제2차 기록관리 현안과제 세미나 토론문(2019. 9. 27.)

‘디지털 시청각기록 장기보존정책 수립 전략’에 대한 토론문

이정남 상무((주)토피도)

1. 장기보존패키지를 대체한 블록체인 기반 기록관리 적용 시 고려사항

시청각기록의 진본성과 무결성을 확보하기 위해 물리적으로 하나의 파일 형태의 장기보존패키지로 패키징을 하게 되면 대상이 되는 기록을 메타정보와 물리적으로 하나의 파일로 보관하면서 기록관리의 요건을 충족하므로 독립적으로 완결된 기록의 형태를 가진다는 장점이 있습니다.

그러나 앞선 발표에서도 언급하였듯이 하나의 파일구조의 장기보존패키지는 패키징을 위해 새로운 파일을 생성하는 과정에서 시청각자료와 같이 콘텐츠의 용량이 큰 경우에 파일 I/O로 인한 부담으로 시스템에 과도한 부하를 야기합니다. 따라서 대안으로 제시한 기록 콘텐츠와 메타데이터의 분산 관리방안이 좋은 해결책이 될 수 있을 것이라고 보고 있습니다.

또한 컨텐츠와 분리하여 분산 관리된 메타데이터를 변경이 용이하여 기록된 정보에 대한 객관적 신뢰성을 확보받기 어려운 DBMS에 저장, 관리하는 것보다는 신뢰성 확보를 위해 블록체인 기술을 활용하여 블록체인 내에 기록함으로써 한번 기록된 정보에 대한 불변성(immutability)을 확보하는 것도 기록관리에 대한 객관적, 기술적 신뢰성을 부여해 주는 좋은 방안이라 하겠습니다.

이러한 흐름은 국내·외의 다양한 구축사례나 연구보고서를 통해서도 확인할 수 있습니다. 근래에 들어 점차 기관 자체의 신뢰를 바탕으로 기관에서 관리하는 기록의 신뢰성을 보장하고자 하는 방식에 대해서 국민들이 의혹을 제기하는 빈도가 증가하는 추세에 따라 영국의 TNA(The National Archives, 국가기록원)나 미국의 NARA(National Archives and Records Administration, 국립문서기록관리청)에서는 새로운 기술인 블록체인을 활용하여 기록관리에 기술적 신뢰를 부영하고자 하는 연구 및 PoC 프로젝트를 수행하고 있습니다.

국가기록원 역시 과기정통부의 공공기관 지원사업의 일환으로, 온-나라 문서시스템

에서 생산된 표준전자문서에 대해서 블록체인을 적용한 신뢰기록관리 플랫폼을 통해 생산 및 관리 현황을 단계별로 투명하게 추적하고 진본성, 무결성을 보장하고자 하는 시범사업을 '19년 4월부터 시작하여 올해 12월까지 수행하고 있습니다.

이 시범사업을 통해 전자문서가 생산시점부터 보존, 관리되는 모든 단계에 대해 감사이력이 투명하고 신뢰할 수 있는 방식으로 관리될 수 있음이 검증되고 나면, 시청각기록으로의 확대 적용이 더 큰 효과와 효율성 향상을 가져올 수 있다고 봅니다. 이는 앞서의 발표를 통해 납득할 수준으로 연구되었다고 생각합니다.

다만 블록체인을 적용함에 있어서 노드시스템에 대한 구성에 대해서는 조금 더 많은 고민이 필요하다고 봅니다. 블록체인에서 참여노드의 수는 저장되는 블록데이터에 대한 신뢰성을 확보하는데 중요한 역할을 수행합니다. 물론 국가기록원에서는 public 블록체인과 같이 임의의 노드를 참여시키지 않고 허가된 참여자들로만 노드로 참여하는 방식의 private 블록체인을 구성하고는 있으나, 이 역시도 노드의 수가 일정 규모 이상이 되지 않으면 국민들의 객관적 신뢰를 확보하기가 어렵습니다.

이를 해결하는 방법으로 참여하는 기록관리 기관을 다양하게 확보하는 것도 중요하고, 또 하나의 보완 방법으로 hybrid 형태의 블록체인 구성에 대해서도 검토해볼 수 있습니다.

Hybrid 형태의 블록체인은 Private과 public을 결합한 방식을 말합니다. 예를 들어, 기본적으로 시청각자료에 대한 메타데이터를 기록관리 기관들 간에 구성된 Private 블록체인에 등록하고, 일정 주기 단위로 각 블록의 마지막 블록에 대한 정보(블록에 대한 해쉬값 포함)를 public 블록체인에 등록하여 private의 작은 노드 수를 보완하는 방안도 있을 수 있습니다.

2. 이용가능성과 무결성을 동시에 확보할 수 있는 방안에 대한 고려

시청각자료의 이용가능성을 확보하기 위해 에뮬레이션 방식을 통해 원문 형태를 그대로 유지하는 경우에는 일반적으로 문서의 해쉬값(hash-value)을 이용하나, 마이그레이션 방식을 사용하는 경우에는 기존의 원문과 실 콘텐츠는 동일하나 해시값은 서로 다른 파일이 생기게 됩니다.

이렇게 되면 마이그레이션을 통해 생성된 파일에 대해서 진본성, 무결성을 검증하는 것이 어렵게 됩니다. 기존에 국가기록원에서는 장기보존패키지 방식을 통해 마이그레

이션(보존포맷변환과 같은)된 파일도 같이 패키징하여 보관하는 방식을 택한 것으로 알고 있습니다.

블록체인을 기반으로 메타데이터를 분산하는 경우에도 동일한 방식을 취할 수 있습니다. 마이그레이션이 발생될 때마다 마이그레이션된 기록에 대해서도 변환 절차에 대한 정보를 담은 메타데이터(언제, 어떤 포맷으로, 누구에 의해 등)와 해시값을 블록에 변환 단계마다 기록함으로써 마이그레이션된 기록과 원문과의 관계성이나 마이그레이션 기록의 무결성 등을 확인하는 방식을 취할 수도 있습니다.

그런데 영국의 TNA의 ARCHANGEL 프로젝트에서는 이에 대한 새로운 방법에 대해서도 연구하고 있습니다. “Tamper-proofing Video Archives using Temporal Content Hashes on the Blockchain”이라는 연구자료를 통해 sha-256과 같은 cryptographic hash가 아니라 Deep Neural Network를 활용하여 TCH(Temporal Content Hash)를 추출하고 이 값을 기준으로 문서의 진본성, 무결성 확보 방안으로 제시하고 있습니다. 다시 말해 블록체인에 시청각기록의 TCH와 TCH를 추출하기 위한 모델정보의 해시값(sha-256)을 기록하여 진본성, 무결성을 검증하겠다는 것입니다.

물론 이러한 연구에 대해서는 충분한 검증이 된 것이 아니기 때문에 좀 더 많은 검토가 필요하다고 생각합니다. 다만 일반 전자문서와 달리 시청각기록이 갖는 특성으로 인해 블록에 기록되어야 하는 정보들이 추가적으로 요구되어질 수 있고, 변경되어질 수 있으므로 이러한 흐름에 대해서는 지속적으로 국내외 기술동향을 검토하고 발전방향을 모색해야 할 것으로 봅니다,