

FAIR 지원을 위한 블록체인 기반의 학술정보공유시스템

김 선 겸^{1*} · 최 동 훈² · 서 태 설²^{1*}한국건설기술연구원 수석연구원²한국과학기술정보연구원 책임연구원

Blockchain-based Scholarly Information Sharing System to Support FAIR

Sun-Kyum Kim^{1*} · Dong-Hoon-choi² · Tae-Sul Seo²^{1*}Senior Researcher, Korea Institute of Civil Engineering and Building Technology, 283 Goyang-daero, Gyeonggi-do, Korea²Chief Researcher, Korea Institute of Science and Technology Information, 245 Daehak-ro, Daejeon, Korea

[요 약]

최근 기술의 발전과 오픈엑세스를 중심으로 한 학술 커뮤니케이션의 변화로 학술정보공유는 연구의 질을 높이고 나아가 국가적 차원의 정보의 확산과 활용에 기여하고 있다. 학술정보공유의 활용은 학술정보에 대한 검색성(findability), 접근성(accessibility), 상호운용성(interoperability) 및 재사용성(reusability)에 달려 있다. 이를 위해 유럽에서는 FAIR 가이드라인이 제정되었다. 블록체인은 분산 컴퓨팅 기반의 위변조 방지 기술로 하나의 거래 정보가 발생하면 블록체인 네트워크에 분산되어 있는 수많은 노드들에게 전파되기 때문에 거래자 간의 신뢰 문제를 해결하는 수단이 된다. 본 논문에서는 FAIR를 지원하기 위한 블록체인 기반의 학술정보공유시스템 모델을 제시하고, 하이퍼렛저 패브릭을 적용하여 구현한 프로토타입을 서술한다.

[Abstract]

Recently, scholarly information sharing contributes to nation-wide information distribution by changes in scholarly communication on advances in technology and open access. Exploiting the scholarly information depends on findability, accessibility, interoperability and reusability. For this purpose, FAIR principles has been established in Europe. A blockchain is a forgery prevention technology based on distributed computing. When a transaction is generated, it propagates to a large number of nodes dispersed in the blockchain network, thereby solving the trust problem between traders. In this paper, we present a blockchain-based scholarly information sharing model to support FAIR and describe its prototype implemented on Hyperledger Fabric.

색인어 : 학술정보공유, 오픈엑세스, 블록체인, 하이퍼렛저 패브릭, FAIR

Keyword : Scholarly information sharing, Open access, Blockchain, Hyperledger fabric, FAIR

<http://dx.doi.org/10.9728/dcs.2021.22.9.1493>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 20 August 2021; **Revised** 07 September 2021

Accepted 07 September 2021

***Corresponding Author; Sun-Kyum Kim**

Tel: +82-31-995-0962

E-mail: sunkyumkim@kict.re.kr

I. 서론

최근 컴퓨팅 기술의 진보와 오픈엑세스 [1] 및 리포지터리를 통한 학술정보 공유 인식의 변화를 통해 학술 커뮤니케이션의 변화가 이루어지고 있다. 이 변화는 연구 성과의 외부 확산과 소장하고 있는 학술정보 공유를 통해 연구의 질을 높이고 나아가 국가적 차원의 정보 유통에 기여한다. 따라서 학술정보공유시스템은 국가 연구 경쟁력 제고를 위한 핵심 인프라로 간주되고 있다 [2].

오픈엑세스의 리포지터리는 학술공유정보시스템 중에 하나로, 오픈엑세스를 지향하는 학술데이터를 기탁하거나 크게는 학술 출판의 과정도 포함하여 연구 성과 확산과 활용에 기여하고 있다. 학술정보의 활용은 학술정보의 검색성(findability), 접근성(accessibility), 상호운용성(interoperability) 및 재사용성(reusability)에 달려 있으며, 이를 위해 유럽에서는 FAIR(findability, accessibility, interoperability, reusability) 가이드라인 [3]을 제정하여 리포지터리 구현에 적용하고 있다.

블록체인 [4]은 암호화폐의 기저(underling) 기술로 알려졌지만, 그 장점으로 인해 다양한 분야에 적용되고 있다. 우리는 학술 커뮤니케이션에 대한 블록체인의 적용 가능성을 살펴보고, 정보의 활용성을 높이기 위해 FAIR를 지원하는 학술정보공유시스템을 제안하고자 한다. FAIR의 지원은 학술정보 기탁자, 리포지터리 관리자, 사용자 간의 검토 및 합의, 이에 대한 근거의 관리를 요구하며, 이것은 블록체인의 적용으로 실현 가능하다. 본 논문에서는 이러한 특성을 갖는 리포지터리의 실현을 위해 FAIR 리포지터리에 대한 블록체인 적용 모델을 제시하고, 이의 프로토타입의 구현을 서술한다.

프로토타입은 블록체인을 적용하여 데이터 정확성을 검증하고 FAIR를 평가하며, 사용자 권한 관리 등 원장에 기록할 내용에 대해 합의를 하고, 재사용자가 검색이 가능하게 한다. 본 논문은 이를 위하여 IBM과 리눅스 재단에서 주도하여 만든 프라이빗 블록체인인 하이퍼렛저 패브릭(Hyperledger Fabric) [5]을 이용하여 개발하였다.

본 논문의 구성은 다음과 같다. 2장은 블록체인과 관련 연구를 정리한다. 3장과 시스템 모델에 대한 설명과 4장에서는 프로토타입 구현 결과를 제시하며, 5장에서 결론으로 마무리 한다.

II. 관련 연구

2-1 블록체인

블록체인은 소규모 데이터들이 P2P(Peer to Peer) 방식으로 연결되어 있는 분산 컴퓨팅 기술 기반의 위변조 방지 기술이다. 블록체인의 필수 요소로 P2P 네트워크, 암호화, 장부, 합의가 있다. 하나의 거래 정보가 발생하면 이 거래 정보는 블록체인 네트워크에 분산되어 있는 수많은 노드들에게 전파되기 때문에 거래자 간의 신뢰 문제를 해결하는 수단이

된다. 이러한 과정 속에서 공증을 위한 합의 및 승인이 수행되며, 이 합의 및 승인에 의해 특정 자격을 갖는 중개자를 제거하는 것을 탈중앙화(decentralization)이라고 부른다. 따라서 P2P 네트워크 참여자는 처리 로직에 의한 거래 데이터의 타당성 확인, 공개키 인프라(public key infrastructure)에 의한 거래 데이터의 암호화, 합의 메커니즘에 의한 단일 블록체인 유지를 통해, 거래 데이터를 수정 불가의 공유 원장에 안전하게 저장하여 공유할 수 있다.

합의는 모든 참여자가 거래 데이터를 하나의 블록체인에 유지하기 위한 것으로, 새로 생성된 암호화 블록의 내용을 참여자가 검증하고 승인하는 과정을 말한다. 이를 위해 마지막 블록이 블록체인에 연결된 이후에 발생한 모든 거래 데이터를 하나의 블록으로 엮고 암호화하여 새로운 암호화 블록을 생성한다. 새로운 암호화 블록의 최초 생성자가 이것을 모든 참여자에게 방송하면, 수신자(최초 생성자 이외의 나머지 참여자)는 수신된 블록의 거래 데이터를 검증하고 거래 데이터의 순서에 문제가 없으면 이를 자신의 블록체인의 끝에 연결한다.

새로운 암호화 블록의 생성은 블록의 해쉬값(주소) 계산을 수반하며, 이는 상당한 양의 계산을 요구한다. P2P 네트워크 참여자는 누구나 새로운 암호화 블록의 해쉬값을 계산할 수 있지만, 이 중에서 최초로 계산한 참여자가 새로운 암호화 블록을 생성하고 나머지 참여자에게 방송하여 승인을 요청하게 된다. 이와 같은 합의 방식을 작업 증명(proof of work) [6]이라고 하며, 비트코인이나 이더리움은 이러한 합의 방식을 사용한다. 엄청난 양의 계산을 요구하는 작업 증명은 공유 원장으로서의 블록체인을 유지하는 전략이기도 하다. 최초 계산자에게 계산 노력을 암호 화폐로 보상하여 블록체인을 일관성 있게 유지한다. 이외에도 지분 증명(proof of stake) [7], 실용적 결함 허용 방식(practical Byzantine fault-tolerance) [8] 등 다양한 방식이 있으며 지속적으로 연구되고 있다.

블록체인은 P2P 네트워크의 모든 참여자가 동등한 권한을 갖고 합의에 참여하는 퍼블릭(public) 블록체인과 합의 참여자를 네트워크 참여자의 일부로 제한하는 프라이빗(private) 블록체인으로 나누어지며 표 1에 퍼블릭과 프라이빗 블록체인의 차이를 정리하였다.

퍼블릭 블록체인에서는 P2P 네트워크의 모든 참여자가 합의에 참여하며, 참여자 수가 증가할수록 합의에 걸리는 시간은 기하급수적으로 증가한다. 이러한 이유로 퍼블릭 블록체인은 거래 데이터의 실시간 처리에 현실적 장벽이 높다. 반면 프라이빗 블록체인은 합의 참여자를 네트워크 참여자 중의 일부로 제한하는 것으로, 합의에 걸리는 시간이 짧고 거래 데이터의 처리 및 관리에 필요한 기능적 요소를 포함하고 있어서 비즈니스 응용에 널리 사용되고 있다.

블록체인의 적용 기준은 탈중앙화 거래에 필요한 합의, 참여자간 수정 불가의 거래 데이터의 영구적 저장 및 공유 등, 블록체인 고유의 기능에 기인한다. 액션츄어가 제시한 블록체인 적용 기준은 거래 원장 공유 여부, 다자간 거래 여부, 거래자간 비신뢰 또는 이해 충돌 여부, 거래 합의의 필요성, 수정 불

가의 거래 기록 유지 필요성, 거래 규칙의 드문 변경 등이다. 최근 적용된 응용을 살펴보면, 이들 기준 요소가 모두 들어맞지는 않더라도, 높은 수준의 프라이버시 및 거버넌스를 요구하는 응용에 블록체인을 적용하기도 한다.

비즈니스에 대한 블록체인 응용은 탈중앙화에 의해 다자간 거래 데이터의 신뢰성을 높이고 참여자 간에 데이터 공유에 따른 비즈니스 비용 절감 및 효율성에 기여하고 있다. 탈중앙화는 높은 비용을 요구하는 기존의 중개자에 버금가는 신뢰도를 보장하는 저비용의 대안적 거래 환경을 생성할 수 있지만, 향후 신뢰 보장을 위해 나타날 수 있는 새로운 중개자의 등장을 방지할 수도 있다. 현재 비즈니스 응용에 대한 블록체인의 적용은 대부분 후자의 목적에 속한다고 볼 수 있고, 이런 추세가 지속되면 자연스럽게 전자의 목적으로 옮겨갈 수도 있을 것이다.

표 1. 퍼블릭 블록체인과 프라이빗 블록체인

Table 1. Public blockchain and private blockchain

Characteristics	Public blockchain	Private blockchain
Member	Anyone	Permissioned user
Main features	Free entry and exit	Faster and higher levels of complements
Consensus algorithm	PoS, PoW	PBFT
Blockchain platform	Bitcoin, Ethereum, Bitcoin cache, etc.	Linux funding, Nasdaq, Hyperledger fabric, etc.

2-2 학술정보공유를 위한 블록체인 적용 현황

학술정보공유에 대한 블록체인의 적용은 다양한 모델 개발을 통해 연구를 진행 중이다.

JP Tennant 등 [9]은 위변조 방지와 보상 제공을 통한 책임성 강화 및 동기 부여 측면에서 블록체인의 동료심사 품질 제고 가능성 제시하였다. M Avital [10]은 블록체인을 활용한 토큰 보상체계를 활용하여 심사 부족 문제 해결을 시도하였으며, Andrea Mambrini 등 [11]은 Referee의 심사품질에 따른 보상을 하는 OA 학술지 Peer Review 에코시스템에 블록체인과 분산저장기술(IPFS)을 적용하였다. K Janowicz 등[12]은 코인 경제 기반 모듈을 통해 블록체인 학술 출판 모듈을 제시하였으며, A Tenorio-Fornés 등 [13]은 탈중앙화 학술출판(Reviewer 평판의 분산 관리, 고유 OA 인프라, 투명한 거버넌스)을 실현을 위해서 블록체인과 분산저장기술(IPFS)을 적용하였다. TK Mackey 등[14]은 학술출판 프로세스의 투명성, 진실성, 확장성 제고를 위해 전통적 출판시스템과 괴리되지 않기 위해 연구되었으며, Z Jan 등[15]은 동료심사의 개방화, 분산화, 보상 반영을 반영하였다.

이러한 연구뿐만 아니라 새로운 학술 커뮤니케이션 패러다임으로써 동료평가와 FAIR 리포지터리 기반으로 블록체인 기술이 새롭게 조명될 필요가 있다. 블록체인은 모든 거래 기록이 저장되는 공유 인프라를 통해, 사용자에게 분산 데이터 관리를 허용한다. 학술 정보를 ‘연구자간 협력에 의해 생성, 변경, 활용 및 공유되는 논문 및 기저(underlying) 데이터로 구성된 하나의 거대하고 동적인 집합체’로 인식한다면, 그 자체가 블록체인 기술에 적합함을 알 수 있다. 블록체인 기반의 학술 커뮤니케이션에서는 연구자가 연구 결과를 생성하거나 연구자 간에 상호 작용할 때마다 그것이 공유 인프라에 저장된다. 블록체인 기반의 학술 커뮤니케이션의 장점은, P2P 네트워크상의 모든 참여자가 동일한 정보를 공유하기 때문에 학술정보의 신뢰성 보장, 소유권 및 추적성 보장, 다양한 척도에 의한 연구 업적 인정 등을 보호 받을 수 있다.

III. 적용 모델

학술정보공유체계의 이해 관계자는 연구자, 출판사(학회 포함), 유통자(정보센터, 도서관 등)이다. 연구자는 연구 프로세스에 따라 연구 결과물을 생산 및 소비하며, 협업 프로세스에 따라 학제간 및 기관간 공동연구를 수행한다. 출판사는 출판 프로세스에 따라 연구 결과물을 저널이나 단행본으로 출판하며, 출판된 학술정보는 후속 연구를 수행하는 연구자들에 의해 소비된다. 출판사는 자사의 저널에 실리는 논문의 품질을 관리하기 위해, 저명한 연구자들의 협력에 의한 동료심사를 거친다. 이러한 이유로 동료심사는 출판 프로세스에서 가장 많은 시간과 노력을 요구하는 서브 프로세스라고 인식되고 있다. 유통자는 출판되거나 출판 중인 학술정보를 수집하여 연구자들에게 제공한다. 과학 연구의 진실성 및 재현성을 위해 동료심사 과정에서 논문에 근거가 되는 데이터 및 데이터에 대한 서술을 보완 자료로 요구하거나 별도의 데이터 논문으로 출판하는 일이 증가하고 있다. 따라서 논문과 그 근거가 되는 데이터는 하나의 묶음으로 유통되는 것이 이상적이다. 이들 이해관계자들이 수행하는 업무 프로세스를 연구자 중심으로 통합 모델링한 학술정보공유체계는 그림 1과 같다.

유통(Distribution)의 전반적인 일(Demand survey, Plan, Collection, Management)를 하는 유통자(Distributor)와 출판(Publication) 업무(Manuscript recruitment, Editorial review, Peer review, Marketing, Editing and design)를 하는 출판사(Publisher), 논문의 저작(Problem definition, Related work, Hypotheses, Experimental design, Experiment, Data analysis, Evaluation of results, Paper work,)과 협력작업(Problem selection, Mock test, Issues and data discloser, Data analysis and submission, Evaluation of results and comparison, Paper co-work)를 하는 연구자(Researcher)들은 서로 유기적으로 연결되어 있으며, 연구자의 중심에 FAIR를 지원하는 FAIR 분산 오픈

리포지터리(FAIR DOR, FAIR Distributed Open Repository)를 두어 논문의 저작과 동료심사를 비롯한 모든 학술 출판의 과정을 변경 관리하며, 과정들이 투명하게 공개되고, 누구나 쉽게 접근할 수 있도록 하여 검색성, 접근성, 상호운용성, 재사용성을 높일 수 있다.

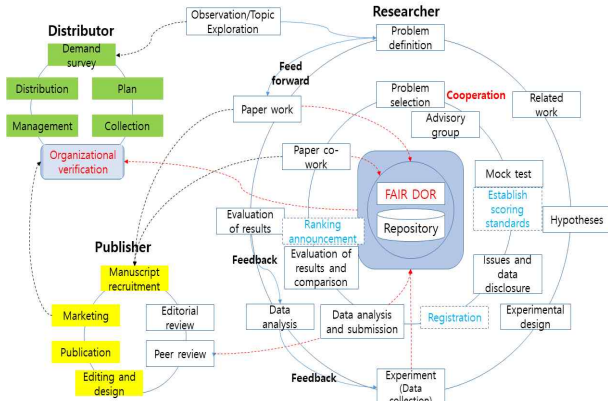


그림 1. 학술정보공유체계의 블록체인 적용점
 Fig. 1. Blockchain application point of academic information sharing system

IV. 프로토타입 구현

4-1 하이퍼렛저 패브릭

리눅스 재단에서 만든 하이퍼렛저(Hyperledger) 프로젝트는 기업에서 적용이 가능한 표준적인 블록체인의 기술을 구현하는 목적으로 만들어진 블록체인 컨소시엄이다. 산업간 블록체인 기술 발전을 위해 조직된 전 세계적인 오픈 소스 협업 활동을 한다. 현재 금융, बैं킹, 사물인터넷, 공급망, 제조 및 기술 분야의 총 130개 이상 회원이 하이퍼렛저 패브릭, 하이퍼렛저 컴포저 등 8개 프로젝트에 현재 참여 중이며, 표준화된 개방형 엔터프라이즈급 분산 원장 프레임워크 및 코드베이스 개발을 위해 노력하고 있다.

이 하이퍼렛저 프로젝트 중 대표적인 하이퍼렛저 패브릭은 모듈형 아키텍처 기반의 블록체인 애플리케이션 또는 솔루션 개발을 위한 근간으로 고안된 비즈니스 블록체인 프레임워크이다. 합의, 멤버십 서비스 등 구성요소의 플러그 앤 플레이를 지원한다. 또한 오픈 소스 모범 사례를 활용하며 비즈니스 환경의 기밀유지와 확장성을 가진다. 특히 가장 큰 특징으로 신뢰, 책임, 투명성을 확립한다. 허가형 네트워크로서 집단으로 정의된 멤버십과 접근권한을 비즈니스 네트워크에 제공하며, 거래의 비밀이 유지됨에 따라 비즈니스의 유연성과 보안성을 제공하여 정확한 암호화 키를 가지고 있는 당사자들에게만 거래가 보이도록 한다. 암호 화폐 사용하지 않기 때문에 거래를 확인하기 위해 채굴이나 값비싼 컴퓨팅을 요구하지 않는다. 이를 위해 스마트 계약의 탑재 논리를 활용하여 네트워크

상의 비즈니스 프로세스를 자동화한다.

하이퍼렛저 패브릭에는 여러 기업들이 프로덕션 비즈니스 네트워크 지원이 가능한 개방형 블록체인 패브릭의 구축을 위해 공급망, 자본 시장, 제조 및 의료 분야의 활용 사례에서 애플리케이션과 보안 블록체인 네트워크 간 테스트를 위해 협력하고 있다.

비즈니스에 대한 블록체인의 적용은 이더리움의 스마트 계약에서 시작되었으나, 비즈니스 응용에 최적화된 블록체인이 속속 등장하면서 새로운 방향으로 전개되고 있다. 하이퍼렛저 패브릭은 비즈니스 환경의 데이터 처리 및 관리에 필요한 각종 기능적 요소를 프레임워크로 제공하고 응용 개발 툴을 생산성 도구로 제공하고 있어서, 개발자의 호응을 얻고 있으며 다양한 비즈니스 응용을 위해 널리 보급되고 있다.

그림 2는 하이퍼렛저 패브릭의 프레임워크를 보여주고 있다. Membership, Blockchain, Transactions, chaincode의 기능으로 동작하고 있으며 각 기능별로 이벤트를 발생시키며 서비스가 된다. Membership은 사용자의 등록과 ID 관리를 담당하고 있으며, Blockchain과 Transaction은 실제로 블록체인 네트워크 상에서 블록체인이 동작하는 합의, P2P 관리, 분산원장 관리, 저장의 세부 기능을 수행한다. chaincode는 컨테이너 및 등록 관리를 통한 보안 관리를 하며 이 프레임워크를 수행하거나 작업하기 위해서는 APIs, SDKs, CLI를 활용해야 한다.

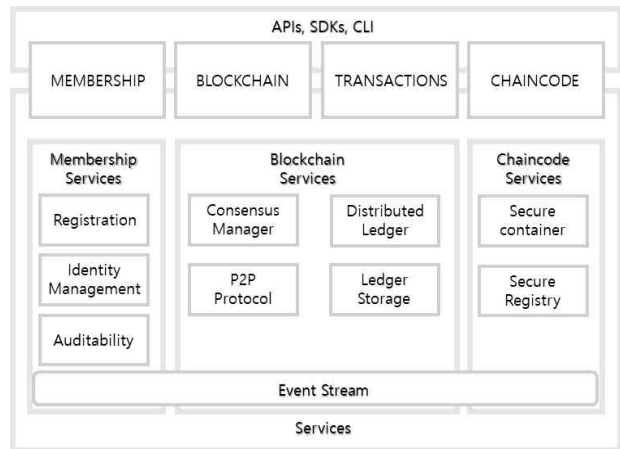


그림 2. 하이퍼렛저 패브릭 프레임워크
 Fig. 2. Hyperledger fabric framework

4-2 프로토타입 시나리오

본 연구는 학술정보공유시스템에 블록체인을 적용하기 위해서 프로토타입을 개발하였다. 기술적 검증을 위해 제안하는 모델은 FAIRness 및 신뢰성을 보장하기 위한 분산 오픈 리포지터리의 기본 기능이다. 하이퍼렛저 패브릭 플랫폼을 이용하며, 프라이빗 블록체인의 장점을 살려 개발하였다. 이 프로토타입은 개별 리포지터리의 Off-chain과 평가/검색이 가능한 블록체인 네트워크의 On-chain으로 구성된다. 블록체인 네트워크는 DOR 마스터를 두어 각각 개별적으로 생성된다. 업선된 전문가 또는 타 기관의 구성원을 평가위원회로 두고

블록체인의 평가(검증)시에 평가자로서 활용되며, 평가된 데이터들은 블록체인으로 연결되어 관리가 된다. 표 2의 프로토타입을 구성하는 사용자의 역할이다.

프로토타입의 사용자는 학술정보 기탁자(또는 출판사, Depositor), 리포지터리 관리자(DOR 마스터, DOR Master) 학술데이터(메타)의 FAIR를 평가하고 승인하는 평가위원회(Committee), 블록체인 검증자(Verifier), 그리고 등록된 학술정보를 검색 및 활용하는 재사용자(Reuser)로 구성된다.

그림 3과 프로토타입의 동작 시나리오이며, 그림 4은 그에 따른 On-chain과 Off-chain의 동작 흐름을 보여준다.

기탁자는 블록체인 네트워크에 데이터 평가를 위해 학술데이터의 엔트리를 On-chain에서 제출하여 작업을 시작한다. 기관이 원하면 해당 데이터는 공개가 되지 않으며, 필요에 따라 평가 시에 공개한다. 선정된 구성원들이 평가를 하며 평가에 통과한 데이터들은 블록이 되어 블록체인으로 연결되며 검증자1과 검증자2에 전달이 되며 각각의 개별 리포지터리에 저장된다. 재사용자는 이러한 블록체인에서 검색을 통해 데이터에 접근 가능하다. 리포지터리 등록 및 관리 모듈, 논문 등록 및 관리(Metadata 링크) 모듈, 메타데이터 관리(FAIR 테스트) 모듈, Front app(데이터 검색 및 메타 데이터 보기)으로 구성된 프로토타입을 개발하였으며, 메타데이터 등록시에, FAIR 검사를 해서 만족하지 않으면 거절, 만족하면 승인하며, app이 아닌 chaincode에서 이루어진다.

표 2. 프로토타입 사용자 역할

Table 2. Prototype user role

Role	Definition
Depositor	Data management, release (On-chain submission)
DOR(Distributed Open Repository) Master	Blockchain network configuration
Committee	Validation and FAIR evaluation of academic data
Verifier	Verification of blockchain data
Reuser	Information retrieval and utilization of academic data

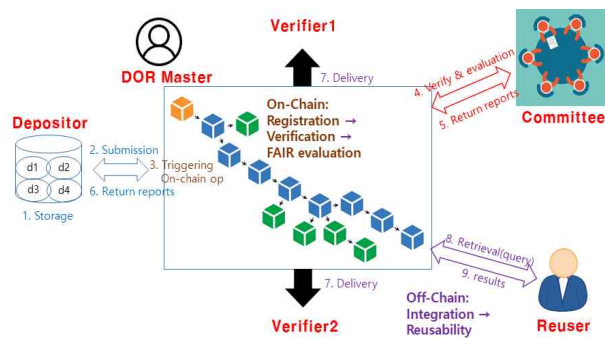


그림 3. 프로토타입 동작 시나리오
Fig. 3. Prototype operation scenarios

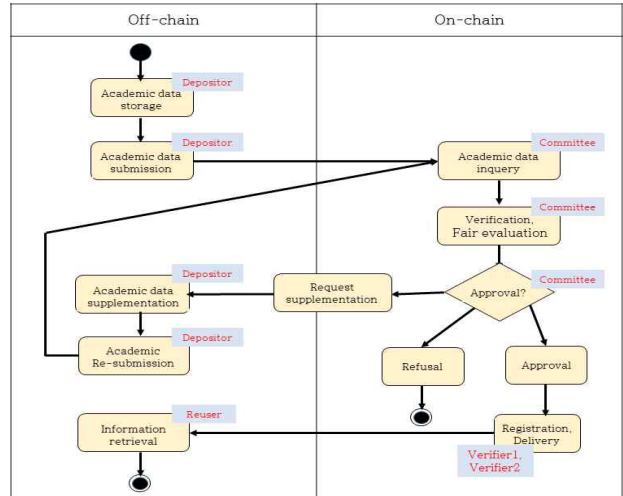


그림 4. on-chain, off-chain 동작 흐름
Fig. 4. on-chain, off-chain operation flow

4-3 프로토타입 모델링

프로토타입은 기본적으로 웹에서 실행할 수 있으며, 이를 위해서 DApp이라고 불리는 웹API에 그림 2의 기반한 블록체인 시스템 프레임워크에 얹는 방식으로 개발되었다. DApp은 블록체인 기반의 애플리케이션으로 스마트 계약을 통해 블록체인에 원장을 읽고 쓰는 작업을 할 수 있다. 그림 5는 제안하는 프로토타입의 동작 프로세스이다.

이 DApp은 API-Server, Commission, Publisher, Web으로 구성되어 있으며 API-server는 네트워크 애플리케이션 개발에 용이한 Node.js [16], 나머지는 부기능을 쉽게 다룰 수 있는 웹 애플리케이션 프레임워크인 Angular.js [17], 블록체인 트랜잭션의 동작을 위한 내부의 chaincode는 GoLang [18]으로 개발되었다. API-Server는 이 프로토타입의 서버 기능을 담당하고 Commission은 위원회에서의 합의 기능을 수행한다. Publisher는 데이터를 저장하고 블록체인에 트랜잭션을 생성하는 출판사이며, Web은 사용자가 검색을 할 수 있도록 한다. 블록체인 네트워크는 gRPC서버와의 연동과 해당 기능들을 하는 peer, Membership 및 Order 서비스로 구성된다. 이 gRPC 서버는 원격제어를 위한 코딩 없이 다른 주소 공간에서 함수나 프로시저를 실행할 수 있는 프로세스 간 통신을 담당하며 DApp에서의 chaincode를 통해 각 Peer에 동작을 지정할 수 있도록 한다. System chaincode는 블록체인 트랜잭션 동작이 아닌 블록체인 구성에 관한 정보를 가지고 있는 chaincode이며, chaincode는 블록체인 트랜잭션을 수행하는 peer이다. Endorsing peer는 위원회에서의 블록체인 생성을 위한 합의를 수행하며, Ledger peer는 블록체인의 일반동작을 수행한다. 합의의 경우 트랜잭션이 생성되었을 때 미리 선택된 Endorsing peer들은 해당 트랜잭션의 데이터를 검토하고 확인하며, 일정 인원수 이상의 Endorsing peer가 수락을 하면 블록이 생성된다. Membership service는 블록체인 네트워크 구성시 각 peer

에게 네트워크의 구성원이 될 수 있도록 권한을 부여하며, Order service의 경우 해당 트랜잭션이 블록체인으로 구성될 때 트랜잭션 발생 순서에 따라 정렬을 할 수 있도록 한다.

본 프로토타입의 네트워크는 Docker 기반의 피어 형태로 구성되며 검증노드는 Docker container에 chaincode를 배포하고 실행한다. 그림 6은 본 시스템이 프로그램상 실제로 동작하는 네트워크의 Docker container별 동작 구성도이며 표 3는 개별 사용자의 기능을 구성하는 Docker container이다. 모든 원안에 들어있는 기능들이 개별적으로 하나의 container에서 동작한다. Off-chain에서는 학술데이터 (scholar info)를 포함하는 생산자와, 검증자의 기능으로 구성되어, On-chain에서는 블록을 생성하기 위한 평가위원회, orderer, ca(인증서)의 합의주체와 합의를 위한 kafka와 zookeeper의 합의 엔진으로 이루어진다.

Off-chain의 상단에 위치한 생산자는 학술데이터를 레지스트리 형태로 메타데이터를 구성하여 on-chain에 제출하며 하단부에 위치한 검증자1과 검증자2는 각각 인증(ca.bc), 리포지터리(couchDB), 피어 동작(peer), 동작접근을 위한 api 로 구성되어 합의 후에 새로 연결되는 블록체인이 갱신이 되면 전달받아 개별 리포지터리에 저장한다. On-chain에서는 기탁자(Depositor)로부터 제출받은 학술데이터를 평가위원회(committee)를 통해 검증 및 FAIR 평가를 하며 Orderer와 ca를 통해 블록 내부의 트랜잭션 순서를 정렬한다. 이 때 합의된 트랜잭션 정렬을 위해 소수일 때는 kafka를, 다수일 경우에는 zookeeper 컨테이너가 동작하며 메시지를 큐잉하여 확장성을 높인다.

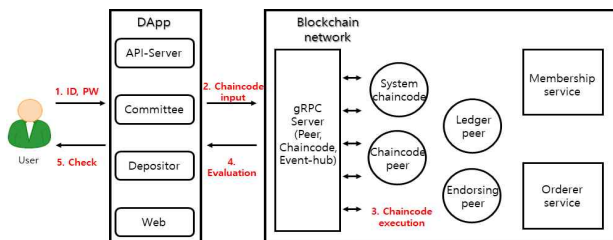


그림 5. 프로토타입 동작 프로세스
Fig. 5. Prototype operation process

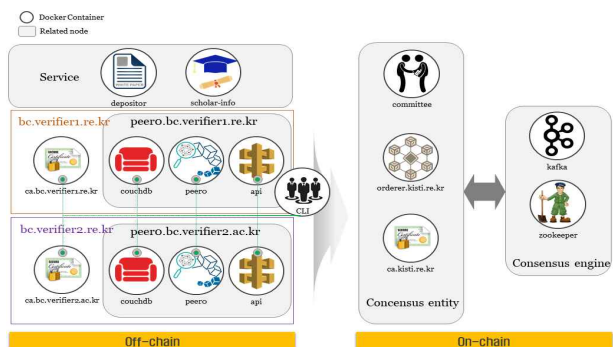


그림 6. 네트워크의 docker container별 동작 구성도
Fig. 6. Operation plot for each docker container in the network

표 3. 사용자별 Docker container
Table 3. Docker container by users

User	Docker container
Depositor	depositor, scholar-info
DOR Master	explorer, explorer-db, ca, orderer, kafka, zookeeper, cli
Committee	committee
Verifier	kisti_couchdb, kisti_ca, kisti_peer, kisti_api pnu_couchdb, pnu_ca, pnu_peer, pnu_api

4-4 프로토타입 동작

본 프로토타입은 그림 4의 시나리오에 따라 그림 7과 같이 동작한다. 1~6과정은 기탁자가 리포지터리에 학술데이터 및 메타데이터를 제출하는 과정이다. 블록체인과 DApp은 REST API서버를 통해서 통신을 수행하며, REST API 서버는 HTTP API기반의 JSON-RPC를 제공하고 하이퍼렛츠 패브릭과 통신하기 위하여 gRPC를 이용한다. 각 수행절차에서는 함수기반으로 호출 및 응답한다.

학술데이터가 등록되고 위원회에서는 논문 및 메타데이터를 확인하고 서명을 하는데, 아래의 그림에서 7~12번이 이에 해당된다. 기탁자가 학술데이터를 등록하는 과정과 평가위원회가 서명을 하는 과정은 모두 블록체인에 등록되며, 본 과정에서 Endorser peer는 chaincode를 시뮬레이션하여 트랜잭션을 승인 여부를 결정한다. API Server는 승인한 Endorser peer들의 서명을 받아서 트랜잭션을 생성하여 Orderer 제출한다. Orderer는 트랜잭션을 확인하고, 합의과정을 통하여 블록을 생성하는데, 본 PoC에서는 Kafka방식의 합의과정을 사용하였다. 최종적으로 생성된 블록은 해당 채널 내의 모든 피어들에게 전송된다.

마지막으로 Web에서는 재사용자(reuser)가 논문을 검색 (13~16번)하고, 특정 논문에 대한 서명(17~20번)을 확인한다.

본 프로토타입은 chaincode에 의해 트랜잭션이 처리된다. 이 chaincode에는 총 7개의 함수로 이루어져 있으며, 각 사용자(각 피어)의 역할로 이루어져 있다. 표 4는 본 프로토타입의 chaincode의 함수를 정의한다. Login 함수는 DOR 마스터와 평가위원회가 로그인을 하기 위한 함수이다. putpaper 및 putdata는 기탁자로 인해 학술데이터 및 메타데이터가 검색을 위해 web과 평가를 위해 평가위원회로 전달되어지는 함수이다. signpaper는 평가위원회에서 메타데이터를 평가 및 승인하는 함수이며, GetPaperContents를 통해 검증자에게 서명된 학술데이터가 전달이 된다. Search 함수는 재사용자가 Web에서 블록체인화된 학술정보를 검색하기 사용되며 GetPaperSign를 통해서 승인된 메타데이터를 확인할 수 있다.

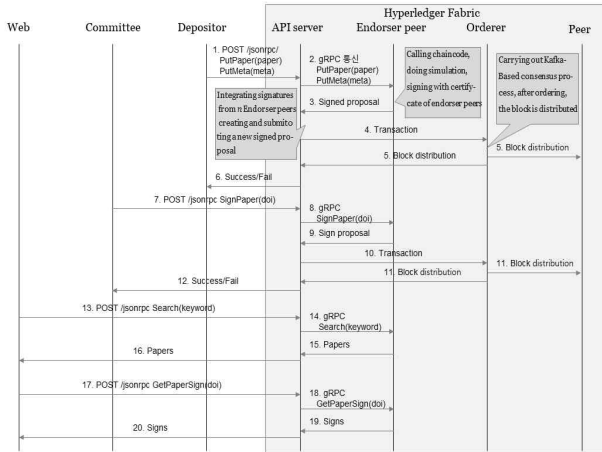


그림 7. 코드 동작
Fig. 7. How the codes work

표 4. 프로토타입 chaincode
Table 4. Prototype chaincode

Functions	Functions in scenarios
Login	DOR Master / Committee : Log in
putpaper	Depositor : Data submission to block chain network
putmeta	Depositor : Metadata submission to block chain network(committee)
signpaper	Committee : Sign and save
GetPaperContents	Verifier : Receiving signed data
Search	Reuser : Information retrieval
GerPaperSign	Resuer : Getting signed metadata in blockchain data

V. 결 론

본 연구는 기존 학술공유정보시스템의 문제점을 해결하기 위해 블록체인을 활용한 FAIRness가 보장되는 학술정보공유 프로토타입을 제안하였다. 기존 시스템에서 동료심사시 투명성이 보장이 되지 않고, 중복 제출과 데이터 소유권 문제가 있었다. 본 연구는 하이퍼렛저를 활용하여 학술 데이터를 위원회를 통해 엄격한 검증을 통해 데이터를 생성하고 이 시스템에 참여하는 기관들에게 모두 분산시킴으로써 동일한 정보를 공유하게 하여 투명성과 신뢰성을 높였다. 또한 사용자들은 이를 다시 검색하고 연구에 다시 활용할 수 있어 결과적으로 FAIR인 검색성, 접근성, 상호운용성, 재사용성을 높일 수 있었다.

본 연구는 프로토타입의 개발이기 때문에 향후에는 이를 보완하여 다수의 참여기관의 본격적인 참여를 통해 합의 알고리즘을 강화하고 인센티브를 제공하는 강화된 학술정보공유시스템을 개발할 예정이다.

감사의 글

본 연구는 2020년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 연구로서, 관계부처에 감사드립니다.

참고문헌

- [1] T.S Seo and H.Y Choi, "Global Dissemination of Domestic Scholarly Journals Using DOI and Open Access", *Journal of Information Management*, Vol. 42, No. 4, pp.1-21, Oct 2011, <https://doi.org/10.1633/JIM.2011.42.4.001>
- [2] J. Cho, "A Comparative Study of Academic Resource Sharing and Service System Between Korea and Japan", *Journal of Korean Library and Information Science Society*, Vol. 43, No. 4, pp.23-45, Dec 2012, <https://doi.org/10.1633/JIM.2011.42.4.001>
- [3] M. D. Wilkinson, M. Dumontier, I.J.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.W. Boiten, L.B. da Silva Santos, P.E. bourne, J. Bouwman, and A.J. Brookes, "FAIR Guiding Principles for scientific data management and stewardship", *Scientific Data*, Vol. 3, No. 1, March 2016, <https://doi.org/10.1038/sdata.2016.18>
- [4] M. SWAN, "Blockchain: Blueprint for a new economy", *O'Reilly Media Inc*, 2015.
- [5] E. Androulaki, A. Barger, V. Bortnikov, C. Cachin, vK. Christidis, A. De Caro, ... and J. Yellick, "Hyperledger fabric: a distributed operating system for permissioned blockchains", in *Proceedings of the thirteenth EuroSys conference*, Porto Portugal, pp. 1-15, 2018, <https://doi.org/10.1145/3190508.3190538>
- [6] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System", *Decentralized Business Review*, 21260, OCT 2008.
- [7] V. Buterin, "What Proof of Stake Is And Why It Matters", *Bitcoin Magazine*[Internet]. Available: <https://bitcoinmagazine.com/culture/what-proof-of-stake-is-and-why-it-matters-1377531463>.
- [8] M. Castro and B. Liskov, "Practical Byzantine Fault Tolerance", *Third Symposium on Operating Systems Design and Implementation*, Vol. 99, No. 1999, pp.173-186, Feb 1999, <https://dl.acm.org/doi/10.5555/296806.296824>
- [9] J.P. Tennant, J.M. Dugan, D. Graziotin, D.C. Jacques, F. Waldner, D. Mietchen et al, "A multi-disciplinary perspective on emergent and future innovations in peer review". *F1000Research*, 6:1151, 2017, <https://doi.org/10.12688/f1000research.12037.3>
- [10] M. Avital, "Peer Review: Toward a Blockchain-enabled

Market-based Ecosystem”, *Communications of the Association for Information Systems*, Vol. 42, No. 1, pp. 28, May 2018, <https://doi.org/10.17705/1CAIS.04228>

- [11] M. Andre, A. Baronchelli, M. Starnini, D. Marinazzo, and M. De Domenic, “PRINCIPIA: a Decentralized Peer-Review Ecosystem”, *arXiv preprint arXiv:2008.09011*, Aug 2020, <https://arxiv.org/abs/2008.09011>
- [12] K. Janowicz, B. Regalia, P. Hitzler, G. Mai, S. Delbecque, M. Frohlich et. al, “On the Prospects of Blockchain and Distributed Ledger Technologies for Open Science and Academic Publishing”, *Semantic web*, Vol. 9, No. 5, pp. 545-555, Aug. 2018, <https://doi.org/10.3233/SW-180322>
- [13] A. Tenorio-Fornes, V. Jacynycz, D. Liop-Vila, A. Sanchez-Ruiz, and S. Hassan, “Towards a Decentralized Process for Scientific Publication and Peer Review using Blockchain and IPFS”, in *Proceeding of the 52nd Hawaii International Conference on System Sciences*, Honolulu:HI, pp. 4635-4644, 2019, <https://doi.org/10.24251/HICSS.2019.560>
- [14] T.K. Mackey, N. Shah, K. Miyachi, J. Short, and K. Cluson, “A Framework Proposal for Blockchain-Based Scientific Publishing Using Shared Governance”, *Frontiers in Blockchain*, Vol. 2, pp. 19, Nov 2019, <https://doi.org/10.3389/fbloc.2019.00019>
- [15] Z. Jan, A. Third, M. Bachler, and J. Domingue, “Peer-reviews on the blockchain”, in *Proceeding of RefResh 2018: 1st Workshop on Reframing Research*, Cologne:Germany, 2018, <https://doi.org/10.3389/fbloc.2019.00016>
- [16] M. Satheesh, B.J. D'mello, and J. Krol, “Web development with MongoDB and NodeJs”, *Packt Publishing Ltd*, 2015.
- [17] B. Green and S. Seshadri, “AngularJS”, *O'Reilly Media Inc*, 2013.
- [18] A.A. Donovan and B.W. Kernighan, “The Go programming language”, *Addison-Wesley Professional*, 2015.



김선겸(Sun-Kyum Kim)

2012년 : 연세대학교 대학원 컴퓨터과
학과(공학석사)

2016년 : 연세대학교 대학원 컴퓨터과
학과(공학박사)

2016년 ~ 2017년: 한국건설기술연구원

2017년 ~ 2019년: 한국과학기술정보연구원

2019년 ~ 2020년: 차세대융합기술연구원

2020년 ~ 현 재: 한국건설기술연구원 미래스마트건설연구본부

※ 관심분야 : 모바일 컴퓨팅(Mobile Computing), 데이터 분석
(Data Analysis, 블록체인(Blockchain))



최동훈(Dong-Hun Choi)

1989년 : Northwestern University
EECS (공학박사.)

2005년 ~ 현 재: 한국과학기술정보연구원 책임연구원

※ 관심분야 : 데이터베이스(Database), 고성능 클라우드 컴퓨팅
(High Performance Cloud Computing)



서태설(Tae-Sul Seo)

2004년 : 한국과학기술원 자동화/설계
공학과 (공학박사)

1986년 ~ 현 재: 한국과학기술정보연구원 책임연구원

※ 관심분야 : 오픈엑세스(Open Access), 메타데이터
(Metadata)