollution event (D)
D. Significant oil spill pollution event (D)

In the above example, the correct answer is indicated by (C), while the wrong answers (distractors) are indicated by (D) (for presentation reasons only in the paper).

- Strategy C (rule-based):

For a rule in the form $d1 \wedge d2 \wedge ... \wedge dm \rightarrow v1 \wedge v2 \wedge ... \wedge vk$, if one of the atoms $d1, d2, ..., dk$ in the body of the rule is in the form C(x), where x is a variable and C is a class, then generate a sentence based on the rule as correct answer. Distractors are generated by substituting C with one of its *super-classes* or one of its *disjoint siblings*.

As an example, classes *pollution_event_Disastrous_oil_spill(?x)* and *pollution_event_response_for_Disastrous_oil_spill(?y)* appear as atoms in the head of the following rule:

1. pollution_event_Disastrous_oil_spill(?e)
   ∧ pollution_event_response_for_Disastrous_oil_spill(?r)
   → concerns_response(?e, ?r)

- Generated Question C:

Which of the following is correct?

A. A disastrous pollution oil spill event concerns a disastrous pollution oil spill event response. (C)
B. A pollution oil spill event concerns a disastrous pollution oil spill event response (D)

C. A disastrous pollution oil spill event concerns a pollution oil spill event response (D).
D. A minor oil spill event concerns a pollution oil spill event response (D).

In current version of QuGAR-OWL, natural language generation is based on the names of ontology classes and properties, provided that they follow certain conventions. Future work should tackle the problem of generating natural language items from domain-specific OWL and SWRL semantics with further study of OWL-to-NLG techniques (e.g. the work presented in Karakatsiotis et al (2007)).

## Conclusion and Future Work

In this paper, building on our previous work on ITS, we present preliminary results of novice and original work towards a) a maritime environmental pollution knowledge base (model, facts, rules), b) the extension of ITS to handle rules for the automatic generation of multiple choice questions, c) the use of the proposed ITS extension to support decision making via education and awareness in the domain of maritime environmental protection. Since this is a work in progress, we need to implement and evaluate the rule-based question generation strategies within the prototype intelligent tutoring system. Furthermore, issues such as interaction and feedback should be explored since currently we only consider interaction within the task of capturing multimedia knowledge by annotating images, and we generate feedback only from the correct/wrong answers. In the current version of the tool, no student module is available, thus personalization or complex interaction with students is not supported. Furthermore, users must obtain basic knowledge from text documents or oral presentations related to the domain. Such basic knowledge is asserted in the knowledge base manually (currently by knowledge engineers). Future work concerns the active participation of decision makers in the knowledge base development process, following a human-centered and collaborative ontology engineering approach supported by Wiki-based argumentation technology. Finally, the problem of generating natural language items from domain-specific OWL and SWRL semantics should be tackled with further study of OWL-to-NLG techniques.

## References

Ceccaroni L, Cortes U, Sanchez-Marre M (2004) OntoWEDSS: augmenting environmental decision-support systems with ontologies. Environmental Modelling & Software 19, pp. 785–797

Gang C, Qingyun D, Hongli M (2008) The Design and Implementation of Ontology and Rules Based Knowledge Base for Transportation. International Conference on Computer Science and Software Engineering, pp.1035-1038

Hatzikos EV, Tsoumakas G, Tzanis G et al (2008) An Empirical Study on Sea Water Quality Prediction. Knowledge Based Systems 21(6):471-478

Karakatsiotis G, Galanis D, Lampouras G et al (2008) NaturalOWL: Generating Texts from OWL Ontologies in Protege and in Second Life. System demonstration, 18th European Conference on Artificial Intelligence, Patras, Greece

Karatzas K, Kaltsatos S (2007) Air pollution modelling with the aid of computational intelligence methods in Thessaloniki, Greece. Simulation Modelling Practice and Theory, 15(10):1310-1319

Kehoe J (1995) Writing multiple-choice test items. Practical Assessment, Research & Evaluation. Vol. 4 No. 9, retrieved February 2009 from http://pareonline.net/getvn.asp?v=4&n=9

Koedinger R, Corbett A (2006) Cognitive Tutors: Technology bringing learning science to the classroom. In Sawyer, K., The Cambridge Handbook of the Learning Sciences, Cambridge University Press, pp. 61–78

Kotis K, Vouros AG (2006) Human-Centered Ontology Engineering: the HCOME Methodology. International Journal of Knowledge and Information Systems (KAIS), 10(1): 109-131

Meech A, Veiga M (1997) Predicting the impact of mercury pollution with a fuzzy expert system. Systems, Man, and Cybernetics, Vol. 2, IEEE press, pp.1056–1061

Papasalouros A, Kotis K, Kanaris K (2008) Automatic generation of multiple-choice questions from domain ontologies. IADIS e-Learning 2008 (eL 2008), Amsterdam

Seon P, Hong-Gee K (2006) Dental Decision Making on Missing Tooth Represented in an Ontology and Rules. Springer Berlin / Heidelberg 2006, Vol. 4185, pp. 322-328

W3C (2004a) OWL Web Ontology Language Reference. http://www.w3.org/TR/owl-features/

W3C (2004b) SWRL: A Semantic Web Rule Language Combining OWL and RuleML. http://www.w3.org/Submission/SWRL/

Wang E, Kashani L, Kim YS (2005) Teaching Strategies Ontology Using SWRL Rules. International Conference on Computers in Education (ICCE), Singapore