

국문 요약

4차 산업혁명 시대에 데이터의 가치는 빠르게 증대되고, 더욱이 코로나바이러스의 확산은 전 세계의 디지털화를 가속했다. 이에 따라 데이터 분석 및 활용이 더욱 중요해지고, 데이터 활용 역량은 개인 및 기관의 핵심 경쟁력으로 대두되고 있다. 이러한 추세를 반영하여 KOICA는 데이터의 원활한 수집·관리·활용을 위해 프로세스 및 시스템 개선, 인프라 최신화로 디지털 KOICA로 도약하고자 디지털 전환 전략을 수립하였다. 현재 KOICA에서 수집·생산되는 다양한 자료에 비해 데이터 관리체계는 다소 미흡하다. 데이터를 통해 개발협력 사업 현황 및 성과를 분석하고, 향후 사업 방향성 및 전략 수립 등에 적극 활용하기 위해서는 동일한 표준과 규정으로 관리되는 양질의 데이터가 매우 중요하며, 이를 가능하게 하는 데이터 거버넌스 정립이 필수적이다. 이에 본 연구에서는 데이터 거버넌스의 정의 및 핵심 요소를 살펴보고, 나아가 KOICA 개발협력 사업의 데이터 거버넌스 수립을 위한 근거자료로 활용될 수 있도록 데이터 현황을 분석하고 향후 방향성을 제시하고자 한다.

데이터의 폭발적인 증가와 그 가치로 인해 많은 기관에서 데이터 거버넌스의 필요성에 대해 강조하고 있으나 데이터 거버넌스에 대한 합의된 정의는 없는 것으로 보인다. 다수의 선행연구를 살펴보면 데이터 거버넌스에 대한 정의는 상이하나, 양질의 데이터를 수집하고 관리하여 조직에서의 다양한 가치 창출에 지속적으로 활용되도록 하는 데 그 목적을 둔다는 점에서는 유사하다. 이에 본 연구에서는 데이터 거버넌스를 데이터의 생성부터 활용까지의 전체 관리 과정에서 필요한 전략, 정책, 기준, 역할, 책임 등에 관한 체계라고 정의하고자 한다.

데이터의 자산으로서의 가치가 급증하고 있다고 하나, 기관에서의 데이터 가치는 아직까지는 저평가되거나 이에 대한 체계적인 관리를 위한 투자로 이어지기까지는 조직 내 상당한 견해 차이가 있는 것으로 사료된다. 데이터의 생성부터 활용까지 전체 사슬의 참여자들 간의 원활한 소통을 위해, 그리고 데이터를 잘 활용하는 동시에 위험과 비용을 줄이기 위해 데이터 거버넌스를 수립할 필요가 있다. 다양한 경로를 통해 생성 및 수집되는 데이터를 수많은 사람들이 필요에 따라 활용하기까지 여러 제약조건들이 존재한다. 또한 데이터는 많으나, 정작 필요한 데이터는 찾기 어렵거나 가공에 상당한 시간이 요구되는 경우가 많다. 이는 데이터의 공급과 수요(혹은 생성과 활용)가 원활하게 연결되어 있지 않기 때문이라고 볼 수 있다. 결국 데이터 거버넌스는

데이터 사슬체계 안에서의 각각의 역할자 간의 연결고리 역할을 하는 프레임워크다. 데이터 거버넌스는 데이터를 전체적으로 관리하는 하나의 운영 체계인 바, 기관의 성격에 따라 사업전략과 목표에 부합하도록 구성되어야 하며, 이를 위해 어떤 데이터가 어디서 생성되어 누가 권한을 갖는지, 누가 어떤 방식으로 활용하는지 살펴보는 것이 필요하다. 데이터 거버넌스의 구축에 정해진 순서는 없다. 다만, 기관의 특성과 상황에 적합한 데이터 거버넌스를 정의하여 작은 부분부터 구축하여 확대해 나아갈 것을 제시한다.

데이터 거버넌스의 주요 요소로 표준화, 품질, 생애주기를 들 수 있다. 데이터 표준화는 다양하게 존재하는 데이터를 특정 기준에 따라 정리하는 것으로, 데이터의 활용 가능성을 높이기 위한 사용자들 간의 또는 기관 내 공통의 언어 및 규칙을 설정하는 과정이다. 이렇게 마련된 기준은 기존의 데이터를 재정비하는 과정에서도, 나아가 일관성 있는 데이터 품질 관리를 위해서 반드시 필요한 절차이기도 하다. 또한 데이터를 공유하는 범위가 넓어지면서 데이터의 잘못된 활용을 방지하고 올바른 활용 및 확산을 위해 양질의 데이터로 관리하는 것이 중요하다. 데이터의 품질은 정해진 기준에 따라 정확히 입력이 되어 있는지, 입력된 데이터 값의 의미를 제대로 전달할 수 있도록 관리되고 있는지를 검토하고 관리하는 과정이다. 이를 위해 세부적인 품질 기준을 마련하고 품질 관리 전담 조직 구성이 필요하며, 표준화를 포함하여 데이터 저장 시 고려사항, 품질 진단 및 개선, 피드백을 통한 오류 관리 및 활용에 대한 성과 평가 등의 관리가 요구된다. 한편, 생성되는 데이터는 무한히 수집관리하여 저장하고 보관하는 것은 불가능하다. 또한 개인정보 등의 민감함 데이터에 대해서는 일정 기한이 지나면 오히려 폐기하고 장기간 저장을 방지하는 관리가 필요하다. 따라서 기관의 데이터는 일련의 기준에 따라 저장 및 관리하고, 그 활용도가 다하면 폐기하거나 별도의 보관장소로의 이관이 필요하다. 이를 위해서는 데이터가 어떻게 생성되고 활용되는지, 더 이상 활용되지 않는 시점은 언제인지 등의 데이터 생성부터 폐기까지의 기준, 그리고 데이터의 성격에 따른 등급체계를 마련하여야 한다.

KOICA에서 주로 생성되는 데이터는 개발협력사업 수행을 통해 얻게 되는 자료, 사업수행과정의 기록, 사업결과 및 평가결과 등을 중심으로 기관 운영과정에서 발생하는 다양한 자료 및 정보라고 볼 수 있다. KOICA의 문서 중심의 업무 프로세스를 고려했을 때 비정형 데이터가 상당 부분 차지하며, 무엇보다도 이러한 데이터의 디지털화는 아직 미흡한 수준이다. 이상적인 데이터 관리를 위해서는 각 데이터의 생애주기 및 단계별 업무 프로세스를 분석하여 업무 및 데이터 관리 시스템 구축이 요구된다. 하지만, 대부분의 기관이 그렇듯, 각 업무 분야별 관리시스템은 각기 다른 시기에 개별 구축되어 결과적으로 여러 시스템에 분산 관리되고 있다.

현재 KOICA의 상당 부분의 업무영역은 시스템을 통해 관리되고 있으며, 각 업무는 다수의 시스템과 연계되어 진행되고 있다. 하지만 시스템의 연계가 제한적이고, 데이터의 디지털화 및 표준화가 미흡하여 데이터 활용에 제약이 발생한다. 이에 따라 KOICA의 핵심이기도 한 프로젝트 데이터를 중심으로 데이터 거버넌스를 구축을 위한 표준화 등에 대한 논의가 절실하다. 통합사업관리 시스템을 중심으로 관리되고 있는 프로젝트 단계별 다양한 문서 및 데이터를 살펴본 결과, 데이터는 상당하지만 효율적인 관리의 필요성을 드러냈다. 특히 수집 및 관리할 데이터에 대한 명확한 가이드라인이 미흡한 바, 각 담당자별 통일되지 않은 기준에 따라 입력 및 수집되고 있다고 볼 수 있다. 이는 활용관점에서는 가치 있는 데이터가 거의 없는 결과를 초래하게 된다. 각 단계별, 업무별 입력 데이터의 정의 및 입력방식에 대한 규칙을 수립하는 것부터, 품질 기준 수립 등 전사적으로 데이터 관리의 중요성을 인식하고 데이터의 디지털화를 위해 노력해가며 디지털환경을 구축해나가야 한다.

개발도상국이라는 특수한 환경 속에 언어 장벽 및 시스템의 한계를 극복하고 양질의 데이터로 관리하는 것은 훨씬 더 많은 노력이 요구될 것이다. KOICA의 사업을 통해 얻어지는 개발협력 데이터는 특수한 만큼, 이에 대한 체계적인 관리 또한 큰 의미가 있을 것이다. KOICA의 데이터는 내부적으로는 사업 방향성 및 전략 수립에 활용 가능하며, 대외적으로는 개발도상국으로의 접근성 확대 및 국제개발협력 분야의 다양한 정보를 공유할 수 있는 계기를 마련할 수 있다. 단기간에 기관의 데이터 거버넌스 수립은 불가능에 가깝다. 데이터 거버넌스 수립은 일회성 프로젝트가 아닌, 지속적으로 관리하며 유동적으로 개선해 나아가야 하는 프레임워크다. 또한 전사적 확대를 위해서는 충분한 자원과 인력의 투입이 반드시 뒤따라야 할 것이다. 따라서 단기적으로는 현재 KOICA의 국별협력 프로젝트 데이터의 표준화를 위한 기준을 마련하는 것에서부터 시작하여 품질 및 생애주기 관리에 대한 전사적인 합의 및 내재화가 필요하며, 장기적으로는 개발도상국에서의 디지털 인프라 구축까지도 고려되어야 할 것이다.