JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 22, No. 3, pp. 483-490, Mar. 2021

신뢰가 보장된 데이터 경제를 위한 자기 주권 데이터 유통플랫폼 설계

김 근 형*
*동의대학교 게임공학전공 교수

A Design of Self-sovereign Data Distribution Platform for a Reliable Data Economy

Geun-Hyung Kim*
*Professor, Game Engineering Major, Dong-eui University, Busan 47340, Korea

[요 약]

코로나-19로 인한 사회적 거리두기 강화로 인해 경제와 사회의 모든 영역에서 비대면 디지털 전환이 가속화되고 있다. 포스트 코로나 시대에도 디지털 변혁이 일어나는 산업 영역은 계속 확장 될 것이다. 이에 따라 웹 기반 인터넷의 디지털 경제 규모도 커질 것이다. 디지털 데이터를 기반으로 한 디지털 경제 규모가 증가함에 따라 경제 참여자들 사이에서 신뢰를 확보하고 프라이버시를 보장해야 할 필요성이 증가하고 있다. 현재 사용자가 웹 기반 디지털 경제 생태계에서 데이터를 제어하지 않기 때문에 탈중앙 웹 연구가 진행 중이다. 본 논문에서는 최근 표준화 및 적용되고 있는 DID 기술을 이용하여 데이터의 자기 주권을 보장하는 플랫폼을 설계하였다. 이를 위해 DID (Decentralized Identifier) 기술과 SSI (Self-Sovereign Identity) 기술을 살펴보고 이를 바탕으로 사물과 데이터의 자기 주권을 보장하는 기술에 기반하여 데이터 유통플랫폼 참여자들 간의 신뢰와 프라이버시를 보장한다.

[Abstract]

Due to the reinforcement of social distancing caused by COVID-19, non-contact digital transformation is accelerating in all areas of the economy and society. Even in the post-COVID-19 era, the industrial area where digital transformation occurs will continue to expand. Accordingly, the scale of the digital economy on the web-based Internet will also increase. As the digital economy scale based on digital data increases, the need to secure trust and guarantee privacy among economic participants increases. Currently, research on the decentralized web is underway because users do not control their data in the web-based digital economy ecosystem. In this paper, we designed a platform that guarantees the self-sovereignty of data using DID technology, which has been standardized and applied recently. To this end, we looked at DID (Decentralized Identifier) technology and SSI (Self-Sovereign Identity) technology and based on these technologies to guarantee the self-sovereignty of objects and data, and the trust and privacy among participants in the data distribution platform are guaranteed.

색인어 : 가상물리소셜 생태계, 블록체인, 데이터 주권, 탈중앙 식별자, 자기주권 신원증명

Key word : Cyber-physical social ecosystem, Blockchain, Data sovereignty, Decentralized identifier, Self-sovereign identity

http://dx.doi.org/10.9728/dcs.2021.22.3.483

Received 24 January 2021; **Revised** 18 February 2021
Accepted 18 February 2021

***Corresponding Author: Geun-Hyung Kim**

Tel: +82-51-890-2271
E-mail: geunkim@deu.ac.kr

Ⅰ. 서 론

코로나-19로 인해 최근 1년 동안 우리의 삶에는 많은 변화가 발생하였다. 특히 사회적 거리두기의 강화로 경제사회 전 분야에서 비대면 디지털 전환이 가속화되면서 산업의 중심이 오프라인에서 온라인으로 옮겨지고 있다 [1]. 초연결, 비대면, 디지털 전환으로 대표되는 뉴 노멀 사회에서 디지털 데이터를 활용한 데이터 경제가 보편화되면서 경제 참여자 간 신뢰 확보의 중요성이 대두되고 있다. 또한 디지털 데이터 경제 활성화를 위해서 데이터에 대한 개인의 주권 및 프라이버시 문제 해결 및 투명성, 책임성, 개인의 권한 강화를 위한 기술 혁신이 필요하다 [2].

사회적 거리두기가 지속되면서 개인들은 인터넷에 연결된 스마트 디바이스를 활용하여 정보 검색, 음악 청취, 게임 하기, 동영상 시청 등의 비대면 활동을 꾸준히 하고 있다. 이러한 비대면 온라인 경제활동이 늘어나면서 대면 경제활동에서 요구되었던 개인의 신원 인증이 온라인 경제활동에서도 요구된다. 개념적으로 신원증명 (identity) 대상의 속성에는 이름, 나이, 성별, 주소와 같은 정보뿐 아니라 취미, 좋아하는 음식, 직업, 특정인과의 관계 등 개인과 관련된 모든 정보가 포함된다 [3]. 현재 실생활에서 사용 중인 신원증명 수단은 주민등록증, 운전면허증, 여권 등인데 이들은 이름, 주소, 나이와 같이 매우 한정된 정보만을 포함하고 있으며 신원증명 속성 중 필요한 속성만 선택하여 증명하는 것이 불가능하여 신원 증명 시 프라이버시가 보호되지 않는 문제가 있다. 또한 실생활에서는 국가별 신원증명 체계의 차이 때문에 한 나라의 운전면허증이 다른 나라에서 인정되지 않는다.

신원증명을 위한 구성요소는 신원증명 대상을 식별하기 위한 식별자 (identifier), 신원증명 대상의 특성인 속성 (attribute), 신원증명 대상의 속성에 대한 소유권을 인증하는 인증수단 (authentication method), 그리고 신원증명을 발급한 발급기관 (issuer) 4가지로 구성된다 [3]. 식별자는 신원증명 대상을 유일하게 식별하는데 사용되는 것으로 학생증의 학번, 주민등록증의 주민등록번호가 그 예이다. 속성은 신원증명 수단 (예: 학생증, 주민등록증 등)에 명시되는 것이며, 실생활의 신분증 체계에서는 신분증 소유자의 실물과 비교할 신분증 사진이 인증수단의 예이다.

현재 온라인 서비스에 가입한 개인의 신원증명 속성(개인 정보)은 서비스 제공자가 소유하고 이를 이용하여 이익을 얻고 있으나 개인 정보의 소유자들은 개인 정보가 언제, 어디서, 어떻게 사용되는지 알지 못한다. 또한 개인 정보가 서비스 제공자의 중앙 서버에서 관리되고 있어 유출 및 도용의 위험성이 높다. 이러한 문제점을 해결하고 신원증명 속성을 사용자 자신이 스스로 관리하고 통제할 수 있도록 신원정보를 저장하고 신원을 증명할 수 있는 탈중앙 식별자 (DID: Decentralized Identifier) [4] 개념이 제안되었다. DID는 개인, 기관, 사물, 디지털 자산 등 온라인에서 독자적으로 관리되며, PKI (Public Key Infrastructure)와 같은 암호화 기법으로 검증 가능한 식별자가 필요한 모든 애플리케이션에 유용하다. 예로 W3C의 검증 가능한 자격증명 (VC: Verifiable Credentials)과 DID를 연계하여 가상공간에서 개인, 기관, 사물, 디지털 자산 등을 식별하고 보안 프라이버시를 보장할 수 있다.

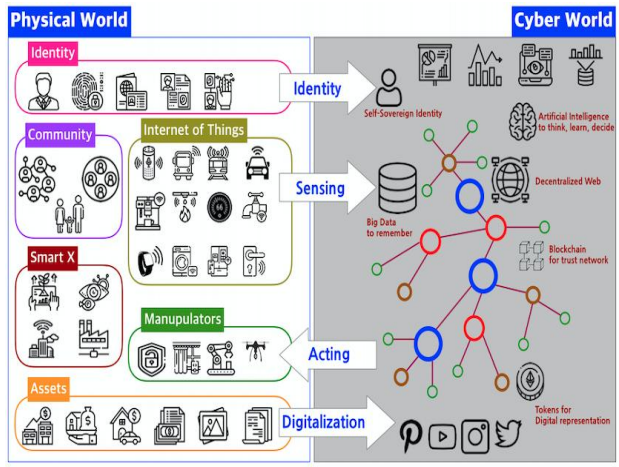

그림 1. 가상공간과 물리공간의 객체 간 상호작용
Fig. 1. Interaction of objects in both cyber space and physical space

그림 1은 가상물리소셜 (cyber-physical social) 생태계에서 물리공간의 물리 객체와 가상공간의 가상 객체 간 상호작용을 나타낸다. 스마트 홈, 스마트 시티, 스마트 헬스케어 등 사람이 중심이 되는 영역의 물리 공간 데이터가 디지털 전환되어 가상공간에서 관리되고 처리되며 처리 결과에 따라 물리 공간 객체를 제어, 통제 등 다양한 상호작용이 이루어진다 [5]. 이 과정 중 데이터 오남용을 방지하고 프라이버시를 보장하기 위해 개인 관련 데이터에 대한 자기 주권이 매우 중요하다.

DID 기술은 W3C에서 표준화 중이며 데이터 주권은 신원증명와 관련한 개인 신원정보의 보호에 집중되어 있을 뿐 가상물리소셜 생태계에서 개인이 생성한 데이터의 주권에 대해서는 아직 연구가 미흡한 상태이다. 본 논문에서는 최근 표준화와 기술 적용이 이루어지고 있는 DID 기술을 활용하여 데이터의 자기 주권을 보장하는 기술을 제안한다. 이를 위해 DID 기술과 자기 주권 신원증명(SSI: Self-Sovereign Identity)과의 관계를 살펴보고 이 기술을 활용하여 개인, 사물, 데이터들을 식별하고 이들 관계를 살펴보고 데이터의 자기 주권 보장 방안을 제시한다. 또한 이러한 데이터 주권을 보장하는 기술을 기반으로 데이터 유통플랫폼에서 참여자 간의 신뢰와 프라이버시를 보장하며 신뢰적인 디지털 자산을 위한 플랫폼을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서 본 논문의 데이터 주권을 보장에 필요한 기술인 탈중앙 식별자, 자기 주권 신원증명, 자기 주권 데이터, 탈중앙 웹 기술에 대해서 살펴보고, 3장에서는 탈 중앙형 데이터 유통플랫폼 구조를 설명한다. 플랫폼 구조를 제안하기 위해 먼저 데이터의 주권을 보장하고 디지털 자산의 신뢰적인 거래를 위한 플랫폼의 요구사항을 도출한다. 그 후 도출한 요구사항을 만족하는 플랫폼 구조를 설계한다. 마지막으로 4장에서 결론을 맺는다.

Ⅱ. 관련 기술

2-1 탈중앙 식별자 (Decentralized identifier)

DID는 글로벌하게 유일한 새로운 유형의 식별자이다. DID의 핵심은 사용자의 신원정보를 타인 또는 기관에 맡기지 않고 사용자가 스스로 자신의 신원정보를 관리하고 통제할 수 있는 것이며 신원을 증명할 때 필요한 정보 또는 신원정보를 가공하여 영지식 증명 (ZKP: Zero Knowledge Proofs) 이 가능한 결과를 전달하여 신원정보 노출을 최소화하여 프라이버시를 보호하는 것이다 [6]. DID는 신원정보를 개인이 소유한 단말에 저장하기 때문에 해커의 대상이 분산되고 개인 정보를 대량으로 탈취할 수 없어 해킹과 개인 정보 유출의 위험성에서 안전하다.

DID 개념이 소개되기 전에는 중앙등록기관이 필요 없는 탈중앙 식별자로 universally unique identifier (UUID)가 제안되었으나 인증을 위해 별도의 인증수단이 요구된다. 또한 식별자가 식별 대상에 지정된 후 변경할 필요가 없는 영구적인 식별자로 uniform resource name (URN)이 제안되었다. 그러나 URN은 중앙등록기관이 필요하며 신원을 증명할 때 암호화 기술이 적용되지 않는 점이 DID와 다르다. DID 기반 인증과정에서 인증수단으로는 DID 문서 (DID document)가 정의되어 사용된다. DID 문서는 DID 소유권을 증명할 수 있는 인증수단으로 DID 주체 (subject), DID 주체 또는 DID 대리인이 자신을 인증하고 DID와의 연관성 증명에 사용될 공개키 또는 익명의 생체인식과 같은 메커니즘, DID 주체의 속성 등을 포함한다 [7].

DID 개념을 구성하는 주요 객체는 DID 주체 (subject), 발급기관 (issuer), 검증기관 (verifier), DID 저장소 (repository) 로 구성된다. 또한 SSI의 기술적 요구사항에 적합한 분산저장소 (DID 저장소)는 블록체인으로 대부분의 SSI 플랫폼에서 DID 저장소로 블록체인 기술의 활용을 검토하고 있다. DID는 DID 주체에 대한 식별자의 역할과 DID 저장소에 저장된 DID 문서를 참조할 수 있는 unified resource identifier (URI) 역할을 한다. DID는 자원에 접근할 때 사용하는 프로토콜을 규정하는 URI 체계 (scheme), DID 문서가 저장된 저장소를 규정하는 DID 방법 (DID method), DID 방법이 규정한 저장소에 DID 문서가 저장된 주소를 규정하는 방법에 특화된 식별자 (method-specific identifier)로 구성된다. 그림 2는 DID 아키텍처를 구성하는 기본 요소로 각 요소의 설명은 표 1과 같다.

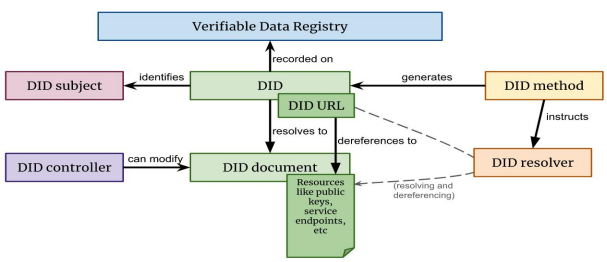

그림 2. DID 아키텍처의 기본 요소
Fig. 2. The basic elements of DID architecture

표 1. DID 아키텍처의 요소에 대한 설명
Table 1. The description of elements in DID architecture

Elements	Description
DID Subject	The entity identified by a DID and described by a DID document
DID Controller	An entity that has the capability to make changes to a DID document
Verifiable Data Registry	A system that facilitates the creation, verification, updating and/or deactivation of DID and DID documents
DID method	A definition of how a specific DID scheme must be implemented to work with a specific verifiable data registry
DID Resolver	A software and/or hardware component that performs the DID resolution function

2-2 자기 주권 신원증명 (Self-sovereign Identity)

본질적으로 SSI 개념은 개인이 자신의 디지털 신원을 서비스 제공자 또는 제 3의 기관이 아닌 자신이 완전히 소유하고 관리 권한을 가지며 개인의 프라이버시를 보호하면서 신뢰를 활성화하는 것이다. SSI는 사용자와 서비스를 분리하기 위해 DID와 블록체인을 사용하며 신원 인증을 위해 제출하는 신원증명에 새로운 속성을 추가하거나 민감한 속성은 제거하여 사용할 수 있다.

디지털 생태계의 SSI는 그림 3과 같이 식별자와 신원증명으로 구성된다. 식별자는 디지털 생태계에서 존재를 표현하는 데이터이고 신원증명은 주체의 신원 증명에 사용되는 정보로 속성과 값의 형태를 가진다. SSI 절차에 사용되는 데이터는 DID, DID 문서, 검증 가능한 자격증명 (VC: Verifiable Credentials), 검증 가능한 표현 (VP: Verifiable Presentations) 으로 구성된다. 자격증명은 일상생활의 일부로 사실 또는 자격을 증명 또는 주장 (claim)하기 위한 데이터로 운전 면허증, 의사 면허증, 졸업장, 여권, 사원증 등이 있다. 자격증명은 발급자가 만든 하나 이상의 주장들로 구성된 집합이다 [8]. 신원증명을 위한 데이터로는 VC와 VP 외에 개인 정보와 주장이 포함된다. 주장은 증명하려는 주체의 신원증명 속성에 대한 정보로 주체-속성-값의 방식으로 표현된다. 이는 해당 주체가 어떤 속성 (property)을 가지며 관련 속성의 값 (value)이 무엇인지 표현한다 [8].

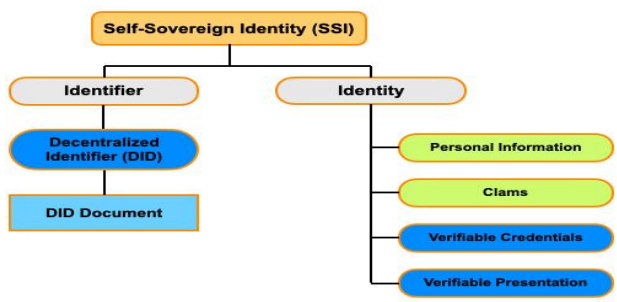

그림 3. 자기 주권 신원증명의 구성 컴포넌트
Fig. 3. The Components in Self-sovereign Identity

크리스토퍼 알렌은 SSI 구현 시스템의 10개 의 자기 주권 신원증명 원칙을 주장했다 [9]. 소브린 재단 (Sovrin Foundation) 백서는 이 원칙을 표 2와 같이 보안 (security), 제어 가능성 (controllability), 이식성 (portability)으로 분류하였다 [10]. 신원정보의 보안 측면으로 주체의 권리보호 (protection), 신원정보의 지속성 (persistence), 주장의 공개 최소화 (minimization), 자기 데이터를 액세스 제어 가능성 측면에서 주체의 독립성 (existence), 자기 주권 신원정보의 통제 (control), 주체의 동의 (consent)에 따른 신원정보 활용이 요구된다. 그리고 신원정보 관련 서비스가 특정 서비스 제공자에 종속되지 않고 원하는 모든 상황에서 신원정보를 사용할 수 있는 이식성 (portability) 측면에서 넓은 신원정보의 사용 가능성 (interoperability), 기반 시스템과 알고리즘의 투명한 공개 (transparency), 자신 데이터에 대한 주체의 접근 (access) 이 보장되어야 한다.

[11]에서 제시한 SSI에 대한 원칙은 신원증명을 중심으로 한 원칙이며 소브린 재단은 생태계 관점에서 12가지의 SSI의 원칙을 제시하였다 [12]. 12가지 SSI 원칙 중에 SSI 생태계 내의 신원증명 보유자가 배제되거나 소외되지 않아야 한다는 공정과 포괄 (equity and inclusion) 원칙, 디지털 신원정보의 프라이버시 보호와 최소의 디지털 신원정보 공개를 요구하는 프라이버시와 최소공개 (privacy and minimal disclosure) 원칙, 신원증명 보유자의 참여를 강제할 수 없다는 참여 (participation) 원칙은 생태계의 거버넌스 관점에서 검토가 되어야 할 원칙이다.

표 2. 자기 주권 신원증명의 원칙
Table 2. The Principles of Self-sovereign Identity

Security	Controllability	Portability
Protection	Existence	Interoperability
Persistence	Control	Transparency
Minimization	Consent	Access

검증 가능한 자격증명 데이터 모델 관점에서 주요 컴포넌트의 역할과 정보 흐름을 그림 4와 같다. 주요 컴포넌트는 VC를 발급하는 발급기관 (issuer), 하나 이상의 검증 가능한 자격증명을 보유하고 그것을 통해서 VP를 생성하는 신원증명 보유자 (holder), 사용자로부터 전달받은 VP를 기반으로 VP의 진위 여부와 사용자가 적절한 VC를 보유하고 있는지 검증하는 검증기관 (verifier), DID 및 DID 관련 정보를 저장하는 블록체인 기반 검증 가능 데이터 저장소 (verifiable data registry), 그리고 검증의 대상인 주체가 있다. 보통 보유자가 검증 대상이지만 다른 경우가 있다. 예로 부모 (holder)가 자식 (subject)의 신원을 증명하는 경우와 건물주 (holder)가 부동산 (subject)를 증명하는 경우 보유자와 검증 대상인 주체가 다르다. 발급기관의 예로는 정부, 회사, 기관, 단체, 개인 등이며 보유자의 예는 학생, 직원, 고객 등이다. 사람, 동물, 사물, 데이터 등 사이버 공간에서 식별자가 필요한 모든 대상이 주체의 대상이 된다. 검증기관의 예로는 웹 사이트, 고용주, 보안 담당자 등이며 검증 가능 데이터 저장소의 예로는 정부 ID DB, 분산원장, 탈중앙 DB 가 있다.

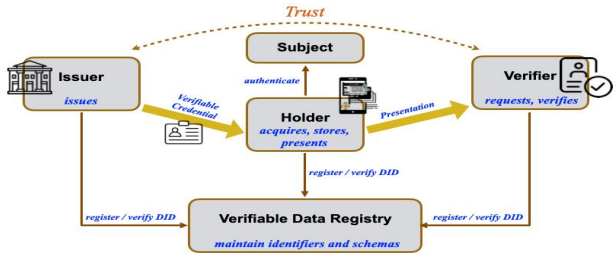

그림 4. 검증 가능 자격증명 데이터 모델의 컴포넌트의 역할과 데이터 흐름
Fig. 4. The role and data flow of components in verified credential data model

그림 4의 검증 가능 자격증명 데이터 모델에서 인증요청 및 인증을 위한 데이터 흐름 예는 다음과 같다.

(1) 보유자는 검증기관에서 주체를 인증하기 위해 해당 내용을 인증할 수 있는 자격증명의 발급기관에 VC 발급을 요청한다.

(2) 발급기관은 보유자 권한의 적법성 권한을 확인 후 보유자의 VC에 서명한 후 발급하고 데이터 저장소에 등록한다. 발급된 VC에는 발급기관과 검증할 주체의 DID가 명시된다.

(3) 보유자는 VC를 취득하고 발급기관이 서명한 내용을 검증한 후 검증기관에 제출할 인증정보를 VC에서 보유자, 인증정보를 추출한 VC 발급기관, VC 주체의 DID를 명시하여 VP를 생성한다. 보유자는 생성된 VP에 서명 후 검증기관에 제출한다.

(4) VP를 수신한 검증기관은 VP가 올바른 보유자의 것인지를 보유자의 DID를 통해 확인하고 VP에 포함된 VC 속성이 검증할 주체의 것인지 VC 내에 명시된 주체의 DID를 통해 확인한다. VP에 포함된 VC의 속성 발급기관은 기관의 DID를 통해 확인한다. VP의 VC 주체의 DID에 포함된 서명은 DID 문서를 통해 검증한다. 검증기관은 검증 가능 데이터 저장소를 통해 해당 내용을 확인하고 내용에 문제가 없으면 검증을 완료한다.

다음 그림 5는 VP, VC, 주장(claim)으로 구성된 검증 가능한 자격증명 데이터 모델의 개념이다. 먼저 주체의 신원정보를 표현하는 주장이 모여서 자격증명이 되고 검증 (proof)이 이루어지면 VC가 만들어진다. VC의 주장 중 일부를 조합하여 검증기관에 제출할 인증정보 (presentation)를 만든 후 검증 가능한 VP가 생성된다. 검증기관에 제출되는 인증정보에는 주체의 신원을 검증하고자 하는 내용이 포함된다. 인증정보는 발급기관이 발급한 자격증명을 참조하여 자격을 증명할 유형, 증명할 주체, 신원 내용인 주장, 발급자, 발급 일자 정보를 추출한다. 추출한 자격증명 정보의 검증은 VC 내 검증 (proof) 필드에 정의된 검증방식에 따라 이루어진다 [8].

그림 6의 예는 주민등록증과 졸업증명서 두 개의 VC의 속성 중 18세 이상이며 학위에 대한 신원정보를 생성하여 VP를 만드는 것이다. 즉 발급받은 VC로부터 인증할 신원정보에 따라 VP를 만들기 때문에 VC는 보관용 신원증명, VP는 제출용 신원증명이라 한다. 사용자는 DID, VC, VP 기술을 활용하여 중앙인증기관의 개입 없이 비대면 본인확인 및 인증요청을 처리할 수 있다.

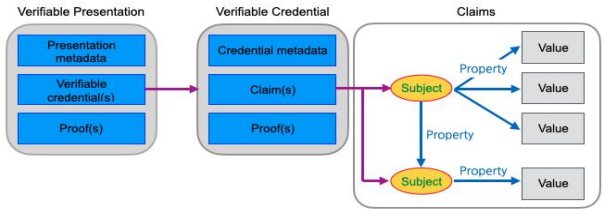

그림 5. 검증 가능 자격증명 데이터 모델 개념
Fig. 5. The concept of verified credential data model

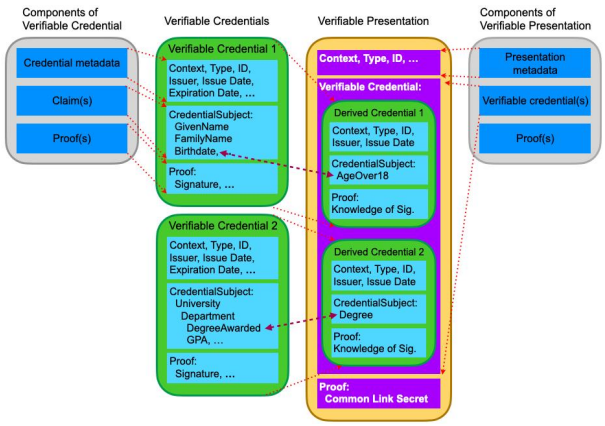

그림 6. 검증 가능 자격증명의 VC와 VP의 관계
Fig. 6. The relationship between VC and VP in verified credential data model

2-3 자기 주권 데이터와 탈중앙 웹

빅 데이터와 인공지능이 기반이 되는 4차 산업혁명 시대에는 데이터 기반 사회라 할 정도로 데이터가 모든 사회의 기반이 될 것이다 [12]. 미래의 데이터 기반 사회가 현재의 웹 환경과 같이 글로벌 플랫폼 기업이 개인의 데이터를 독점하고 개인의 데이터로 창출되는 가치와 그로 인한 수익을 독점하는 사회로 유지된다면 매우 심각한 문제가 될 것이다. 이러한 독점은 사용자의 신원정보, 생성 데이터, 애플리케이션 또는 서비스를 개별 플랫폼별로 통합하여 처리하고 관리하기 때문에 발생한다. 글로벌 플랫폼 기업의 서비스 제공 구조는 그림 7과 같이 데이터와 서비스가 통합된 형태로 자신이 관리하는 데이터 저장소에 공유된 사용자 데이터에 따라 경쟁 우위를 확보하는 사일로 구조 (silo architecture)이다.

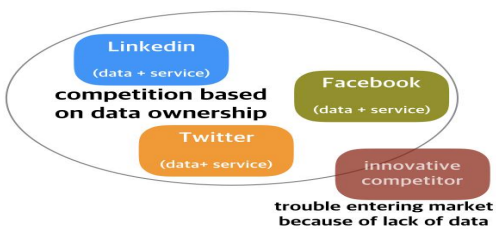

그림 7. 플랫폼이 소유하고 제어하는 개인 데이터
Fig. 7. The personal data owned and controlled by platforms

현재 웹 생태계에서 플랫폼 기업은 보유한 사용자의 개인 데이터를 독점 활용하여 새로운 가치를 제공하고 수익을 창출하기 때문에 보유한 데이터가 부족한 기업의 시장 진입 장벽이 높아 서비스의 혁신이 어려운 상황이다. 또한 플랫폼 기업이 사용자의 모든 데이터를 저장 관리하면서 개인 데이터의 대량 유출, 오남용, 악용, 단일 장애 지점 (SPF: Single Point of Failure), 프라이버시 침해, 데이터의 검열 등의 문제가 있다.

따라서 이러한 문제를 해결하기 위해서 세계적으로 법/규제 및 기술적 대응이 활발하게 이루어지고 있다. 소수의 플랫폼 기업이 데이터를 독점하고 경제적 자산으로 데이터의 가치가 높아지면서 데이터의 경제적, 사회적 가치를 창출할 수 있는 데이터의 소유권과 통제권을 데이터의 생산자가 가지는 데이터 주권 (data sovereignty)과 개인의 프라이버시가 보장되며 가명과 익명의 사용이 보장되는 신뢰성이 지원되고 데이터 주권을 보장하는 탈중앙 웹 생태계가 이슈로 부상하고 있다.

개인의 데이터 주권과 프라이버시 보호를 위해 법과 규정으로 EU의 general data protection regulation (GDPR)[13], 미국 캘리포니아주의 california consumer privacy act (CCPA)[14], 중국의 사이버보안법 [15], 우리나라의 데이터 3법(개인정보보호법, 정보통신망법, 신용정보법) [16]이 제정되었다. 이들 법은 개인 데이터의 수집, 처리, 이동, 삭제 등에 대한 기업의 책임 강화, 프라이버시에 관련한 데이터 주체의 권리, 익명 정보 개념 등을 규정하고 있다. 또한 국가가 국민의 데이터 주권을 보호하기 위해 기업이 데이터의 수집 국가 내에 데이터를 저장하고 처리하는 데이터 지역화 (data localization) [17]를 추진 중인 국가가 늘고 있다.

데이터 주권 보장을 위한 데이터 3법의 개정으로 개인 데이터의 자기 결정권을 강화하려는 마이데이터 [18] 개념이 주목받고 있다. 마이데이터는 데이터의 주체가 개인 데이터에 대한 결정권을 갖고 원하는 곳에서 원하는 방식으로 활용하여 개인이 혜택을 누려야 한다는 패러다임 전환이다. 마이데이터는 금융거래, 통신, 구매, 진료, 여행, SNS 등 서비스 이용 중에 생성된 개인의 데이터에 대해 주체가 접근하고 저장하고 활용할 수 있는 환경을 조성하는 것이 목적이다.

데이터 주권은 자기 주권 신원증명 개념과 같이 자신의 데이터를 보호하고 공정하게 데이터의 사용 방법과 목적을 결정하는 권리 [19]로 개인이 정보 주체로서 자신 데이터를 직접 확인, 통제할 수 있는 권리를 보장하는 개념으로 이해된다 [20]. 즉 데이터의 자기 주권은 물리적인 자산에 대한 권리와 같이 정보의 권리를 개인에게 부여하는 것으로 스스로 자신의 데이터가 어디서 어떤 목적으로 어떻게 사용될지를 결정할 수 있는 권리이다. 본 논문에서는 DID와 SSI 개념을 활용하여 데이터의 자기 주권이 보장되는 데이터 유통플랫폼 구조를 제안한다.

2-4 데이터 유통플랫폼

현재 인터넷의 데이터 유통플랫폼으로 공공데이터 포털 [21] 및 데이터스토어 [22]이 있다. 공공데이터 포털은 공공기

관에서 만들어지는 데이터를 저장하며 서비스 제공자는 이를 이용하여 서비스를 개발한다. 데이터스토어는 누구나 쉽게 데이터를 온라인에서 판매하고 구매하는 데이터 오픈마켓으로 본 논문에서 고려하는 데이터 유통플랫폼의 기능과 같으나 본 논문의 데이터 유통플랫폼은 탈중앙 웹으로 중앙형 데이터스토어와 달리 플랫폼 제공자와 달리 데이터 주권이 데이터 소유자에게 있다.

또한 블록체인 기반 데이터 유통플랫폼과 관련해서 블록체인 기반 사물인터넷 데이터 유통플랫폼 연구가 진행되고 있으나 DID, VC, VP 기술이 통합된 데이터 유통플랫폼 관련 연구는 아직 초기 단계이다.

Ⅲ. 자기 주권 보장 데이터 유통플랫폼 설계

최근 미디어 콘텐츠(비디오, 오디오, 사진, 텍스트 데이터 등)를 주고받을 수 있는 미디어 중심 사물인터넷 (media-centric Internet of Things), 개인이 다양한 나만의 콘텐츠를 직접 제작하고 공유할 수 있는 1인 미디어, 소셜 네트워크에서 개인의 미디어 콘텐츠를 서로 공유할 수 있는 소셜 미디어의 등장으로 개인이 제작하는 미디어 콘텐츠의 양이 증가하고 있다. 현재 사용중인 미디어 콘텐츠 플랫폼은 중앙 집중형 웹 구조로 미디어 콘텐츠의 제어 및 관리가 소수 기업에서 이루어져 미디어 콘텐츠의 자기 주권이 보장되지 않는다.

코로나-19로 사회적 거리두기가 일반화되어 원격교육, 원격의료, 온라인 쇼핑 같은 비대면 경제활동이 증가하고 있으며 포스트 코로나 시대에도 비대면 경제활동이 지속될 것이 예상되므로 경제 주체 간 비대면 신뢰성 확보가 필수적이다. 본 연구에서는 신뢰성과 자기 주권이 보장되는 데이터 유통플랫폼 구조를 제안한다. 제안하는 플랫폼은 탈중앙 웹 (decentralized web) 기반 플랫폼으로 사용자 신원정보와 데이터의 신뢰성과 프라이버시를 보장할 수 있는 요소 기술로 구성된다. 플랫폼의 요소 기술 정리를 위해 플랫폼의 요구사항을 먼저 도출한다.

3-1 요구사항

포스트 코로나 시대에도 비대면 데이터 경제활동이 지속적으로 확대될 것이 예상되므로 신원정보와 데이터의 소유권과 사용자의 프라이버시를 보장하는 것이 필요하다. 즉 사이버 공간에서 비대면 데이터 경제활동의 신뢰성을 확보하여 비대면 경제 활동을 활성화할 수 있는 데이터 유통플랫폼의 요구사항을 다음과 같이 도출하였다.

- 웹의 분산성과 보편성 원칙을 따른다[23].
- 신원정보와 데이터의 신뢰성과 보안성을 보장한다.
- 언제 어디서나 데이터를 검색, 액세스 및 유통 가능한 개방형 플랫폼이어야 한다.
- 데이터 소유자는 데이터 접근 제어 권한을 가진다.
- 데이터 소유자는 데이터를 선별적으로 공유할 수 있다.
- 디지털 자산의 소유권을 보장하고 거래를 통해 디지털 자산 유통이 가능하다.
- 데이터는 다양한 애플리케이션에서 활용 가능해야 한다.
- 사용자, 데이터 등 모두 요소는 DID를 사용하여 식별한다.
- 사용자는 서비스별 식별자 사용이 가능하며 익명 또는 가명의 식별자 사용이 가능하다.
- 요소의 신원정보는 VC 구문으로 기술하며 검증기관에 제출할 신원정보는 VP 구문에 따라 기술한다.

3-2 데이터 유통플랫폼 구조 설계

데이터가 특정 애플리케이션에 종속되지 않게 하여 여러 애플리케이션에서 공동으로 활용하도록 본 논문에서 제안하는 데이터 유통플랫폼의 구조에서는 데이터 주권을 보장하기 위해 그림 8과 같이 애플리케이션과 데이터를 분리한다. 본 논문에서 제안하는 유통플랫폼 구조는 요구사항을 모두 반영하여 설계하였다.

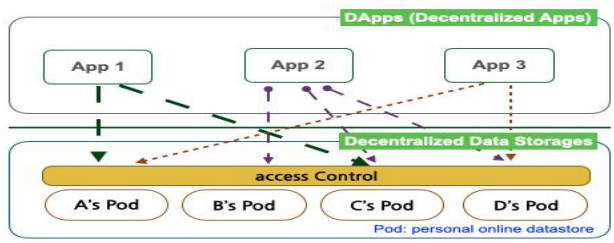

그림 8. 애플리케이션과 데이터의 분리
Fig. 8. The separation of application and data

애플리케이션과 데이터가 분리된 구조에서 사용자 데이터는 링크 데이터 (linked data) 표준을 따라 애플리케이션이 네트워크에 분산 저장된 데이터들을 통합 처리할 수 있는 기반을 제공한다. 데이터는 저장 노드가 피어-투-피어 (peer-to-peer) 네트워크 형태로 연결된 탈중앙 데이터 저장소 (decentralized data storage)에 저장된다. 탈중앙 데이터 저장소에 데이터가 분산 저장되며 데이터와 애플리케이션과 분리되어 있어 단일 장애 지점 문제, 동일 데이터가 여러 애플리케이션에 중복 저장 및 관리되는 문제가 해결되며 데이터의 무결성 보장이 된다.

제안하는 데이터 유통플랫폼에서 데이터 생산자와 소비자는 DID 기반 신원증명을 수행한다. 데이터 생산자는 블록체인 기반 탈중앙 데이터 저장소에 데이터를 저장한다. 데이터 소비자는 생산자의 데이터를 사용하기 위해 데이터의 접근 권한을 얻고 데이터 소비자의 탈중앙 데이터 저장소에 저장된 데이터 (디지털 자산)의 접근과 처리는 데이터 생산자와 소비자 간 데이터 유통의 신뢰성을 보장하기 위해서 블록체인 내에 저장된 스마트 계약 (smart contract)을 통해 수행하는 탈중앙 애플리케이션 (DApp: decentralized application)을 통해 수행된다. 데이터의 거래 내용은 블록체인의 분산원장에 기록되어 거래의 투명성과 거래 정보의 무결성을 보장한다.

다음 그림 9는 제안하는 플랫폼의 소프트웨어 아키텍처로 플랫폼의 사용자는 플랫폼 내의 애플리케이션 또는 다른 사용자와의 상호작용 과정에서 본인을 인증하기 위해서 DID를 사용한다.

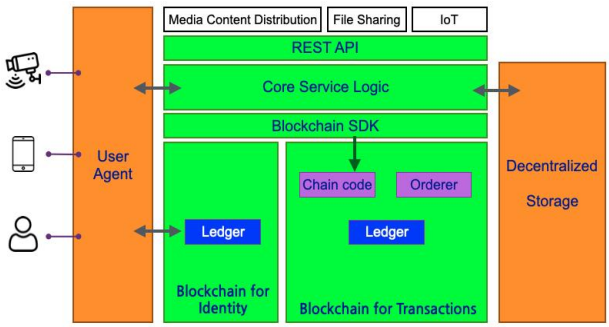

그림 9. 제안하는 플랫폼의 소프트웨어 아키텍처
Fig. 9. Proposed software architecture of platform

데이터 유통플랫폼에서 디지털 자산 역할을 하는 데이터 저장 및 데이터 거래 내용 저장 관리와 신원증명을 위해 저장 관리되는 정보영역으로 구분한다. 이는 데이터를 저장하는 탈중앙 데이터 저장소 (decentralized storage), 신원증명 관련한 저장소 (blockchain for identity), 데이터 거래와 관련하는 스마트 계약 (smart contract), 거래와 관련한 체인 코드 (chain code), 거래 정보를 저장하는 저장소 (blockchain for transactions) 이다. 플랫폼에서는 두 개의 블록체인을 통해 데이터 관리계층과 신원증명 관리계층을 구성한다. 데이터 관리계층은 탈중앙 저장소의 피어-투-피어 네트워크를 통해 안전하고 투명한 데이터 교환을 담당하며, 신원증명 관리계층은 플랫폼 참여자가 자신의 신원증명 데이터의 접근을 제어할 수 있도록 자신의 데이터에 접근할 수 있는 대상과 접근 권한을 제어하는 기능을 제공하여 데이터의 자기 주권을 보장한다. 제안하는 플랫폼의 소프트웨어 아키텍처는 앞에서 살펴본 세 개의 저장소와 데이터 유통 관련 핵심 서비스 로직 (core service logic), 블록체인 software development kit(SDK), representational state transfer(REST) API 기능으로 구성된다.

핵심 서비스 로직은 데이터 유통, 파일 공유, 개인의 데이터를 활용하는 DApp 개발에 필요한 기능을 제공한다. 데이터 거래를 위한 블록체인은 탈중앙 저장소에 저장된 데이터의 프라이버시 보호 및 검색, 그리고 향후 데이터 거래 내용 검증을 위한 거래 내용을 저장한다. 신원정보를 위한 블록체인은 사용자의 신원증명을 위해 필요한 데이터 (public key, serviceEndpoint, DID document 등)를 탈중앙 관리한다. 플랫폼에 거래되는 데이터는 탈중앙 저장소에 저장되고 있어 데이터 거래를 위한 블록체인에는 거래되는 데이터가 저장된 탈중앙 저장소에 저장된 데이터의 인덱스 정보를 저장한다. 프라이버시에 민감한 데이터는 누구나 열람할 수 있으며 변경 불가능한 블록체인에 저장되지 않으며, 저장해야 하는 경우 데이터를 암호화하여 저장한다.

3-3 데이터 거래 흐름

제안 플랫폼의 데이터 유통을 위한 데이터 거래 흐름을 그림 10에 나타냈다. 데이터 생산자는 생성한 데이터를 탈중앙 저장소에 저장하기 위해 dataStore 메시지를 통해 핵심 서비스 로직 컴포넌트에 전송한다. 데이터가 탈중앙 저장소에 저장된 후 스마트 계약에 해당하는 체인 코드 (chain code)를 호출하여 분산원장에 데이터 생성을 기록한다.

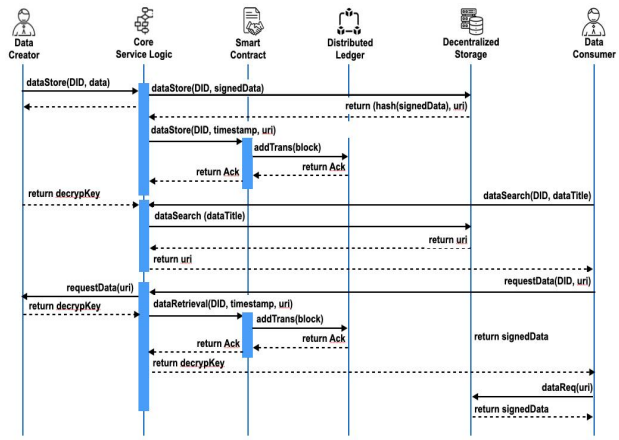

그림 10. 데이터 유통을 위한 거래 데이터 흐름
Fig. 10. Transaction flow for data distribution

데이터 소비자는 dataSearch 메시지를 통해 원하는 데이터를 검색하고 requestData 메시지를 통해 데이터 생산자에게 데이터 요청을 통보하고 데이터 생산자는 데이터를 액세스할 수 있는 키 (예, decrypKey)를 전달하여 사용 허가를 표현하면 핵심 서비스 로직 컴포넌트는 데이터 소비자와 생산자 간의 데이터 거래 기록을 분산 원장에 기록한다. 핵심 서비스 로직 컴포넌트는 데이터 거래를 분산 원장에 기록한 후 데이터 소비자에게 서명된 데이터와 키를 데이터 소비자에 전달한다. 데이터 소비자는 전달받은 데이터를 사용한다.

Ⅳ. 결 론

포스트 코로나 시대에도 디지털 전환이 발생하는 산업영역이 지속 확대될 것이며 이에 따라 웹 기반 인터넷상에서 디지털 경제 규모도 증가할 것이다. 디지털 데이터 기반의 디지털 경제 규모가 증가함에 따라 경제 참여자 간의 신뢰 확보 및 프라이버시 보장의 필요성이 높아지고 있다. 현재 웹 환경의 디지털 경제 생태계에서 사용자가 자신의 데이터에 대한 제어권을 가지지 못하는 문제점 때문에 탈중앙 웹 연구가 진행 중이다. 또한 코로나-19로 가속화된 디지털 경제활동이 포스트 코로나 시대에도 일상생활이 될 것이 예상되는데 현재의 웹 생태계는 개인의 데이터의 제어권을 주체가 가지지 못하고 있어 데이터 주권의 확보가 중요하다. 본 논문에서는 코로나-19와 4차 산업혁명으로

인해 등장할 가상물리소셜 생태계의 가상공간에서 미디어 콘텐츠를 비롯한 디지털 자산의 주권을 보장하고 디지털 자산의 신뢰적인 거래를 보장하는 탈중앙 웹 기반의 미디어 콘텐츠 유통 플랫폼 구조를 제안하였다. 이를 위해 W3C에서 표준화하고 있는 탈중앙 식별자와 검증 가능한 자격증명을 기반으로 자기 주권 신원증명 생태계의 신원 및 디바이스 인증체계를 검토하고 데이터 주권 보장을 위해 제정된 국내외 법 규제 및 기술적 요소를 살펴보았으며 생태계 참여자의 신뢰와 프라이버시를 보장하며 미디어 콘텐츠와 같은 디지털 자산의 거래의 신뢰성을 보장하는 플랫폼의 요구사항, 구조, 데이터 흐름을 제안하였다. 향후 제안한 구조를 다양한 데이터 생산자를 고려하여 블록체인 오픈소스 기반 가상물리소셜 생태계를 구축할 계획이다.

참고문헌

[1] H. W. Kim, H. Y. Kim, K. J. Song, M. H. Lee, and P. H. Han, COVID-19 and Digital Transformation, 2020 Highlights and 2021 Outlook, SPRI FOCUS, Dec. 2020.

[2] B. Y. Kim, Non-contact Service Treands and Implications due to COVID-19, FKII Issue Report 2020-1, April, 2020.

[3] D. Yoon, *Self-Sovereign Identity Architecture Analysis*, Jpub, 2020.

[4] Decentralized Identifiers (DIDs) v1.0, Core architecture, data model, and representations[Internet]. Available: https://www.w3c.org/TR/did-core.

[5] P. Wang, L. T. Yang, J. Li, J. Chen, and S. Hu, "Data fusion in cyber-physical-social systems: State-of-the-art and perspectives," *Information Fusion*, vol. 5. pp.42-57, Nov. 2019.

[6] A Primer for Decentralized Identifiers[Internet]. Available:https://w3c-ccg.github.io/did-primer/.

[7] Use Cases and Requirements for Decentralized Identifiers[Internet]. https://www.w3.org/TR/did-use-cases/.

[8] Verifiable Credential Data Model 1.0, Expressing verifiable information on the web[Internet]. Available: https://www.w3.org/TR/vc-data-model/.

[9] A. Muhle, A. Gruner, T. Gayvoronskaya, C. Meinel, A Survey on Essential Components of a Self-Sovereign Identity," *Computer Science Review,* 30, pp. 80-86, 2018.

[10] White Paper: The Inevitable Rise of Self-Soverign Identity, Sovrin Foundation, 2017.

[11] Principles of SSI[Internet]. Available: https://www.sovrin.org /principles-or-ssi/.

[12] The Economist, The Data Economy: The world's most valuable resource, May 2017.

[13] GDPR Homepage[Internet]. Available: https://gdpr,eu/.

[14] California Consumer Privacy Act(CCPA) Homepage [Internet]. Available: https://oag.ca.gov/privacy/justice.

[15] KIEP Report, A Research on Cyber Security Policy of China, Korea Institute for International Economic Policy, July, 2020.

[16] KISA Report, Prospects on the Revisions of Thee Data-related Bills, vol. 2, February, 2020.

[17] Impact of Data Localization Requirements on Commerce and Innovation[Internet]. Available: https://www.americanactionfrom.org/impact-of-data-localization-requirements-on-commerce-and-innovation/.

[18] KDA Report, My Data Service Guide, December, Korea Data Agency, 2019.

[19] European Commission, European Data Economy, 2017

[20] Weekly Technology Trend, Availability and Development Status of Blockchain Technology in the Data Sovereignty Era, no. 1956, IITP. 2020.

[21] Public Data Portal[Internet], Available: https://data.go.kr.

[22] Data Store[Internet], Available: https://datastore.or.kr.

[23] L.-D. Ibanez, E. Simberl, F. Gandon, H. Story, "Redecentralizing the Web with Distributed Ledgers," IEEE intelligent Systems, vol. 32, no. 1, pp. 92-95, 2017.

김근형(Geun-Hyung Kim)

1986년: 서강대학교 대학교 (공학사)
1988년: 서강대학교 대학원 (공학석사)
2005년: 포항공과대학교 대학원(공학박사)

1988년～1990년: LS 산전 연구원
1990년～1993년: 삼성종합기술원 선임연구원
1993년～2007년: KT BcN본부 수석연구원
2007년～현 재: 동의대학교 게임공학 전공 교수
※관심분야 : 탈중앙 웹, 자기 주권 데이터, 데이터 보호, 블록체인, 인공지능, 설명 가능한 인공지능