

Crawling한 오픈 데이터 기반의 오픈 소스 스마트 반응경로 탐색시스템

정준수, 박소현, 신동일[†]

명지대학교

(dpshin@gmail.com[†])

저장매체 기술의 발전에 따라 수많은 정보가 온라인상에는 등록되고, 축적되었다. 그중에서 연구자들은 실험 전 원하는 물질, 반응, 안전에 대한 정보를 필수적으로 찾아보며, 적절한 것을 선택하는데 시간이 오래 걸린다. SciFinder, Reaxys 등과 같은 유료 DB 또는 S/W를 사용한다면 원하는 정보를 빠르게 제공받을 수 있지만, 이러한 것들은 일반적으로 매우 비싸기 때문에 소규모 업체에서는 사용하기 힘들다. 이러한 단점을 해결하기 위해 오픈 데이터와 오픈 소스 기반의 반응경로 탐색시스템을 제안하였다. DB는 온라인상에 공개된 문헌이나 PubChem, ChemSpider 등의 기관의 공개정보를 crawling하여 자동으로 수집하였다. 제안 시스템에서는 이러한 오픈 데이터 기반의 DB를 이용하여 목표물질을 만들기 위한 반응경로들을 탐색한다. 적절한 탐색을 위해서 depth-first search 알고리즘과 pruning 알고리즘을 조합하여 탐색속도를 향상하였으며, 유해성 data를 통해 경로상 존재하는 물질들이 갖는 위험성을 분석한다. 생성된 결과는 목표물질까지 진행되는 다양한 경로들을 간단한 노드로 표현하며, 자세한 정보는 별도의 탭에서 나타난다. 이를 통해 연구자가 적절한 반응경로를 선택하는데 걸리는 시간을 줄이고, 경제성 높은 반응경로를 선정하며, 실험 시 안전사고를 예방하는데 도움이 될 것으로 기대된다.