

## 습지 환경에서 수질관리를 위한 온톨로지 시스템 구축

지귀환<sup>o</sup>, 임소영\*, 김영민\*, 오염덕\*

<sup>o</sup>\*충주대학교 컴퓨터정보공학과, 교신저자

e-mail: {wlrmlghks}@gmail.com, {gray\_47, minkoon13, rdoh}@cjnu.ac.kr

# The construction of Ontology System for Water Quality Management based on Marsh Environment

Gwi Hwan Ji<sup>o</sup>, Soyoung Im\*, Youngmin Kim\*, Ryumduck Oh\*

<sup>o</sup>\*Chungju National University Computer Information Technology, corresponding author

### ● 요약 ●

최근 IT산업의 발전으로 인해 광범위한 정보가 존재하는 웹 기반 환경에서 원하는 정보의 검색과 검출을 위한 시멘틱 웹 (Semantic Web) 기반의 온톨로지에 대하여 연구가 다양하게 진행 중 이다. 본 논문에서는 온톨로지 기법을 이용하여 인간의 다양한 편의와 친환경적인 습지의 수질 관리에 대한 통합적이고 편리한 관리를 위해 WSN 기반의 습지 수질 관리에 특화된 온톨로지 설계 및 구축을 진행 하였다.

키워드: 상황인지, 온톨로지, 습지환경, 수질관리, WSN

## I. 서론

습지는 홍수조절과 사면 침식방지, 동물의 서식처 및 지하수 공급과 수질개선(침식에 의한), 지구 온난화 방지의 다양한 혜택을 제공하는 중요한 역할을 수행한다. 국외의 습지 관리보다 현저하게 낮은 국내 습지 환경 관리에서 특히, 세계 물 부족 국가로 인식 되어진 수질 관리[1]에 초점을 맞추어 최근 활발하게 진행중인 시멘틱 웹(Semantic Web)의 온톨로지를 통하여 습지환경 곳곳에서 수집되는 센서들의 데이터 관계를 정의하고 온톨로지로 설계하여 습지환경 수질 관리에 효율적인 상황인지 시스템의 구축 방법 기술이 필요하다.

일반적으로 직무에 관계된 전문가가 아니면 습지 환경의 수질 등급에 대한 특정한 변화 데이터나 용어에 관련한 의미를 파악하기 힘들다. 예를 들어 습지 내의 pH와 용존 산소량의 적절한 농도를 모를 경우 수집한 데이터가 어떤 의미의 데이터인지 중요성을 알 수 없을 것이고 수질 자체의 상태를 쉽게 파악할 수 없다. 따라서 기존의 방법으로 웹과 문서를 이용한 정보를 검색하여 그 데이터의 값이 어떤 상태인지 파악하려면 상당한 시간과 노력이 필요하다. 습지 환경의 특정 변화를 온톨로지 기반 환경으로 정의하면, 일반 사용자도 센싱된 데이터 중 의미 있는 데이터의 발견과 접근이 쉽고 보다 빠른 검색과 활용이 가능하다. 즉, 습지 수질에 대한 전문화된 지식이 없이도 전문가뿐만이 아닌 일반 사용자도 편리하

고 빠른 습지안의 수질 분석을 가능케 할 수 있다.

본 논문의 목적은 습지환경 수질등급과 온톨로지의 접목을 통하여 향후 시멘틱 웹의 발전과 더불어 습지환경 전반에 걸친 다양한 환경 관리를 통합적으로 운영하고 상황인지 기능을 적용할 수 있도록 구성하였다. 습지 환경의 요소 중 수질관리만을 다룬 온톨로지 구축에 초점을 맞추어 연구를 진행하였으며, 대상영역 설정은 수질등급과 상호 의미 관계에 대한 온톨로지 구축에 대하여 제안하였다.

## II. 관련 연구

시멘틱 웹을 위한 도구로 saltlux에서 개발한 온톨로지 라이브러리 시스템 COMET(온톨로지 수집 관리 툴)은 사용자가 간단하게 클라이언트 모듈을 다운받아 자신이 사용하는 온톨로지 모델링 툴에 플러그인 할 수 있도록 개발되었다.[2] 현재까지 많은 온톨로지가 개발되었고 지금도 개발이 진행 중이며 COMET과 같은 라이브러리 시스템으로 온톨로지의 수집과 관리가 이루어져 온톨로지의 필요성은 더욱 극대화 될 것으로 전망된다. 각계 각 분야에 걸친 온톨로지 구축과 더불어 이전까지 없었던 습지 수질관리 데이터의 온톨로지 구축은 매우 중요한 부분이다.

### 2.1 온톨로지

온톨로지란 어떤 관심 분야를 개념화하기 위해 명시적으로 정형화한 명세서로서 개념 구조의 명사와 관계에 대하여 규정하는 전산적 대상이라 정의할 수 있다.[3] 특정 영역을 표현하는 데이터

본 연구는 교육과학기술부와 한국산업기술진흥원의 지역혁신 인력양성사업으로 수행된 연구 결과임.

모델로서 영역(Domain)에 속하는 개념과, 개념 사이의 관계를 기술하는 정형(Formal) 어휘의 집합으로도 정의된다. 즉, 각 사물에서 공통점을 찾아내고 이를 하나의 집합 또는 범주로 나타내기 위해 의미, 지식의 쓰임새 등을 분명하고 자세하게 설명하는 것을 말한다.

시맨틱 웹의 핵심이라 할 수 있는 온톨로지의 구축은 이미 다양한 분야에서 설계가 이루어지고 있다.[4,5,6]

온톨로지는 크게 두 가지 접근법에 의해 생성되고 아래 [표1]에서 보듯이 사용자의 개인적인 차이는 있으나 본 논문에서는 보편적으로 사용하는 두 번째 방식인 명시적 온톨로지 구축 방법을 사용한다.

[표 1] 온톨로지 접근법

온톨로지 접근법		
	단어와 단어들간의 의미적 관계를 통계적 방법으로 발견하는 방법	전문가 인터뷰를 통한 명시적 온톨로지 구축 (에디터)
Tool	OntoLT, OntoLearn, Text2Onto	ezOWL, SWOOP, OilEd, Protege 2000
장점	사람의 직접적 개입이 없음	정확성, 편의성, 확장성
단점	정확성 부족	사람의 초반 개입 필요

## 2.2 Protege OWL (Ontology Tool)

프로티지는 국립의료도서관(National Library of Medicine), NSF(National Science Foundation), DARPA(Defense Advanced Research Projects Agency)의 후원을 받아 Stanford 의과대학의 의료정보학과(Stanford University School of Medicine, Stanford Medical Informatics)에서 지식 기반의 구조를 작성하기 위한 시스템으로 15년간의 연구 기간을 거쳐 개발되었다. 강력한 확장성을 갖춘 시맨틱 웹 언어를 사용한 개념모델링과 온톨로지 편집 및 구축을 위한 시각적 도구로서 오픈 소스 개발 환경이다.[7]

또한, 다양한 import를 지원하여 온톨로지의 시각화 및 데이터베이스와의 연동 등을 구현하여 현재 가장 널리 사용되는 시맨틱 웹 언어 온톨로지 개발 툴로서 의학계 및 한의학, 건축등의 온톨로지 개발에 사용되어, 본 논문의 온톨로지 구축에 사용하였다.

## III. 본 론

### 3.1 온톨로지 구성

온톨로지 구축에 있어서 Noy & McGuinness(2001)가 제안한 방법론(Ontology Development 101)에 근거하여 구축단계를 정해 다음의 각 단계를 거치게 된다.

- [단계 1] 온톨로지 도메인과 범위 결정
- [단계 2] 온톨로지 재사용 여부 결정
- [단계 3] 중요한 용어 목록 결정
- [단계 4] 클래스와 클래스의 계층 구조 결정
- [단계 5] 클래스슬롯과 속성 결정

[단계 6] 슬롯의 패시(facets) 결정

[단계 7] 인스턴스 생성

위의 단계 진행을 위해서 도메인의 특성과 온톨로지의 사용 용도를 고려한 온톨로지의 구성영역을 정한다. [그림1]은 본 논문에서 구축하려는 도메인과 서브클래스의 관계에 대한 간략한 모습이다.



[그림 1] Class

[표2]는 수질에 관련된 일련의 중요한 단어와 그 개념이다. 한국수자원공사에서 발췌한 것으로 [표2]에 언급되지 않은 많은 용어들이 존재한다. 이들 용어의 정확한 명세화를 위해선 해당 도메인 전문가들과의 협의에 의한 개념들과 그들의 관계들의 구조를 정하는 단계를 거쳐야 한다. 하지만 본 논문은 수질 등급의 추론으로 어중과 수질기준에 관한 BOD의 PPM 기준에 대한 속성과 그들과의 관계를 통하여 온톨로지 구축을 하였다.

[표 2] 수질 용어

수질 용어				
pH-수소이온 농도(hydrogen ion exponent)	DO-용존산소량(dissolved oxygen)	BOD-생물학적 산소 요구량(Biochemical Oxygen Demand)	COD-화학적 산소 요구량(Chemical Oxygen Demand)	ppm-100만분의 1(parts per million), 1L속에 존재하는 경우 mg수
중성-pH7, 열기성-pH7이상, 산성-pH7이하 등으로 1~14까지 구분	수중에 용해되어 있는 산소의 양을 말하는데 많은 물의 용존 산소량은 7~ 14ppm	하수, 폐수 중 함유되어 있는 유기물질이 수중의 호기성 미생물(박테리아, 플랑크톤)에 의해 분해될 때 요구되는 산소량	하수, 폐수 중에 함유되어 있는 유기물질을 화학적으로 산화시킬 때 소비되는 산소량	ppm-100만분의 1(parts per million), 1L속에 존재하는 경우 mg수

수질등급의 자료는 여러 조건들의 조합으로 전문화된 자료로서 일반인의 활용하기에 이해와 접근의 문제점을 내포하고 있다. 본 논문의 온톨로지 구축에서는 위의 [표2]의 속성과 [표3]에서 보여주는 수질기준의 57개 법정정수검사항목 법적기준치 또한 준수하여야 하므로 [그림1]에서와 같이 각 클래스와 클래스안의 인스턴스간의 관계 정립을 명확히 해야 한다.

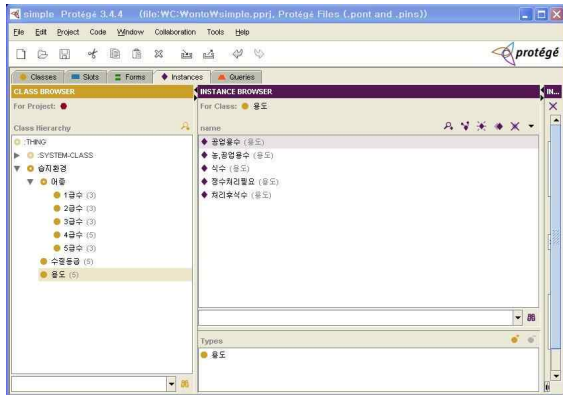
[표 3] 수질등급 법정기준

법정기준 법정정수검사항목(57개 항목) : 2009.1월부터 수질검사 법정기준 구분 항목 수질기준				
미생물에 관한 기준 (4항목)	건강상 유해영향 무기물질에 관한 기준 (11항목)	건강상 유해영향 유기물질에 관한 기준 (17항목)	소독제 및 소독부산물질에 관한 기준 (10항목)	심미적 영향물질에 관한 기준 (18항목)

### 3.2 프로티지를 이용한 온톨로지 구축

본 논문에서는 실험 대상영역을 수질환경으로 정하고 서버 클래스 안에 수질등급을 두어 위에 언급했던 표들의 속성과의 관계를 통하여 상황인지 온톨로지 구축을 하였다. 특히 위의 언급된 기존의 데이터 값은 자원의 미비함을 보완하기 위해 [그림1]에서 보았던 어종을 들으로써 위의 속성과 수질 기준의 관계를 정의하였다. 이로써 급수 제약조건의 관계를 정형화하여 수질등급을 바로 표출함과 동시에 식수 및 농업용수 등의 사용 가능성을 추론할 수 있도록 [그림2]와 같이 제한적인 부분만을 명시하여 프로티지를 이용한 온톨로지를 개발 하였다.

식수 수질 등급에 대한 한정된 도메인 내에서 수질관련 데이터의 신빙성있는 자원의 부족으로 어종과 몇 가지 조합만으로 온톨로지를 구축한 모습이다. 그러나 몇 가지 팩터(factor)를 추가한다면 [그림2]에서 추론 가능한 용도의 5가지 선택보다 더 다양하고 포괄적인 상황인지 온톨로지 구축이 가능하다.



[그림 2] 온톨로지 구축

[그림3]은 OWL로 표현된 온톨로지의 출력 화면이다.

온톨로지를 기반으로 하여 수질등급의 확인과 용어의 정의 및 인스턴스의 관계에 대한 용도의 특성을 보여주고 있으며, 프로티지 도구에서 적용한 query, from의 커스마이징 및 인스턴스와의 관계 그리고 슬롯 등의 각 기능을 정의하고 있다.



[그림 3] XML Coding

### IV. 결론 및 향후계획

시맨틱 기술을 이용한 수질 데이터의 온톨로지 구축이라는 점은 단순히 사람만이 인식할 수 있는 정보가 아닌 컴퓨터와 인간의 질의를 서로가 이해하는 자료의 지능화이다. 단순한 질의를 통해 그동안 단편적인 값 들 만들 도출하는 것이 아닌, 관련된 모든 자료 및 통계, 관계와 추론까지 해결 할 수 있는 기술로서 인간과 자연의 중요한 부분인 수질관련 질의를 정의할 수 있다.

수질과 관련된 많은 데이터의 제약이 있었지만, 향후 연구에서는 분석기를 통한 수질등급과 거의 근사한 결과 값을 도출할 수 있도록 더 많은 서브클래스와 관계를 통합할 예정이다. 이를 통해 단순한 분석기의 수질등급만이 아닌 시맨틱 웹의 장점인 관계를 통한 추론과 나아가 습지의 수질 내 넓은 범위 안에 뿌려진 무선 센서의 센싱 데이터 수집과 통계를 이용하여 이벤트의 검출, 정보 뿐만 아니라 조건부확률(Bayes Rule)을 이용한 데이터 마이닝 결과를 학습하게 한다. 그리고 이 학습 결과에 따른 추론 가능한 집단지성 온톨로지 구축을 통해 점차 분석기의 값에 근사한 값을 도출 가능토록 할 것이다. 또한, 이런 학습결과에 따른 추론 가능한 집단지성 온톨로지를 스마트폰과 같은 단말에 연동하여 유비쿼터스 환경에서 좀 더 간단하게 시맨틱 웹의 접근을 가능토록 하는 것이 이번 연구의 진일보적인 발전 방향이 될 것이다.

### 참고문헌

- [1] [http://www.kwater.or.kr/0678301250\\_43602.pdf](http://www.kwater.or.kr/0678301250_43602.pdf), 8010471250\_39398.pdf
- [2] <http://semanticwiki.saltlux.com/index.php/Comet> comet 코멧 (2008.08.13)
- [3] Mahesh Kavi. Ontology Development for Machine Translation: Ideology and Methodology. Technical Report MCCS-96-292. Computing Research Laboratory. New Mexico State University 1996.
- [4] 임수연, 박성배, 이상조, “의미관계 정볼르 이용한 약품 온톨로지의 구축과 활용”, 정보과학회 논문지, 소프트웨어 및 응용 제 32권 5호, 2004.5
- [5] 윤진혁, 유상봉, 김인한. “건축용어의 온톨로지 구축과 활용” 한국전자거래학회지, 8, 2, pp49~62. 2003
- [6] 탁문희, 김정화, 심준호. “상품 온톨로지 모델링 도구의 설계”. 한국컴퓨터종합학술대회(B). 2005
- [7] <http://protege.stanford.edu/wiki/>