ViWaT: A Lightweight Ontology of the Vietnamese Wastewater Treatment Management with Biological Methodologies

Thanh Ma[1], Thanh-Nghi Do[1], Nguyen-Khang Pham[1], Minh-Thu Tran-Nguyen[1], Tri-Thuc Vo[1] and Salem Benferhat[2]

[1]CICT, Can Tho University, Vietnam
[2]CRIL, Artois University&CNRS, UMR 8188, France

Abstract
Wastewater directly impacts individuals' health; hence, this topic is always a flame to motivate scientists to research treatment solutions. In terms of economy and infrastructure, building wastewater treatment systems is a significant and prominent challenge and problem in developing countries. Therefore, to construct a foundation for determining wastewater treatment solutions and building their management systems, we have proposed a lightweight ontology for wastewater treatment (WT) management. In particular, we concentrate on the keys of Biological Methodologies because these approaches are suitable for the environment and topography of Vietnam, especially at low cost. Accordingly, we pay attention to implementing this modeling by analyzing and classifying the primary biological and geographical features to serve in searching WT companies. Moreover, we leverage description logic that underlies the web ontology language OWL2-EL for representation, reasoning, and querying WT ontology-based modeling.

Keywords
Wastewater Treatment, Ontology, Description Logic, Biological Methodologies

1. Introduction

Environmental pollution is one of the significant challenges of countries; each government is implementing novel policies to deal with the increasingly severe pollution problem, especially in wastewater solutions. For that reason, managers and scientists are paying attention to these aspects to seek suitable solutions, i.e., leveraging the power of the technology for wastewater treatment (WT) and constructing a WT system. Many wastewater treatment approaches [1] are currently proposed, including mechanical, chemical, and others.

- *Mechanical wastewater treatment method:* Using cleaning equipment to remove large-scale solids or residues in wastewater.
- *Methods of treating physicochemical wastewater:* In wastewater containing inorganic acids or alkalis, it should be neutralized to obtain a PH of 6.5 - 8.5 before the wastewater conveys into the receiving source or treatment technology.
- *Biological wastewater treatment method:* A treatment method based on the living activities of microorganisms in wastewater. These tiny microorganisms will continually metabolize organic matter by synthesizing new cell walls solely. During biological wastewater treatment, toxic substances in wastewater will be converted by microorganisms into harmless substances.

Generally, biological solutions [2] obtain the highest efficiency, safety, and environmental benevolence. Especially the cost of the project deployment is downward and suits agricultural governments and developing countries. Wastewater treatment via the biological strategy treats dissolved organic substances in wastewater and several other inorganic pollutants, including H_2S, $sulfide$, $ammonia$, $nitrogen$, and others, based on microbial activity to degrade organic contaminants. Namely, microorganisms use organic matter and some minerals in wastewater as nourishment to grow and develop. The low-cost wastewater treatment methods are as follows: (1) Biological lake [3] (anaerobic lake, arbitrary lake, and aerobic lake thoroughly treated reservoir), (2) Centralized treatment *(wastewater stabilization biological pond, submerged filtered field, interrupted sand filtration, wastewater treatment system by plant, anaerobic wastewater treatment method)*, (3) wastewater treatment on site (septic tank, biogas tank, underground filtered field)

Moreover, Vietnam is an agriculturally developing country with many different terrains; for example, the North has most hills, the South has multiple plains, and the central territory has perennial crops. Each area has soil sorts in which the environment is different. On the other hand, each region's wastewater classification de-

ENIGMA-23, September 03–04, 2023, Rhodes, Greece
mtthanh@ctu.edu.vn (T. Ma); dtnghi@ctu.edu.vn (T. Do); pnkhang@ctu.edu.vn (N. Pham); tnmthu@ctu.edu.vn (M. Tran-Nguyen); vtthuc@ctu.edu.vn (T. Vo); benferhat@cril.fr (S. Benferhat)

pends on the individuals and that area's living habits. Therefore, the wastewater treatment must suit each region's topography and mud (soil).

In general, managing wastewater treatment in Vietnam is still challenging because individuals' awareness of waste disposal and the cost of operating and deploying treatment plants still need improvement. Moreover, the data management of wastewater treatment companies needs to be sufficient and centralized. Therefore, we propose a lightweight ontology for wastewater management by biological methods widely applied in Vietnam.

This paper analyzes Vietnam's biological methods and develops a lightweight ontology for addressing domestic wastewater treatment plants (referring to [4]). Here, we need an OWL file to be the foundation for building the actual application. We will use description logic to demonstrate the implementation of the code in this paper because the DLs family [5] provides a solid foundation of Web Ontology Language (OWL). One of OWL2's profiles is OWL2-EL [6, 7], dedicated to applications that use huge volumes of data where query-answering is the most significant reasoning task. In addition, this paper also focuses on querying wastewater treatment companies and their methods used in that area. In particular, most of the knowledge on waste treatment is still being determined and insufficient. The storage method is still fragmented and inconsistent, leading to difficulties in searching, inferring, and classifying. For the above reasons, implementing the description logic $\mathcal{EL}_{\bot}$ as a fundamental core is a suitable and desirable selection.

In the research process, we decomposed our approach into three prominent stages: (1) the first aspect is to construct an ontology to store WT's facts which are implementing in Vietnam; (2) the second direction is to develop mobile applications and websites established on the ontology proposed to synthesize the knowledge and to allow the company and corporations that can update on current and future technologies. Our targeting application is a channel for sharing the knowledge and recommendation system of suggesting appropriate biological methods for each location. (3) the third angle is to propose solutions to solve the inconsistency and uncertainty [8, 9, 10] when storing ViWaT's knowledge. In this paper, we mainly concentrate on the first step by mentioning the characteristics of Vietnam related to wastewater treatment. After that, we put all our effort into presenting how to implement ViWaT ontology using description logic $\mathcal{EL}_{\bot}$ for reasoning and query-answering. Our main challenge was identifying the key concepts and attributes needed through our collective pieces of knowledge and expert opinions to drive the application development in the next step.

The remainder of this paper is structured as follows: in the next section (Section 2), we give an overview of wastewater treatment in Vietnam. Section 3 will provide a brief description of the popular biological methodologies. Moreover, we also put the structure of biological wastewater treatment in this section. In Section 4, we discuss and illustrate how to implement an ontology through description logic $\mathcal{EL}$ for representation, reasoning, and query-answering. Finally, section 5 concludes the paper.

2. Wastewater Treatment in Vietnam

Vietnam's municipalities [11, 12] generally evolve significantly fast, and constructing urban drainage and wastewater treatment is increasingly challenging. Therein, the investment issue in urban infrastructure technical systems in general and urban water collection and drainage systems, in particular, needs to be solved. Many provinces and metropolises have asynchronous infrastructure planning, failing to keep up with urban development. In addition to that, the need for more technical infrastructure for wastewater treatment is quite expected. Many inappropriate technological treatment systems have led to untreated urban domestic wastewater being blatantly discharged into the environment, threatening the ecological environment and becoming a great challenge for cities in Vietnam.

According to environmental management reports, household wastewater accounts for 64% of Vietnam's customary wastewater. However, only about 13% of this wastewater is treated according to the process. In Vietnam, 60% of households discharge wastewater into generally shared sewers. However, most of the wastewater discharges directly into the surface drainage system. And only 10% of the wastewater is treated, while 90% of the households dump the wastewater into the septic tank. Several techniques of wastewater treatment in Vietnam are presented in Table 1.

Wastewater treatment in Vietnam [13, 14] has some differences compared to other countries, whereas furthermore encloses similarities. Several differences and similarities in wastewater treatment in Vietnam compared to other policies are as follows:

1. *Scale and infrastructure*: Vietnam, especially in rural areas, has an underdeveloped wastewater treatment infrastructure and needs a comprehensive public treatment system. Meanwhile, several developed countries have invested in extensive and modern public wastewater treatment systems.
2. *Management and enforcement*: Vietnam is improving its regulatory and enforcement processes related to wastewater treatment. However, there are challenges in complying with and monitoring

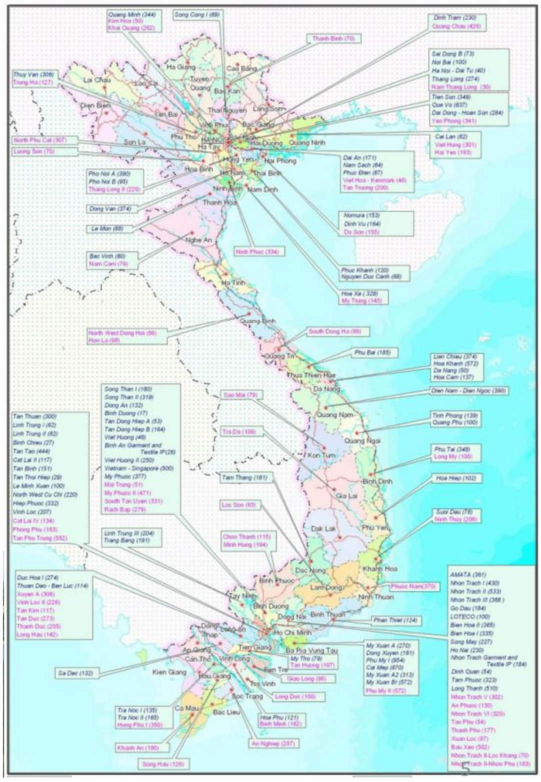


Figure 1: Vietnam map to show the distribution of industrial parks.

wastewater treatment regulations. Some developed countries have robust environmental management systems and more straightforward enforcement procedures.

3. *Awareness and education*: Awareness of wastewater management and treatment in Vietnam is advancing, but there is a need for education and public awareness on the importance of environmental protection and wastewater treatment. Meanwhile, developed countries have a clearer vision of water management and environmental awareness from the public.

Moreover, according to statistics from the General Department of Environment[1], 209 industrial parks have been operated nationwide, with a total area of 47, 300 hectares. About 80% of industrial parks in Vietnam have a centralized wastewater treatment system, and the remaining 20% have not or are investing in a centralized wastewater treatment system. Regarding wastewater treatment measures, in Vietnam, industrial parks and companies classify according to technology groups as follows *(the industrial parks and companies are shown in Figure 1)*:

- *Traditional technology* with biological treatment by activated mud and biological filter tank;
- WT technology with *aerobic biological treatment* by adherent growth organisms;
- WT technology with *long-lasting activated sludge biological* treatment.

Most domestic wastewater in Vietnamese households [15] originates from the kitchen, bathroom (gray water), and black water from the toilet. Black water is usually treated in indoor septic tanks. Gray water is discharged directly into the water system. In several households, black and gray water are released into the soil to fertilize plants. From there, soil pollution and other pollution start to emerge. Generally, most domestic wastewater is treated by the septic method. The flow of Vietnam's wastewater is shown in Figure 2

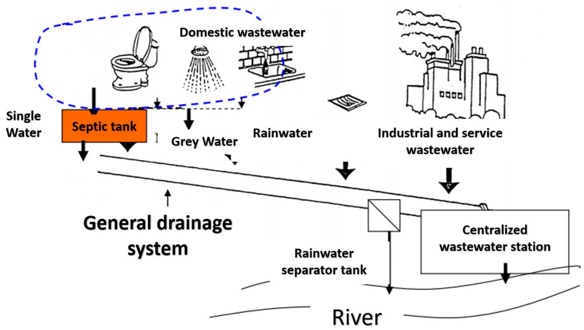


Figure 2: Wastewater flow in Vietnam.

3. Biological Methodologies

Biological wastewater treatment [16, 17, 18, 19] is an essential and integral stage of the wastewater treatment strategy, and it treats wastewater arriving from either residential buildings or industries. They are often called the "Secondary Medicine" process, which removes any contaminants left over after primary treatment. Chemical treatment of wastewater uses chemicals to react with pollutants present in the wastewater, whereas biological treatment uses microorganisms to degrade wastewater contaminants. This treatment relies on bacteria, nematodes, algae, fungi, protozoa, and rotifers to break down hazardous organic wastes using normal cellular processes to stable inorganic forms. This section briefly presents several popular biological methods used in Vietnam. Classification of biological techniques based on the activity of microorganisms is as follows:

[1] https://monre.gov.vn/

ID	Companies	Cities	Year Begins	Wattage (m^2/day)	Wastewater System	Process / Techniques
1	Kim Lien	Ha Noi	2005	3,700	General	Aerobic, Anaerotank, Anaerobic (activated mud)
2	Truc Bach	Ha Noi	2005	2,500	General	Aerobic, Anaerotank, Anaerobic (activated mud)
3	Bac Thang Long	Ha Noi	2009	7,000	General	Aerobic, Anaerotank, Anaerobic (with denitrification)
4	Yen So	Ha Noi	2012	120,000	General	Aerobic, Anaerotank, Anaerobic (Reactor tank)
5	Binh Hung	Ho Chi Minh	2009	141,000	General	Traditional activated sludge
6	Binh Hung Hoa	Ho Chi Minh	2008	30,000	General	Aerobic lake
7	Canh Doi (Phu My Hung)	Ho Chi Minh	2007	10,000	Private	Oxidation ditch
8	Son Tra (Phu My Hung)	Ho Chi Minh	2009	15,000	Private	Aerobic, Anaerotank, Anaerobic (activated sludge)
9	Son Tra	Da Nang	2006	15,900	General	Anaerobic pond with tarpaulin
10	Hoa Cuong	Da Nang	2006	36,418	General	Anaerobic pond with tarpaulin
11	Ngu Hanh Son	Da Nang	2006	11,629	General	Anaerobic pond with tarpaulin
12	Bac Giang	Bac Giang	2010	8,000	General	Oxidation ditch
...	...	...	...	...	...	...

Table 1
The number of synchronized concepts for each level.

1. *The aerobic method* [20] is a method that uses groups of aerobic microorganisms. For their operation, the microorganism must continuously deliver oxygen and maintain a temperature between 20 and 40^oC. This method is applied widely in Vietnam since Vietnam's temperature is in the range of microorganism activities. It is the decomposition process of organic and inorganic compounds in the presence of oxygen and microorganisms, and at the same time, increasing biomass or intracellular decomposition thanks to a part of organic and energy extraction. The process mechanism goes as follows:
   - Organic compounds **without nitrogen**:

   $$C_xH_yO_z + (x + \frac{y}{4} - \frac{z}{2})O_2 \longrightarrow xCO_2 + \frac{y}{2}H_2O$$

   - Organic compounds **with nitrogen**:

   $$C_xH_yO_z\mathrm{N} + (x + \frac{y}{4} - \frac{z}{2} - \frac{3}{4})O_2 \longrightarrow xCO_2$$
   $$+(\frac{y}{2} - \frac{3}{2})H_2O + NH_3$$

   - Biomass synthesis:

   $$C_xH_yO_zN + (x + \frac{y}{4} - \frac{z}{2} - \frac{23}{4})O_2 + NH_3 \longrightarrow$$
   $$2C_5H_7NO_2 + (x-10)CO_2 + (\frac{y}{2} - 7)H_2O + NH_3$$

   - Intracellular destruction:

   $$C_2H_7NO_2 + 5O_2 \longrightarrow 5CO_2 + 2H_2O + NH_3$$

2. *The anaerobic method* [20, 21] decomposes organic substances without oxygen, resulting in methane, carbon dioxide, nitrogen, hydrogen, and others.

   $$(CHO)_nNS \longrightarrow CO_2 + H_2O + CH_4$$
   $$+NH_4 + H_2 + H_2S + \textit{Microbial cells}$$

3. *Anoxic/Anaerotank method*: In an environment where oxygen is unavailable, aerobic microorganisms cannot function; instead, anaerotank microorganisms leverage other sources of oxygen. As a result, the O atom and $NO2^-$ and $NO3^-$ molecules will appear. Then, the decomposition process produces energy and produces more microorganisms:

   $$CHC + NO_2^- \mid NO_3^- + \text{microorganisms} \longrightarrow CHC$$
   $$+CO_2 + N_2 + \text{more microorganisms}$$

From the popular biological approaches and expert knowledge, we propose a schema of the structure of Biological Wastewater Treatment, presented in Figure 3. The companies have applied various methods depending on their region. Namely, several companies are involved in A2O technology [22] (a.k.a. AAO) with three manners (anaerobic, anoxic, and oxic). i.e., Kim Lien, Truc

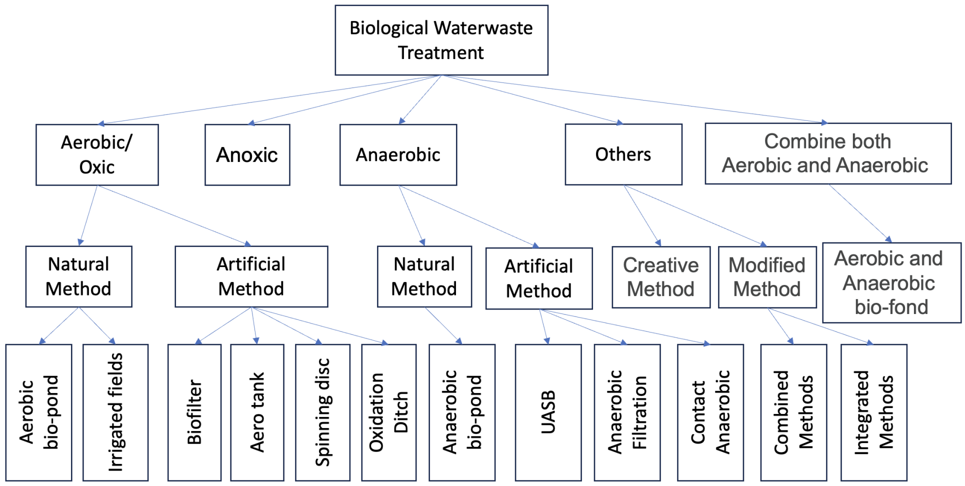


Figure 3: Hierarchy of Wastewater Treatment via Biological Methodologies

Bach, North-Thang Long, some companies utilize SBR [23] (Sequence Batch Reactor) system for the treatment of biological wastewater containing high organic matter and nitrogen, treatment of wastewater with activated sludge in a fill-and-discharge manner, i.e., Tu Son, Bac Ninh, Yen So. Other companies use "Anaerobic Pond with float cover", i.e., Ngu Hanh Son, Son Tra, Hoa Cuong (in Da Nang province). Moreover, because of the soil and type of wastewater in that area, several enterprises have modified some active ingredients and biological formulations, i.e., Ho Tay, Cau Nga (they operate the modified SBR). Hence, to manage the specific information of companies in different regions, we named the instances after the company name and correction method, i.e., AAO-KimLien, and SBR-HoTay. Although companies have mostly stayed the same since the original technology, our ontology allows using instances of the original name. Regarding the unexplored and innovative approaches, we have created an "others" concept to store new solutions.

4. Description Logic

$\mathcal{EL}$ is a family of lightweight DLs, which underlies the Ontology Web Language profile OWL2-EL, considered one of the main representation formalisms to express terminological knowledge [24]. The main ingredients of DLs are individuals, concepts, and roles, which correspond at the semantic level to objects, sets of objects, and binary relations between objects. More formally, let N_C, N_R, N_I be three pairwise disjoint sets where N_C denotes a set of atomic concepts, N_R denotes a set of atomic relations (roles), and N_I denotes a set of individuals. In this paper, we consider $\mathcal{EL}_\bot$ concept expressions [25] which are built according to the following grammar:

$$C ::= \top \mid \bot \mid N_C \mid C \sqcap C \mid \exists r.C$$

where $r \in N_R$. Let $C, D \in N_C$, $a, b \in N_I$, and $r \in N_R$. An $\mathcal{EL}$ ontology $\mathcal{O} = \langle \mathcal{T}, \mathcal{A} \rangle$ (a.k.a. knowledge base) comprises two components, the TBox (Terminological Box denoted by $\mathcal{T}$) and ABox (denoted by $\mathcal{A}$). The TBox consists of a set of General Concept Inclusion (GCI) axioms of the form $C \sqsubseteq D$, meaning that C is more specific than D or C is subsumed by D, and axioms of the form $C \sqcap D \sqsubseteq \bot$, meaning that C and D are disjoint concepts. $C \equiv D$ which is a shortcut for $C \sqsubseteq D$ and $D \sqsubseteq C$. The ABox is a finite set of assertions on individual objects of the form $C(a)$ or $r(a, b)$.

The semantics is given in terms of interpretations $\mathcal{I} = (\Delta^\mathcal{I}, \cdot^\mathcal{I})$, which consist of a non-empty interpretation domain $\Delta^\mathcal{I}$ and an interpretation function $\cdot^\mathcal{I}$ that maps each individual $a \in N_I$ into an element $a^\mathcal{I} \in \Delta^\mathcal{I}$, each concept $A \in N_C$ into a subset $A^\mathcal{I} \subseteq \Delta^\mathcal{I}$, and each role r into a subset $r^\mathcal{I} \subseteq \Delta^\mathcal{I} \times \Delta^\mathcal{I}$.

A summary of the syntax and semantics of $\mathcal{EL}_\bot$ is shown in Table 2. An interpretation $\mathcal{I}$ is said to be a model of (or satisfies) an axiom Φ in the form of the left column in the table, denoted by $\mathcal{I} \models \Phi$, when the corresponding condition in the right column is satisfied. For instance, $\mathcal{I} \models C \sqsubseteq D$ if and only if $C^\mathcal{I} \subseteq D^\mathcal{I}$. Similarly, $\mathcal{I}$ satisfies a concept (resp. role) assertion, denoted by $\mathcal{I} \models C(a)$ (resp. $\mathcal{I} \models r(a, b)$), if $a^\mathcal{I} \in C^\mathcal{I}$ (resp. $(a^\mathcal{I}, b^\mathcal{I}) \in r^\mathcal{I}$). An interpretation $\mathcal{I}$ is a model of an ontology $\mathcal{O}$ if it satisfies all the axioms and assertions in $\mathcal{O}$. An ontology is said to be consistent if it has a model. Otherwise, it is inconsistent. An axiom Φ is entailed by

Syntax	Semantics
$C \sqsubseteq D$	$C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
r	$r^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
a	$a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$
$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
$\top$	$\Delta^{\mathcal{I}}$
$\bot$	$\emptyset$
$\exists r.C$	$\{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}} s.t. (x, y) \in r^{\mathcal{I}}, y \in C^{\mathcal{I}}\}$

Table 2
Syntax and semantics of $\mathcal{EL}_\bot$

an ontology, denoted by $O \models \Phi$, if Φ is satisfied by every model of O. We say that C is subsumed by D w.r.t. an ontology O iff $O \models C \sqsubseteq D$. Similarly, we say that a is an instance of C w.r.t. O iff $O \models C(a)$. An interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ is said to be *fulfilling* when each concept name in the ontology is non-empty in $\mathcal{I}$, i.e., for each concept $C_i \in N_C$, $\cdot^{\mathcal{I}}(C_i) \neq \emptyset$.

The main reasoning task that is considered in terminological ontologies is classification. It consists in computing all the entailed subsumptions ($C \sqsubseteq D$) (and equivalences ($C \equiv D$)) that hold between atomic concepts of an ontology, or the concepts $\top$ or $\bot$. Such a procedure is described in [24], which first consists in transforming the ontology into a normal form using a set of rules, and then performing a classification reasoning process using the set of inference (completion) rules (see [24] for more details). In this paper, we assume that our ontology is provided in *a normal form*, to which we apply completion rules for classification. This classification step is to normalize the ontology. The reason of conducting normalization is to handle and transform the complex axioms into the axioms of all atomic concepts to be simpler for the representation process. We define the $\mathcal{EL}_\bot$ normal form as follows:

Definition 1 (Normal Form $\mathcal{EL}_\bot$). *An $\mathcal{EL}_\bot$ TBox is in normal form if all concept inclusions have one of the following forms:*

$$A \sqsubseteq B, A \sqcap B \sqsubseteq C, A \sqsubseteq \exists r.B, \exists r.A \sqsubseteq B, A \sqcap B \sqsubseteq \bot$$

where $A, B, C \in N_C$ are atomic concepts and $r \in N_R$.

Query-Answering: A query is a first-order logic formula, denoted $q = \{\vec{x} \mid \Phi(\vec{x})\}$, where $\vec{x} = \{x_1, \ldots, x_n\}$ are free variables and n is the arity of q and atoms of $\Phi(\vec{x})$ are of the form $A(u_i)$ or $r(u_i, v_j)$ with $A \in N_C$, $r \in N_R$ and u_i, v_j are terms, i.e., instances of N_I or variables. Furthermore, when $\Phi(\vec{x})$ is of the form $\exists \vec{y}.conj(\vec{x}, \vec{y})$ where $\vec{y}$ are bound variables called existentially quantified variables, and $conj(\vec{x}, \vec{y})$ is a conjunction of atoms of the $A(u_i)$ or $r(u_i, v_j)$ with $A \in N_C$ and $r \in N_R$, q is said to be a conjunctive query (CQ). Note that if we have $n = 0$, q is called a Boolean query (BQ). Here, in the case of a BQ with no bound variables is denoted as a ground query. In addition, when q only contains one atom with no free variables, it is then known as an instance query.

For the BQ, we have $\mathcal{I} \models q$ if and only if $\Phi^{\mathcal{I}} = True$ and $\mathcal{K} \models q$ if and only if $\forall \mathcal{I} \models \mathcal{K} \longrightarrow \mathcal{I} \models q$. For the CQ, q with free variables $\vec{x} = (x_1, \ldots, x_n)$, a tuple of constraints $\vec{a} = (a_1, \ldots, a_n)$ is said to be the certain answer for q over $\mathcal{K}$ if the BQ $q(\vec{a})$ obtained by replacing each variables x_i by a_i in $q(\vec{x})$, evaluates to $True$ for every model of $\mathcal{K}$. Therefore, CQ answering can be reduced to BQ answering.

In this paper, we implement the description logic $\mathcal{EL}_\bot$ because it can create large ontologies with complexity O(n). Furthermore, the constraints of this background are largely compatible and appropriate with the knowledge of wastewater management.

In the next section, we will implement this setting for our ontology. Specifically, we conduct representation, reasoning, and query-answering knowledge of water treatment management in Vietnam.

5. Representation of ViWaT using DL $\mathcal{EL}_\bot$

In this section, we present predominantly how to implement ontology-based modelling of Vietnamese wastewater treatment management using description logic $\mathcal{EL}_\bot$ language. For the presentation, the arisen common questions is as follows: “Which companies and regions should use the Aerobic Lake method?”, What is the treatment method in the Red River Delta? To answer these questions, from the knowledge presented in Section 3, our ontology will focus on the following characteristics: (1) terrain types in Vietnam; (2) sorts of tools and devices used; (3) companies divided by level; (4) regions of the Vietnamese territory; (5) wastewater treatment methods.

In particular, for (1), we have divided the terrain type into three categories: plain, midlands, and hilly. For (2), we focus on the devices and pieces of equipment. For (3), this paper concentrates mainly on managing companies that are split into three levels: lower, middle, and top. We represent (1), (2), and (3) in Figure 4 (in the left side). For (4), we divide Vietnamese territory into four regions: Nord, Central, South, and Highlands. Within each region, we have subdivided down to the level of provinces. It is presented in Figure 4 (in the right side). We divide each level to facilitate querying information and solving problems involving overlapping areas and locations. Moreover, government policies will be decentralized to the provinces; therefore, the distinct groups/levels can store different instances that are comfortable for management.

For (5), based on the structure of the figure 3, we construct a hierarchy of the ViWaT ontology. The concepts

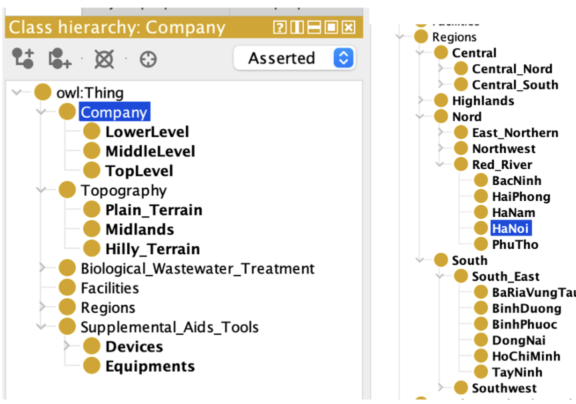


Figure 4: Representing Hierarchy of terrain sorts, kinds of tools, company's level, and Vietnam's regions subdivided down to the level of provinces.

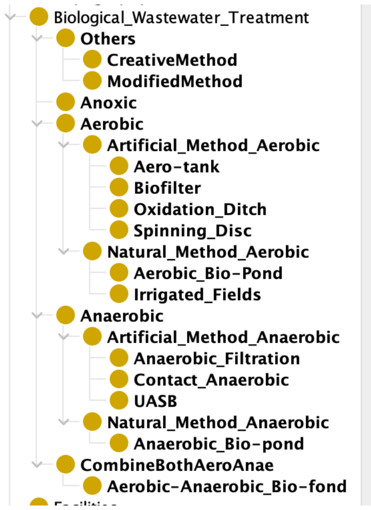


Figure 5: Hierarchy of biological waste water treatment methods.

will store pieces of information on wastewater treatment methods. Here, we allow the company to coordinate multiple solutions for its wastewater treatment. Since description logic $\mathcal{EL}_\bot$ does not allow a "union" concept ($\sqcup$) (but only an "intersection" ($\sqcap$)), therefore, we cannot represent this situation in DL $\mathcal{EL}_\bot$. To solve the above problem, we consider solution combinations as new solutions and store them in the "Creative Method". Figure 5 represents a hierarchy of WT methods. Furthermore, to keep each method's originality and ensure no overlap between techniques during the inference process (reasoning), we create a "disjoin" constraint between several concepts, i.e., $Anoxic \sqcap Anaerobic \sqsubseteq \bot$. That is the reason why we used Description logic $\mathcal{EL}_\bot$ instead of only using DL $\mathcal{EL}$

In the next sub-sections, we will represent ontology by description logic $\mathcal{EL}_\bot$ in the syntactical and semantical side. Specifically, we will present the knowledge in TBox and ABox.

5.1. TBox of ViWaT

For a knowledge base $\mathcal{KB}$, the TBox $\mathcal{T}$ will represent the structure of the ontology. For ease of presentation, we will only express several axioms to focus on answering the common questions mentioned in the section. These axioms will be relevant to ABox and query-answering, which will be presented in the following subsection. Note that we will split the presentation into two parts for TBox: the hierarchy of the ViWaT ontology and the r relationship between the two concepts. We assume the "Thing" concept will be a $\top$ concept. The first subsumption structure of ViWaT ontology will be presented as follows:

$$\mathcal{T}_1 = \begin{cases} Company \sqsubseteq Thing \\ Topography \sqsubseteq Thing \\ BiologicalWaterwaterTreatment \sqsubseteq Thing \\ Region \sqsubseteq Thing \\ Tools \sqsubseteq Thing \\ Midlands \sqsubseteq Topography \\ PlainTerrain \sqsubseteq Topography \\ MiddleLevel \sqsubseteq Company \\ TopLevel \sqsubseteq Company \\ Anaeotank \sqsubseteq BiologicalWaterwaterTreatment \\ Anaerobic \sqsubseteq BiologicalWaterwaterTreatment \\ Aerobic \sqsubseteq BiologicalWaterwaterTreatment \\ ArtificialMethodAerobic \sqsubseteq Aerobic \\ ArtificialMethodAnaerobic \sqsubseteq Anaerobic \\ NaturalMethodAerobic \sqsubseteq Aerobic \\ ArtificialMethodAerobic \sqsubseteq Anaerobic \\ UASB \sqsubseteq ArtificialMethodAerobic \\ AnaerobicFiltration \sqsubseteq ArtificialMethodAerobic \\ Devices \in Tools \\ Equipments \in Tools \\ Anoxic \sqcap Anaerobic \sqsubseteq \bot \\ Nord \sqsubseteq Regions \\ RedRiver \sqsubseteq Nord \\ HaNoi \sqsubseteq RedRiver \\ HaiPhong \sqsubseteq RedRiver \\ \ldots \end{cases}$$

Moreover, to represent the knowledge between concepts, we have several the relations as follows:

$$\vec{r} = \{hasPlaceAt, useTool, applyMethod, Topology, isOfCity, others\}$$

We will use the axioms $A \sqsubseteq \exists r.B$ to express the knowledge. For example, *Kim Lien company located in HaNoi*; we will have a representation as $Compapny \sqsubseteq hasPlaceAtHaNoi$ or $TopLevel \sqsubseteq hasPlaceAtNord$ since Kim Lien Company is a big company so it will belong to TopLevel. These rules will enrich the application implementation. From the inference ability of the description logic, we implement mostly the relationship for the high level (level 0) instead of for the low level (level 3). i.e., $Company \sqsubseteq \exists hasPlaceAt.Region$. However, we also implement low levels (level 3) to keep inference independence, i.e., $MiddleLevel \sqsubseteq \exists applyMethod.Aerobic$.

$$\mathcal{T}_2 = \begin{cases} Company \sqsubseteq \exists hasPlaceAt.Region \\ MiddleLevel \sqsubseteq \exists applyMethod.Aerobic \\ Company \sqsubseteq \exists applyMethod.UASB \\ LowerLevel \sqsubseteq \exists applyMethod.Anaeotank \\ MiddleLevel \sqsubseteq \exists useTool.Devices \\ TopLevel \sqsubseteq \exists useTool.Equipment \\ Company \sqsubseteq \exists hasPlaceAt.HaNoi \\ LowerLevel \sqsubseteq \exists hasPlaceAt.HoChiMinh \\ \ldots \\ HaNoi \sqsubseteq \exists hasTopology.PlainTerrain \\ HoChiMinh \sqsubseteq \exists hasTopology.PlainTerrain \\ HaGiang \sqsubseteq \exists hasTopology.HillTerrain \\ BuonMeThuoc \sqsubseteq \exists hasTopology.Midlands \\ \ldots \end{cases}$$

Note that, our TBox includes both of parts $\mathcal{T}_1$ and $\mathcal{T}_2$ to illustrate the knowledge about subsumption and existence rules. Namely, $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2$. In the next subsection, we will present the facts in ViWaT's ABox.

5.2. ABox of ViWaT

This subsection illustrates several notable cases of the great companies in the main cities, i.e., HoChiMinh City, DaNang City, and the Hanoi capital. Therein, our main purpose is for query-answering to verify and evaluate the feasibility of the implementation process, including *"Kim Lien company in Hanoi has applied Anaerobic method and Binh Hung Hoa company in Vietnam's South has applied one solution for wastewater treatment."*. The significant feature is that Kim Lien has used all three wastewater treatment methods (Anaerobic, Aerobic, and Anoxic). Moreover, all three manners are scattered in different places in the north of Vietnam. We need to determine which city Kim Lien company is using Anaerobic. For the second example, we want to ask where Binh Hung Hoa company is located in the South and what solution to use to treat wastewater. Based on the TBox in the above section, we implement the ABox as follows:

$$\mathcal{A} = \begin{cases} TopLevel(KimLien1) \\ TopLevel(KimLien2) \\ TopLevel(TrucBach) \\ TopLevel(BinhHungHoa) \\ Company(SonTra) \\ Middle(HoaCuong) \\ LowCompany(QuocHuy) \\ HaNoi(DongDaDistrict) \\ HaNoi(CauGiayDistrict) \\ HaNoi(HoGuomLake) \\ HoChiMinh(PhuNhuan) \\ HoChiMinh(TanPhu) \\ UASB(QH\_UASB) \\ AerobicBio-Pond(``BHH\_AerobicLake") \\ Anaerobic(''KimLien\_Anaerobic") \\ hasPlaceAt(KimLien, DongDaDistrict) \\ applyMethod(QuocHuy, QH\_UASB) \\ hasPlaceAt(BinhHungHoa, TanPhu) \\ hasPlaceAt(KiemLien2, HaNoi) \\ applyMethod(KimLien1, KimLien\_Anaerobic) \\ applyMethod(KimLien2, KimLien\_Aerobic) \\ applyMethod(KimLien3, KimLien\_Anoxic) \\ applyMethod(BinhHungHoa, BHH\_AerobicLake) \\ \ldots \end{cases}$$

We allow a company to have multiple instances with the indexes marked, i.e., KimLien with 1,2, and 3. The following section presents query-answering the questions placed in subsection 5.2

5.3. Query-Answering ViWaT

Based on available information, we assume that Kim Lien company is applying the Anaerobic method to answer the question of where this company is; and the name of Binh Hung Hoa company to answer the question of where this company is and what its wastewater treatment is. To answer the above question, we coded with the following query:

$$Q(x, y) \rightarrow Company(x) \wedge hasPlaceAt(x, y)$$

$$\wedge\ applyMethod(x, KimLien\_Anaerobic)$$

The query means that searching the $KimLien\_Anaerobic$ method applied with x and x is located at y. The answer from ViWaT ontology will be x = "$KimLien1$" and y = "$HaNoi$"

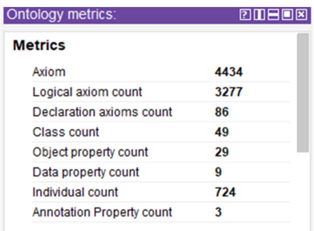
The second question's query will be coded as follows:

$$Q(y, z) \rightarrow Company(BinhHungHoa)$$

$$\wedge hasPlaceAt(BinhHungHoa, y)$$

$$\wedge\ applyMethod(BinhHungHoa, z)$$

The second query means searching where BinhHungHoa company is located with variable y and what the method applied with the variable z. The result from our ViWaT ontology will be y = "$TanPhu$" and z = "$BHH\_AerobicLake$". It means that BinhHungHoa company has used *Aerobic Lake* to treat the wastewater.



Axiom	4434
Logical axiom count	3277
Declaration axioms count	86
Class count	49
Object property count	29
Data property count	9
Individual count	724
Annotation Property count	3

Figure 6: The statistic of concepts, axioms and individuals.

In general, our ontology responds well to the management of information about wastewater treatment of companies. We have also stored more detailed information, i.e., the performance capacity of the company, the beginning year, the biological procedure used, and others (see Figure 6).

We end this paper with a discussion on query-answering in the presence of contradictory information. Indeed, the presence of "disjoin" constraints between concepts naturally raises the issue of managing conflicting information. There are a large number of methods in the literature that support query answering even if the knowledge base is inconsistent. Some methods use the notion of a "repair" which is a maximal consistent subset of the knowledge base. In general, an inconsistent knowledge base admits a large number of repairs, which makes the task of answering queries computationally difficult (e.g., [26]). This remains true even if we consider that the TBox is stable and reliable and only the assertions of the ABox can be questioned (we then talk of assertional repairs).

Depending on the used description language, some inconsistency management methods based on assertional repairs can be handled in a tractable way. We find for example the the well-known Intersection of ABox Repair (IAR) semantics [27], the so-called the grounded repair [28], or the so-called Elect method [29] defined for partially ordered DL-lite knowledge bases.

The Elect method is interesting because it selects a single consistent subbase of the ABox which contains the set of so-called accepted or elected assertions. An assertion f of an ABox $\mathcal{A}$ is said to be elected if for every conflict C involving it, there exists at least one other assertion g of C such that f is strictly more reliable than g. When all the assertions of the ABox $\mathcal{A}$ have the same level of reliability then the Elect method simply coincides with the one based on IAR-semantics. It is in this sense that the Elect method is considered as an extension of IAR-semantics.

Work in progress consists of exploring one of the extensions of the Elect method while maintaining its tractability. The idea is to propose an iterative version of Elect which consists in progressively removing from the ABox assertions which are inconsistent with the repair defined by Elect (wrt a TBox $\mathcal{T}$). Let us denote by Elect($\mathcal{B}$) the repair obtained using the Elect method on the set of assertions $\mathcal{B}$ (and with respect to a given TBox $\mathcal{T}$). At initial step (step 0), we simply let $\mathcal{A}_0 = \mathcal{A}$ (the initial Abox). At step i, if there exists an assertion f of $\mathcal{A}_i/Elect(\mathcal{A}_i)$ such that f is inconsistent with Elect($\mathcal{A}_i$) (and a TBox $\mathcal{T}$) then f is removed from $\mathcal{A}_i$. In this case, we let $\mathcal{A}_{i+i} = \mathcal{A}_i - \{f\}$ and we repeat again this step. The algorithm stops when at step n there is no assertion $f \in \mathcal{A}_n$ which is inconsistent with Elect($\mathcal{A}_n$) and the TBox $\mathcal{T}$. Elect($\mathcal{A}_n$) will then be the final repair of the initial ABox $\mathcal{A}$. This iterated process makes it possible to obtain a better repair than if a single application of Elect was carried out. Note that the computational overhead induced by this extension concerns additional consistency test checks, which are often performed in polynomial time in lightweight description logics.

6. Conlusion

We have laid the first foundations for the development of a lightweight ontology that meets the needs and actual situation in Vietnam to support the management of wastewater treatment of companies. Furthermore, we have provided an application based on ontology-based modeling for Vietnamese wastewater treatment (VWT) through description logic $\mathcal{EL}$ to aim at creating the firm foundation for building a semantic web of VWT by OWL2-EL. Using lightweight description logic for representation, reasoning, and querying is discussed in this paper.

The work presented in this paper is one of the initial steps in a long-term effort to create a universal and large repository for a mobile application and website for future works. Moreover, we will also pay attention to handling specific inconsistent cases in ViWaT.

Acknowledgments

Acknowledgements: This research has received support from the European Union's Horizon research and innovation programme under the MSCA-SE (Marie Skłodowska-Curie Actions Staff Exchange) grant agreement 101086252; Call: HORIZON-MSCA-2021-SE-01; Project title: STARWARS (STormwAteR and WastewAteR networkS heterogeneous data AI-driven management).

References

[1] I. V. Sevostianov, Y. V. Ivanchuk, O. V. Polishchuk, V. L. Lutsyk, K. V. Dobrovolska, S. Smailova, W. Wójcik, A. Kalizhanova, Development of the scheme of the installation for mechanical wastewater treatment, Journal of Ecological Engineering 22 (2021).

[2] C. Grandclément, I. Seyssiecq, A. Piram, P. Wong-Wah-Chung, G. Vanot, N. Tiliacos, N. Roche, P. Doumenq, From the conventional biological wastewater treatment to hybrid processes, the evaluation of organic micropollutant removal: a review, Water research 111 (2017) 297–317.

[3] H. Singer, S. Müller, C. Tixier, L. Pillonel, Triclosan: occurrence and fate of a widely used biocide in the aquatic environment: field measurements in wastewater treatment plants, surface waters, and lake sediments, Environmental science & technology 36 (2002) 4998–5004.

[4] L. Ceccaroni, U. Cortés, M. Sanchez-Marre, Ontowedss: augmenting environmental decision-support systems with ontologies, Environmental Modelling & Software 19 (2004) 785–797.

[5] F. Baader, D. Calvanese, D. McGuinness, P. Patel-Schneider, D. Nardi, The description logic handbook: Theory, implementation and applications, Cambridge university press, 2003.

[6] R. Mohamed, Z. Loukil, F. Gargouri, Z. Bouraoui, Evolution of prioritized el ontologies, in: International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, Springer, 2022, pp. 859–870.

[7] J. Mendez, B. Suntisrivaraporn, Reintroducing cel as an owl 2 el reasoner., in: Description Logics, 2009, pp. 1–11.

[8] S. Benferhat, C. Cayrol, D. Dubois, J. Lang, H. Prade, Inconsistency management and prioritized syntax-based entailment, in: IJCAI, volume 93, Citeseer, 1993, pp. 640–645.

[9] S. Belabbes, S. Benferhat, Computing a possibility theory repair for partially preordered inconsistent ontologies, IEEE Transactions on Fuzzy Systems 30 (2021) 3237–3246.

[10] S. Benferhat, Z. Bouraoui, T.-T. Ma, K. Tabia, Representing vietnamese traditional dances and handling inconsistent information, in: Information Processing and Management of Uncertainty in Knowledge-Based Systems: 19th International Conference, IPMU 2022, Milan, Italy, July 11–15, 2022, Proceedings, Part II, Springer, 2022, pp. 379–393.

[11] N. V. C. Le Hoang Viet, T. H. H. Ngan, N. V. Phu, Effectiveness of aquatic product processing wastewater treatment by submerged aerobic biological filter tank, Can Tho University Science Journal (2015) 94–101.

[12] T. T. Canh, T. C. Tan, N. Q. Nga, N. K. V. Truong, Study on wastewater treatment by biotechnology combined with usbf reverse filter, Science & Technology Development Journal, Vol. 9 (2006).

[13] H. N. Nguyen, T. T. Luong, Situation of wastewater treatment of natural rubber latex processing in the southeastern region, vietnam, Journal of Vietnamese Environment 2 (2012) 58–64.

[14] N. V. Anh, T. D. Ha, T. H. Nhue, U. Heinss, A. Morel, M. Moura, R. Schertenleib, Decentralized wastewater treatment-new concept and technologies for vietnamese conditions, in: 5th Specialised conference on small water and wastewater treatment systems, Istanbul, Turkey, 2002, pp. 24–26.

[15] M. Sabelfeld, L. Streckwall, B. Xuan-Thanh, S.-U. Geißen, Optimization potentials for wastewater treatment and energy savings in industrial zones in vietnam: Case studies, Case Studies in Chemical and Environmental Engineering 5 (2022) 100169.

[16] C. L. Grady Jr, G. T. Daigger, N. G. Love, C. D. Filipe, Biological wastewater treatment, CRC press, 2011.

[17] N. J. Horan, et al., Biological wastewater treatment systems: theory and operation., John Wiley & Sons Ltd., 1989.

[18] H. Vanhooren, J. Meirlaen, Y. Amerlinck, F. Claeys,

H. Vangheluwe, P. A. Vanrolleghem, West: modelling biological wastewater treatment, Journal of Hydroinformatics 5 (2003) 27–50.
[19] W. Reineke, M. Schlömann, Biological waste water treatment, in: Environmental Microbiology, Springer, 2023, pp. 467–491.
[20] A. Talaiekhozani, A review on different aerobic and anaerobic treatment methods in dairy industry wastewater, Goli A, Shamiri A, Khosroyar S, Talaiekhozani A, Sanaye R, Azizi K. A review on different aerobic and anaerobic treatment methods in dairy industry wastewater. Journal of Environmental Treatment Techniques 7 (2019) 113–41.
[21] G. Lettinga, J. Field, J. Van Lier, G. Zeeman, L. H. Pol, Advanced anaerobic wastewater treatment in the near future, Water Science and Technology 35 (1997) 5–12.
[22] J. Wang, K. Chon, X. Ren, Y. Kou, K.-J. Chae, Y. Piao, Effects of beneficial microorganisms on nutrient removal and excess sludge production in an anaerobic-anoxic/oxic (a2o) process for municipal wastewater treatment, Bioresource technology 281 (2019) 90–98.
[23] S. Mace, J. Mata-Alvarez, Utilization of sbr technology for wastewater treatment: an overview, Industrial & engineering chemistry research 41 (2002) 5539–5553.
[24] F. Baader, S. Brandt, C. Lutz, Pushing the $\mathcal{EL}$ Envelope, in: Proceedings of the 19th International Joint Conferences on Artificial Intelligence (IJCAI'05), 2005, p. 364–369.
[25] F. Kriegel, Most specific consequences in the description logic $\mathcal{EL}$, Discrete Applied Mathematics (2020) 172–204.
[26] M. Bienvenu, On the complexity of consistent query answering in the presence of simple ontologies, in: Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, 2012.
[27] D. Lembo, M. Lenzerini, R. Rosati, M. Ruzzi, D. F. Savo, Inconsistency-tolerant semantics for description logics, in: Web Reasoning and Rule Systemss, RR 2010, Bressanone/Brixen, Italy, 2010, pp. 103–117.
[28] M. Bienvenu, C. Bourgaux, Querying and repairing inconsistent prioritized knowledge bases: Complexity analysis and links with abstract argumentation, in: Principles of Knowledge Representation and Reasoning (KR), Virtual Event, 2020, pp. 141–151.
[29] S. Belabbes, S. Benferhat, J. Chomicki, Handling inconsistency in partially preordered ontologies: the Elect method, Journal of Logic and Computation 31 (2021) 1356–1388.