

상수도시스템 수질사고의 전개양상을 고려한 비정형정보 의미분석 Semantic analysis of unstructured information considering the step in progress of water quality accidents in the water supply systems

홍성진*, 문기훈**, 양성훈***, 유도근****

Sungjin Hong, Gihoon Moon, Seong Hun Yang, Do Guen Yoo

.....
요 지

상수도시스템의 과정 중 최종 단계인 급수단계에서 지역전반에 수질문제가 발생할 경우, 직간접적인 피해의 해결은 장기간 지속될 수 있다. 본 연구에서는 실시간 비정형정보의 빅데이터 분석을 통해 상수도시스템에서 수질사고 문제의 파급력과 2차 피해 등의 연결 관계 변화 추적을 위한 기초적 분석을 수행하였다.

과거 대규모 수질사고가 발생한 바 있는 인천광역시 유충발생 사고를 대상으로 뉴스 기사 웹크롤링 절차를 정립하고, 그 결과를 분석하였다. ‘인천 유충’이 최초 보도되었던 2020년 7월 13일부터 이후 1년을 대상으로 네이버 통합검색에 의해 표출되는 뉴스 기사를 웹크롤링하였으며, 프로그래밍을 통한 불용어 제거 및 관련성 검토를 통해 총 920건의 기사를 분석하였다. 수질사고의 전개양상에 따라 사고발생, 확산, 수습, 그리고 보상의 4단계로 임의 구분하여 분석하였다. 의미분석을 위한 토픽모델링 기법은 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA) 방법을 적용하였으며, 긍부정 감정분석은 KNU 한국어 감정사전(KNU sentiment lexicon)을 활용하여 수행하였다. 토픽 모델링 결과, 사고 발생에서부터 확산, 수습, 보상의 단계에 맞춰 적절한 주제어의 조합에 따른 기사들이 도출되었으며, 단계별 긍부정 기사 비율역시 사고의 전개단계에 따라 적절히 나타남을 확인하였다. 제시된 수질사고 관련 비정형정보 분석 방법론과 결과는 과거 사고 사례 분석을 통한 검색 및 긍부정 키워드 확정, 키워드 발생 비율 변동(사고전과 후)에 따른 상황판단 기준 설정 등에 활용이 가능하다.

핵심용어 : 상수도시스템, 수질사고, 전개양상, 비정형정보, 의미분석

감사의 글

본 연구는 한국수자원공사(K-water)의 개방형혁신 R&D 사업(21-BT-001)의 일환으로 수행되었습니다.

* 정회원 · 수원대학교 공과대학 토목공학과 공학석사 · E-mail : yd1163124@naver.com
** 정회원 · 수원대학교 공과대학 토목공학과 석사과정 · E-mail : godurum3@naver.com
*** 비회원 · 수원대학교 공과대학 건설환경공학과 학부연구생 · E-mail : ch15316@naver.com
**** 정회원 · 수원대학교 공과대학 건설환경에너지공학부 조교수 · E-mail : dgyoo411@suwon.ac.kr