

| 주제 : 연구데이터 리포지토리 구축 실무 |

데이터 리포지토리 GDR 구축 경험과 교훈 Lessons from Data Repository GDR (Geoscience Data Repository) Building Experience

한종규, 한국지질자원연구원 국토지질연구본부 지질연구센터, jghan@kigam.re.kr

JongGyu Han, Geological Research Division

100년의 역사를 지닌 한국지질자원연구원(KIGAM)은 국내 유일의 지질자원 전문연구기관으로서 그간 생산한 조사·연구데이터는 우리나라 과학기술의 귀중한 역사적·학술적 가치가 큰 유산으로써 보존 가치가 매우 크다고 할 수 있다. 하지만 현재 KIGAM의 상황은 최종성과물 위주로 자료관리가 이루어지고 있으며, 조사·연구 과정에서 생산된 암석·토양·지하수샘플이나 조사·탐사장비를 통해 얻어지는 자료는 연구자 또는 연구실·팀에서 개별적으로 관리하고 있다. 이러한 자료관리체계는 자료의 공동 활용이 어렵고, 자료를 보유하고 있는 연구자의 퇴직이나, 조직개편으로 인한 팀·실의 분리 과정에서 자료의 손실과 훼손 가능성이 높고, 누가 어디에 어떤 자료를 무슨 형태로 보관하고 있는지 찾기 어려워 자료의 재활용도가 떨어질 뿐만 아니라, 이로 인한 중복 조사·연구 가능성도 배제할 수 없다. KIGAM은 지질자원분야 국가데이터센터 구축을 목표로 연구과정에서 생산되는 연구데이터의 체계적인 관리와 공유, 활용체계 구축을 위해 2015년도에 기획사업을 통해 중장기 로드맵을 포함한 추진전략을 수립하였으며, 한국과학기술정보연구원(KISTI)의 DataNest를 기술이전받아 지질자원 연구데이터 리포지토리 시스템(GDR: Geoscience Data Repository)를 개발하였다. GDR 시스템 개발을 위해 연구데이터 분류코드를 작성하였으며, 2016년부터 데이터관리계획(DMP: Data Management Plan)을 주요사업 연구계획서 양식에 포함시켜 제출하도록 하였다. 과거 KIGAM은 연구데이터를 수집, 관리하기 위해 몇 차례에 걸쳐 시도를 했지만 실패한 경험을 가지고 있다. 실패 요인에는 ① 관련 정책, 제도, 조직, 인력, 예산 등 데이터 관리 인프라 부재, ② 연구사업에서 생산된 데이터는 개인소유라는 인식 및 공유 의식 부족, ③ 데이터 관리 활동은 귀찮은 것이고, 시간 낭비라는 인식, ④ 데이터 관리·공개·공유 활동에 대한 보상체계 부재 등을 꼽을 수 있다. 즉, 제도를 포함한 인프라 부족과 경영진과 구성원의 인식부족이 제일 큰 원인으로 판단된다. 성공적인 연구데이터 관리를 위해서는 지속적이고 꾸준한 투자가 이루어져야 하지만 경영진의 의지에 따라 사업이 중단되기도 한다. 이러한 과거의 실패 요인에 대한 해결 없이 지난 1년 6개월 정도의 GDR 운영은 지지부진하였다. 이러한 문제점을 해결하기 위해서는 국가차원의 제도적 뒷받침이 따라야 한다. 즉 국가 R&D 성과물 관리 차원에서 연구데이터를 주요 성과물로 관리해야 할 것으로 판단된다. 연구사업계획서에 DMP를 포함시키고, 연구주제 및 분야별로 데이터센터(혹은 데이터 리포지토리)를 지정하고, 국가 R&D에서 생산되는 연구데이터를 의무적으로 제출하도록 하는 것이다. 또한 데이터센터의 안정적이고 지속적인 운영을 위해 연구사업비 항목에 데이터 관리비를 신설하여 데이터센터의 운영비로 사용하도록 하면 예산문제도 어느 정도 해결 될 수 있을 것으로 본다. 또한 데이터 제출 및 인용도에 따라 데이터 생산부서 혹은 생산자에게 평가점수를 부여하는 등 보상체계 마련을 위한 연구도 필요할 것으로 보인다. 국가 R&D 연구데이터의 수집, 관리, 공유, 활용을 제대로 성공시키려면 국가 R&D 최고정책결정자의 지속적인 관심과 지원이 필수적이다.

주제어: 연구데이터, 데이터 리포지토리, 지질정보, 자원정보, 지질자원