

분산화와 개방성 기반의 메타데이터 레지스트리 확장 모델

정 동 원*

An Extended Metadata Registry Model Based on Decentralization and Openness

Dongwon Jeong*

“이 논문은 2019년도 군산대학교 교수 장기국외연수경비의 지원에 의하여 연구되었음.” 또한 “이 연구는 2019년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2019R111A3A01060826).”

요 약

이 논문에서는 기존 메타데이터 레지스트리의 모델의 문제점을 해결할 수 있는 확장 모델을 제안한다. ISO/IEC 11179-메타데이터 레지스트리는 데이터 상호운용성 향상을 위해 개발된 국제 표준이다. 메타데이터 레지스트리 모델은 중앙식 메타데이터 관리와 폐쇄형 메타데이터 평가 체계를 제공하기 때문에 이 모델의 사용은 확장성과 활용성을 저하시키는 문제점을 지닌다. 이 논문에서는 메타데이터 레지스트리 관리 체계의 한계를 분석하고 메타데이터 레지스트리 확장 방안을 제시한다. 제안한 확장 모델은 분산형 관리 및 참여의 개방성을 제공한다. 또한 제안 모델을 구현할 수 있는 가능한 구현 모델을 기술하며, 마지막으로 비교 평가 및 제안 모델의 장점을 기술한다.

Abstract

This paper proposes an extended model for resolving the problems of the existing metadata registry model. ISO/IEC 11179-Metadata registries is an international standard that has been developed to improve data interoperability. The metadata registry model provides a centralized metadata management and closed metadata verification scheme, and thus the use of the model compromises extensibility and usability. This paper analyzes the limitations of the existing metadata registry management scheme and proposes an extended metadata registry model to address the limitations. The extended model proposed in this paper enables a decentralized management and open participation. This paper also shows a possible implementation model to realize the proposed model, and a comparative evaluation and the advantages of the proposed model are described.

Keywords

metadata registry, decentralization, openness, collective intelligence, blockchain

* 군산대학교 소프트웨어융합공학과 교수
- ORCID: <http://orcid.org/0000-0001-9881-5336>

• Received: Dec. 03, 2019, Revised: Jan. 09, 2020, Accepted: Jan. 12, 2020
• Corresponding Author: Dongwon Jeong
Software Convergence Engineering Department, Kunsan National University, Korea,
Tel.: +82-63-469-8912, Email: djeong@kunsan.ac.kr

1. 서 론

클라우드 서비스의 확산, 빅데이터 분석 영역의 확대, 이기종 분산 센서 네트워크 정보의 통합 및 융합 서비스 증가 및 오픈 데이터와 오픈 API 활용의 증가 등으로 데이터 교환 및 공유의 중요성 대두되고 있다. 시스템과 서비스들의 다양성과 복잡성이 증가함에 따라 데이터 교환과 공유의 중요성이 배가되고 있으며, 이러한 데이터 상호운용성이 유효하게 작동할 때 정확하고 풍부한 양질의 서비스를 사용자에게 제공할 수 있다.

지금까지 다양한 분야와 영역에서 데이터 상호운용성을 향상시키기 위한 수많은 연구가 수행되어 왔으며, 특히 규범적인 메타데이터 관리를 통해 데이터 이질성 문제를 근본적으로 해결하기 위한 표준화 연구가 지속적으로 진행되어 왔다[1]-[3]. 가장 대표적인 연구로서, ISO/IEC JTC 1/SC 32에서는 메타데이터를 등록·관리 및 활용을 위한 표준인 ISO/IEC 11179-메타데이터 레지스트리(MDR, Metadata registry)를 개발하였으며[3], 지금까지 지속적인 개정 작업을 수행해 오고 있다. ISO/IEC 11179는 메타데이터의 표준화 절차 및 표준 메타데이터를 관리하기 위한 메타모델 등을 명세하며, 동적 메타데이터 관리, 메타데이터 재사용 체계 등을 통해 데이터 공유 및 교환의 용이성을 제공한다.

앞서 언급하였듯이, 데이터의 교환과 공유를 위한 국제 표준인 ISO/IEC 11179는 여러 가지 장점을 지닌다. 이러한 장점에 기인하여, 의료, 교육, 교통, 환경 등 다양한 분야에서 표준 메타데이터를 정의하고 활용하기 위해 ISO/IEC 11179 모델을 적용하고 있다[4]-[11]. 또한 ISO/IEC JTC 1/SC 32에서 개발한 이 표준은 다른 표준화 분야에도 적용되어 데이터 상호운용성 향상을 위한 연구가 널리 적용 및 활용되고 있다.

메타데이터 레지스트리의 목적은 데이터 공유와 교환을 용이하게 하여 데이터 상호운용성을 향상시키는 데 있다. 이러한 목적을 이루기 위한 메타데이터 레지스트리는 동적 메타데이터 관리 및 재사용 메커니즘 등을 제공한다. 비록 메타데이터 레지스트리가 다양한 장점을 지니고 있으나, 중앙 집중식 관

리 모델, 특정 집단에 의한 표준화 체계 및 이로 인한 확장성과 활용성 저하 등의 문제점을 지닌다.

이 논문에서는 상기 언급한 메타데이터 레지스트리의 한계를 개선하고 향상시킬 수 있는 메타데이터 레지스트리 확장 방안을 제안한다. 이 논문의 목적은 기존 모델의 한계를 극복할 수 있는 확장 방안, 즉 메타모델을 제안하고 이에 대한 타당성을 논의하는 데 있다. 즉, 실질적인 구현 모델이나 구현 레벨에서의 평가에 초점을 두지 않는다. 이 논문에서 제안한 방법은 제한된 집단에 의한 메타데이터 표준화가 아닌 제한되지 않은 다수가 참여할 수 있는 개방형 구조를 지닌다. 또한 중앙 집중식 관리 체계의 한계를 개선할 수 있는 분산형 관리 체계를 지원한다. 이 논문에서는 보다 유연하고 확장성 있는 메타데이터 확장 모델을 보이고 기존 메타데이터 레지스트리 모델과의 정성 평가를 통해 제안 방안의 타당성을 기술한다.

이 논문의 구성을 다음과 같다. 제2장에서는 메타데이터 레지스트리 적용 사례와 그 한계를 기술하고, 제3장에서는 이 논문에서 제안하는 확장 방안의 개념과 제안 방안을 위한 메타모델을 기술한다. 제4장에서는 확장 모델을 구현하기 위한 제안으로서의 구현 모델과 비교 평가 결과를 기술하며, 마지막으로 제5장에서는 결론 및 향후 연구를 기술한다.

II. 메타데이터 레지스트리 및 문제 정의

2.1 메타데이터 레지스트리와 그 활용 사례

ISO/IEC 11179-메타데이터 레지스트리는 데이터 교환 및 공유를 목적으로 개발되었으며, 총 6부로 구성된 국제 표준이다. 현재까지 두 번의 개정을 거쳐 개발되었으며, 현재 세 번째 개정 작업이 진행되고 있다. 메타데이터 레지스트리는 동적 메타데이터 관리, 표준화된 메타데이터 등록 및 재사용을 통해 데이터 이질성 문제를 해결한다.

지금까지 의료, 교육, 교통, 환경 등의 다양한 분야에서 메타데이터 레지스트리를 준용한 시스템을 구현하여 활용하고 있다[4]-[11]. 대표적으로, Metadata online registry(METeOR)[4], Cancer data standards

repository(caDSR)[5] 및 교육 분야 활용 사례[6] 등을 예로 들 수 있다. METeOR은 호주 건강보건국(Australian institute of health and welfare, AIHW)에서 개발한 시스템으로, 국가에서 인증한 메타데이터 표준을 정의하고 활용을 위한 다양한 기능을 제공한다[4]. caDSR은 NCI clinical trial network(NCTN/ETCTN)와 Lead protocol organizations(LPOs) 간 일관성과 정확성 있는 정보 교환을 위해 개발이 시작되었다. caDSR은 NCI(National cancer institute) 주도로 ISO/IEC 11179 전문가의 컨설팅을 거쳐 개발되었으며, 도메인의 요구 사항을 수용하여 추가적인 기능을 개발하여 제공한다[5]. 교육 분야에서도 일관성 있고 정확한 정보 공유를 위해 활용되고 있으며, 콜로라도, 메인, 몬타나, 오레곤 및 워싱턴 등 미국의 많은 주에서 메타데이터를 정의하여 활용하고 있다[6].

상기 언급한 활용 예 외에도 메타데이터 레지스트리의 활용성을 향상시키기 위한 연구[12], 메타데이터 레지스트리 개념을 사물인터넷 환경에 적용한 연구[13][14] 및 임상 실험 분야에 적용한 연구[15] 등이 진행되었다.

2.2 문제 정의

ISO/IEC 11179-메타데이터 레지스트리는 데이터 교환 및 공유를 위해 개발되었으며, 규범적인 메타데이터 정의, 등록, 재활용을 통한 데이터 상호운용성을 향상시킨다. 메타데이터 레지스트리와 유사한 메타데이터 표준들이[1][2] 개발되었으나 이 표준들은 특정 분야에 한정적이며 정적인 하향식 접근 방법을 제공한다. 반면, 메타데이터 레지스트리는 특정 분야에 종속적이지 않으며, 동적 메타데이터 관리가 가능한 상향식 접근 방법을 지원한다. 이러한 동적 메타데이터 관리는 제한적이고 수동적인 메타데이터 관리의 한계를 해결함으로써 높은 확장성을 제공한다.

앞서 기술하였듯이, 메타데이터 레지스트리가 데이터 상호운용성 향상에 많은 장점을 제공함에도 불구하고 여전히 몇 가지 한계를 지닌다. 메타데이터 레지스트리는 메타데이터 표준화 및 메타데이터 관리를 위해 중앙 집중식 관리 체계를 제공한다. 이

러한 중앙 집중식 관리는 그 구조적인 특성으로 인해 효율적인 측면이 있다. 그러나 메타데이터는 특정 분야에만 종속되지 않고 다양한 분야, 다양한 시스템 및 다양한 서비스와 연계되어 활용되기 때문에 중앙 집중식 관리 체계는 한계를 지닌다. 무엇보다 어느 기관이 관리 기관으로서의 권한을 소유할 것인가에 대한 문제는 기관의 특성과 구조, 비용 문제 등의 측면에서 결정이 쉽지 않다.

또 다른 한계는 한정되고 제한된 집단에 의해 메타데이터 표준화 과정이 이루어진다는 점이다. ISO/IEC 11179를 적용하여 개발된 많은 시스템과 서비스에서 정의한 메타데이터는 해당 기관에서 구성된 특정 전문가 집단에 의해 정의된다. 이러한 한정된 집단에 의해 검토 및 검증 과정을 거쳐 메타데이터를 표준화하기 때문에 다양한 다수의 의견과 전문성이 반영되지 못한다는 한계를 지니며, 아울러 상호운용성 측면에서 충분한 기여를 하지 못한다.

메타데이터 레지스트리의 목적은 데이터 교환과 공유를 통한 데이터 상호운용성 향상에 있다. 그러나 앞서 언급한 이러한 한계는 그 확장성과 활용성을 저해하는 요인으로 작용한다. 따라서 이 논문에서는 이러한 한계를 극복할 수 있는 확장 방안을 제안한다.

III. 메타데이터 레지스트리 확장 모델

3.1 개념 모델

그림 1은 이 논문에서 기존의 메타데이터 레지스트리 모델의 문제점과 그 문제점을 해결하기 위한 접근 방법을 보여준다. 그림 1(a)에서, ISO/IEC 11179-메타데이터 레지스트리는 크게 두 가지 장벽, 즉 중앙 집중식 관리 체계와 제한된 특정 집단에 의한 표준화로 인해 확장성과 활용성이 낮아짐으로써 장점을 극대화하지 못한다.

그림 1(b)는 ISO/IEC 11179-메타데이터 레지스트리의 한계를 극복하기 위한 제안 방법의 개념도를 보여준다. 이 논문에서 제안하는 확장 방안의 핵심적인 요소는 분산화(Decentralization)와 개방성(Openness)이다. 분산형 메타데이터 관리 체계를 통해 중앙 집중식 관리 체계의 한계를 개선할 수 있다.

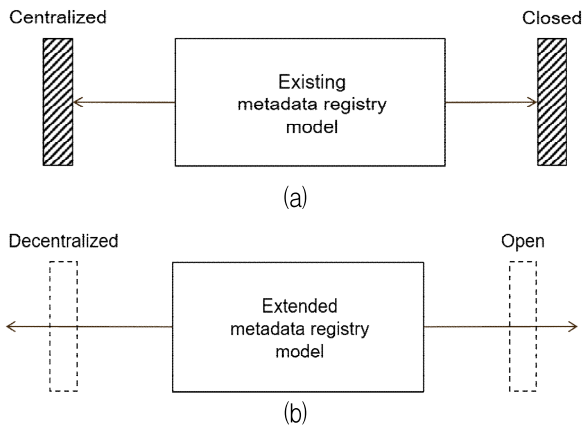


그림 1. 접근 방법의 개념
Fig. 1. Concept of approach

메타데이터 레지스트리의 또 다른 한계는 제한된 특정 집단에 의한 메타데이터 표준화에 있다. 이러한 표준화 체계는 메타데이터 레지스트리가 지니는 궁극의 목적인 데이터 상호운용성 향상을 극대화하기 어렵다. 따라서 이 논문에서는 이러한 한계를 극복하기 위해 메타데이터 표준화에 대한 참여를 개방하여 폭넓은 집단지성에 의해 메타데이터를 표준화할 수 있는 체계를 제안한다. 이러한 개방형 메타데이터 표준화 체계는 참여와 합의 의사 결정 구조를 확장시킬 수 있다.

3.2 메타데이터의 분산화

이 논문에서 제안하는 메타데이터 레지스트리 확장 방안 중 첫 번째 개념은 메타데이터의 분산화이다. 그림 2는 메타데이터 관리를 위한 분산화 개념 구조를 보여준다.

그림 2에서, 메타데이터는 다수의 노드에 분산되어 관리된다. 전체 메타데이터 집합은 각각 Metadata_set-1, Metadata_set-2, Metadata_set-3 및 Metadata_set-4에 분산 및 저장된다. 이 때, 각 메타데이터 집합은 서로 다른 메타데이터를 저장하지 않고 참여에 따른 상호 중복된 메타데이터를 관리하게 된다. 따라서 전체 메타데이터 집합은 분산된 메타데이터 집합의 합이 되며, 각 분산 메타데이터 개수보다 적은 수를 지니게 된다. 표준화 후보로 제출된 메타데이터를 포함한 전체 메타데이터 집합을 M , 분산된 메타데이터 집합을 DM_i 라고 표현할 때, 다음과 같이 정의된다.

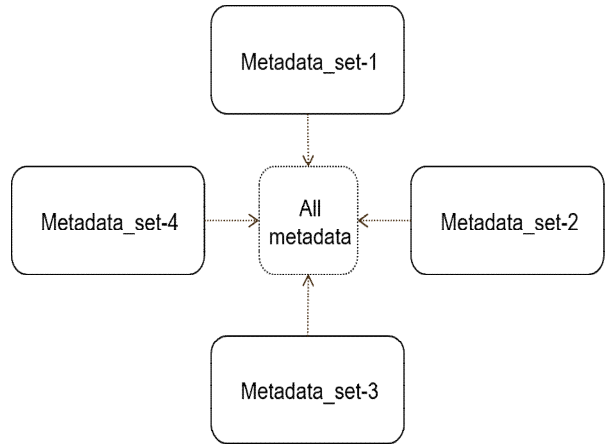


그림 2. 메타데이터 분산 구조
Fig. 2. Decentralization of metadata

$$M = DM_1 \cup DM_2 \cup \dots \cup DM_n \quad (1)$$

$$n(M) \leq n(DM_1) + n(DM_2) + \dots + n(DM_n) \quad (2)$$

식 (2)에서 $n(M)$ 과 우변의 값이 동일한 경우는 표준화를 위해 제출된 메타데이터에 대한 참여가 이루어지지 않았음을 의미하며, 서비스의 초기 단계에 일시적으로 발생할 수 있다.

3.3 집단지성에 의한 합의 의사 결정 프로세스

메타데이터 레지스트리 모델에서의 메타데이터 표준 결정은 임의의 제출된 메타데이터를 특정 전문가 집단이 판단하고 검증하는 프로세스이다. 이러한 프로세스는 ISO/IEC 11179 제6부[16]에 명세되어 있으며, 이를 도식화하면 그림 3과 같다.

기존 메타데이터 레지스트리 모델에서, 일반 사용자인 제안자(Submitter)가 메타데이터를 식별한 후 필요한 정보를 기술하여 표준 메타데이터 후보로 제안한다. 중간 검토 및 평가 역할을 수행하는 검토 집단(Steward)은 새롭게 제출된 메타데이터를 검토하고 필요할 경우 제안자에게 기술(Description)에 대한 보완을 요청한다. 이러한 과정을 거쳐 메타데이터에 대한 검토 작업이 완료되면 최종적으로 최종 확인 집단인 관리 위원회(Control committee) 또는 등록자(Registrar)를 통해 표준 메타데이터로 확정하여 메타데이터 레지스트리에 등록한다.

그림 3에서 알 수 있듯이, 현재의 메타데이터 레지스트리 모델은 특정 제한된 집단에 의한 평가와 확정 과정을 거쳐 표준 메타데이터를 등록한다. 이러한 제한된 일부 전문가 집단에 의한 표준화 과정, 즉 폐쇄적인 메타데이터 표준화 및 등록 체계는 다양한 참여자의 의견을 반영하기가 어려우며, 무엇보다 닫힌 환경에서 합의가 이루어지기 때문에 생산된 메타데이터의 높은 활용성을 제공하지 못한다.

본 논문에서는 상기 기술한 한계를 극복할 수 있는 방안으로 집단지성과 개방성에 기반한 확장된 메타모델을 제안한다. 제안 모델은 참여자에게 열린 개방형 환경을 제공한다. 즉, 관심이 있는 모든 참여자의 참여를 허용하고 참여한 참여자들의 의견을 반영하여 합의된 메타데이터 표준을 결정할 수 있는 등록 체계를 제공한다. 이는 집단지성[17]에서 제공하는 구조이자 가장 큰 장점인 합의 의사 결정 (Consensus decision making) 구조와 동일하다. 그림 4는 이 논문에서는 제안하는 집단지성 기반의 메타데이터 표준화 합의 결정 과정을 보여준다.

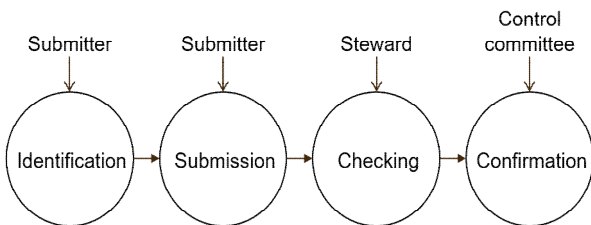


그림 3. 메타데이터 레지스트리 등록 절차
Fig. 3. Registration procedure of metadata registry

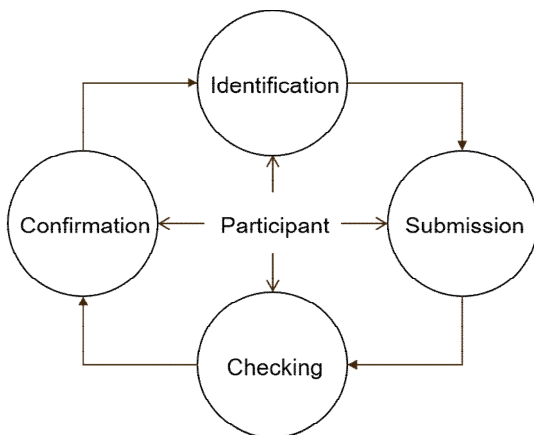


그림 4. 제안 모델에서의 등록 절차
Fig. 4. Registration procedure of proposed model

그림 4에서 제안 모델에서의 등록 절차, 즉 집단지성 기반의 메타데이터 표준화 합의 의사 결정 과정은 참여자들이 모든 프로세스에 참여하고 최종 등록 여부를 결정한다.

따라서 제안 모델에서는 사용자가 제안자이자, 중간 평가자이며 최종 확정 기능을 수행하는 관리 위원으로서의 역할을 수행한다. 제안한 확장 메타모델을 물리적인 시스템으로 구현하여 서비스할 경우, 최소한의 시스템 관리를 위한 관리자를 구성하여 운영할 수 있으나, 메타데이터 표준화를 합의 과정에서의 실질적인 행위를 위한 특정 제한 집단으로 구성하지 않는다.

IV. 구현 방안 및 평가

4.1 제안 모델과 블록체인

이 논문에서 제안하는 메타데이터 레지스트리 확장 모델의 핵심은 분산과 개방이다. 중앙 집중식 메타데이터 관리의 한계를 극복하기 위해 메타데이터 분산 관리를 지원하고, 제한된 특정 집단에 의한 닫힌 표준화 구조가 아닌 집단지성 기반의 열린 합의 의사 결정 구조를 지원한다. 이 논문에서 제안하는 확장 방안의 구현은 다양한 방법으로 가능하다. 이 논문에서는 이러한 제안하는 확장 방안을 구현할 수 있는 가능한 한 방법으로서, 블록체인 기반 구현 모델을 제시한다. 서두에 언급하였듯이, 이 논문의 목적은 기존 모델의 한계를 극복할 수 있는 메타레벨에서의 확장된 모델 제안하고 이에 대한 타당성을 논의하는 데 있으며, 구현과 구현 결과에 대한 평가에 초점을 두진 않는다. 그러나 제안한 방안을 구체화할 수 있는 구현 방안을 제시함으로써 관련 분야에서의 구현 연구에 기여할 수 있다.

블록체인 기술은 최근 많은 관심과 함께 가상 화폐 분야를 포함하여 다양한 분야에 적용하기 위한 연구가 지속적으로 수행되고 있다. 블록체인은 보안성 강화를 위해 개발된 기술로서, 사토시 나카모토 (Satoshi Nakamoto)가 블록체인을 가상 화폐 분야에 적용한 연구 내용이 공개되면서 널리 알려지고 확산되었다.

블록체인은 보안성이 높은 기술로서, 탈중앙화, 개방화, 데이터 분산화라는 특성을 지닌다[18]. 블록체인 기술은 참여자들이 트랜잭션을 독립적으로 검증하고 참여자들의 집단정보를 이용하여 증명(Authentication)하는 구조이다. 또한 블록체인에서의 데이터는 P2P(Peer-to-Peer) 네트워크와 분산 타임스탬핑 서버를 이용하여 자율적으로 관리된다. 블록체인은 개방된 환경으로, 기본적으로 참여자에 대한 특별한 허가(Permission)가 요구되지 않는 구조이다.

앞서 기술한 이러한 블록체인 기술은 이 논문에서 제안하는 메타데이터 레지스트리 확장 모델의 핵심 개념들과 매우 일치하는 특성을 지닌다. 따라서 이 논문에서 제안한 메타데이터 레지스트리 확장 모델을 구현하기 위한 한 방법으로서 블록체인 기술을 적용할 수 있다.

4.2 블록체인 기반 메타데이터 분산 구조

그림 5는 메타데이터 레지스트리 확장 모델을 구현하기 위해 블록체인 기술을 이용한 시스템 구조를 보여준다. 참여자들은 각각 독립적인 노드(P1, P2, P3, P4)를 구성하며, 각 노드는 자신들이 참여한 메타데이터 집합을 지닌다. 각 노드의 참여자는 제한된 권한을 지니지 않고 메타데이터 발굴(Identification), 제출(Submission), 검토(Checking) 및 확정(Confirmation) 단계까지 참여할 수 있는 개방된 권한을 지닌다. 임의의 참여자가 제출한 메타데이터(표준 메타데이터 후보)는 참여자들의 검토와 합의를 통해 표준 메타데이터로 등록된다.

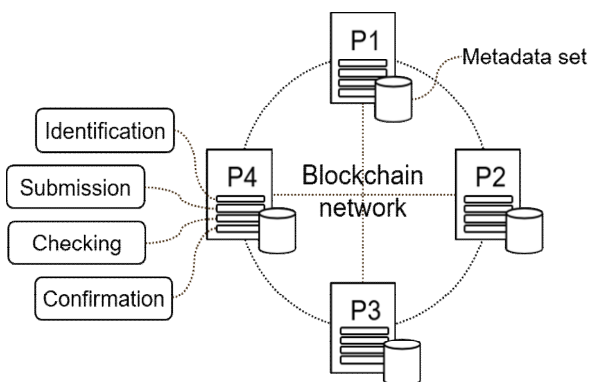


그림 5. 블록체인 기반 확장 구조
Fig. 5. Blockchain-based extended structure

4.3 비교평가

이 논문에서는 기존의 메타데이터 레지스트리가 지니는 한계를 극복하고 확장성과 활용성을 향상시킬 수 있는 확장 모델을 제안하였다. 또한 제안 모델을 구체화할 수 있는 블록체인 기반 구현 모델을 보였다. 제안한 확장 모델은 기존 메타데이터 레지스트리 모델에 비해 여러 가지 측면에서 장점을 지니며, 표 1은 기존 메타데이터 레지스트리 모델과 제안 모델을 비교 평가 결과를 보여준다.

표 1. 정성 평가 결과
Table 1. Qualitative evaluations

Comparative item	Comparison
Diversity of participation	$ PM \leq EM $
Scale of participation	$ PM \leq EM $
Degree of agreement	$ PM \leq EM $
Reliability of metadata	$ PM \leq EM $
System scalability	$ PM < EM $
Usability of metadata	$ PM \leq EM $
Contribution of interoperability	$ PM \leq EM $
Security	$ PM < EM $
Efficiency of storage	$ PM > EM $
Rapidity of standardization	$ PM > EM $

※ PM(Existing model), EM(Proposed-extended model)

먼저, 참여의 다양성(Diversity of participation) 측면에서, 기존 모델은 특정 제한된 집단에 의한 메타데이터 표준화 과정을 제공하기 때문에 참여가 제한적이다. 반면, 제안 모델은 개방형 참여 구조를 통해 다양한 분야의 참여자가 메타데이터 표준화 과정에 관여한다. 따라서 제안 모델은 다양한 참여자가 메타데이터 표준화에 참여하기 때문에 기존 모델에 비해 높은 참여의 다양성을 지원한다.

실질적인 운영 과정에서 제안 모델은 기존 모델과 동일한 환경에서 운영되며, 이론적으로 추가 참여가 없는 상태로 메타데이터 표준화가 진행된다 해도 동일한 수준의 다양성을 지원한다. 이러한 다양성은 참여 규모(Scale of participation)의 차이를 가져오게 되며, 제안 모델은 기존 모델의 참여 규모 이상의 규모를 지원한다. 따라서 참여의 규모 측면

에서 제안 모델이 지원하는 참여의 규모가 기존 모델에 비해 크다.

앞서 언급하였듯이, 제안 모델이 기존 모델에 비해 참여의 다양성과 참여 규모 측면에서 장점을 보인다. 이러한 참여의 다양성과 참여 규모의 개선은 보다 나은 수준의 메타데이터 표준화 합의 의사 결정 결과, 즉 높은 수준의 참여 합의성(Degree of agreement)을 제공한다.

메타데이터의 신뢰성(Reliability of metadata) 측면에서, 이 논문에서 정의한 메타데이터 신뢰성은 합의 과정을 통해 표준화된 메타데이터에 대한 대표성을 의미한다. 앞서 기술하였듯이, 제안 모델은 특정의 제한된 집단만이 아닌 참여를 희망하는 모든 참여자가 메타데이터 표준화 과정의 모든 단계에 참여할 수 있는 열린 개방형 의사 결정 구조를 지닌다. 이러한 구조는 참여의 다양성, 참여 규모의 확장성, 높은 참여 합의성을 제공한다. 따라서 제안 모델을 통해 표준화된 메타데이터의 대표성이 높은 수준으로 수렴하며 이는 메타데이터 신뢰성을 향상시킨다.

시스템 확장성(System scalability) 측면에서, 제안 모델은 중앙 집중식 구조가 아닌 탈중앙화를 지향한 분산 구조를 지원한다. 즉, 이 논문에서 제안한 확장 모델은 P2P 네트워크 기반 분산형 구조를 지원한다. 따라서 기존 모델의 구조적인 특성에서 비롯된 한계를 개선함으로써 높은 시스템 확장성을 제공한다.

메타데이터 활용성(Usability of metadata) 측면에서, 등록된 표준 메타데이터를 재사용함으로써 데이터 교환 및 공유를 용이하게 하여 불일치 문제 해결이 ISO/IEC 11179 메타데이터 레지스트리 모델의 궁극적인 목적이다. 이를 위해서는 등록된 메타데이터의 활용성이 매우 중요한 요소이다. 제안 모델은 참여의 다양성, 참여 규모, 참여 합의성 및 메타데이터 신뢰성 측면에서 기존 모델에 비해 높다. 특히 표준화 합의 결정 과정에서 개방성을 통해 참여자의 적극적 참여를 유도하기 때문에 활용성 또한 높아지게 된다. 또한 이러한 메타데이터의 활용성 향상은 데이터 상호운용성 향상으로 이어지며, 따라서 기존 모델에 비해 높은 상호운용성의 기여도

(Contribution of interoperability)를 제공한다.

메타데이터 레지스트리의 목적이 보안성(Security) 강화는 아니지만 실질적인 시스템 운영 측면에서 보안성은 중요한 요소이다. 이 논문에서는 구현을 위한 물리적인 시스템 모델로서 블록체인 기술을 이용한 모델을 제시하였다. 블록체인은 보안성이 높은 기술이며, 따라서 제안한 확장 모델을 기반으로 구현 시스템은 매우 높은 보안성을 제공한다.

제안 모델은 기존 모델과 비교하여 여러 가지 측면에서 장점을 지닌다. 반면, 제안 모델의 구조적인 특성상, 몇 가지 단점도 지닌다.

먼저, 제안 모델은 구조적인 특성으로 인해 메타데이터의 중복이 발생하고 블록체인 네트워크 구성을 위한 추가적인 공간이 요구된다. 따라서 참여하는 노드를 포함한 전체 저장공간의 효율성(Efficiency of storage) 측면에서 기존 모델에 비해 낮은 효율성을 보인다. 그러나 이러한 특성은 중앙 집중식 서버 구성에 따른 부하 등을 고려할 때 장점으로 작용할 수 있다. 무엇보다 메타데이터의 표준화와 이를 통한 상호운용성 향상 측면 등을 고려할 때 저장공간의 효율성 저하의 한계가 단순히 단점으로서만 작용한다고 판단하기 어려운 측면도 있다.

또 다른 단점으로서, 표준화의 신속성(Rapidity of standardization)이 기존 모델에 비해 낮을 수 있다는 점이다. 기존 모델은 메타데이터 표준화 각 단계에서 역할을 수행할 특정 집단을 사전에 구성한다. 이러한 구조에서는 제안자가 제출한 후보 메타데이터에 대한 표준화를 신속하게 진행할 수 있다. 반면, 제안 모델은 기존 모델과 동일한 수준의 참여자를 참여시키고 참여를 희망하는 모든 참여자가 참여할 수 있는 개방형 합의 결정 구조를 지닌다. 이러한 구조적 특성으로 인해 제안 모델은 기존 모델에 비해 표준화의 신속성이 저하된다. 이러한 신속성 저하가 제안한 확장 모델의 한계로 작용하는 측면도 있다. 그러나 보다 다양한 참여자의 의견을 모아 신뢰성 높은 메타데이터 표준화 생산하고 이를 통해 데이터 공유 및 교환의 신뢰성을 확보에 기여한다는 측면에서 긍정적인 효과가 더 높다고 판단된다.

메타데이터 레지스트리의 궁극적인 목적은 데이터 상호운용성 향상이다. 따라서 앞서 언급한 저장

소의 효율성과 메타데이터 표준화의 신속성이 핵심 평가 요소는 아니며, 제안 모델이 제공하는 다양한 특성에 비해 중요성이 높지 않다고 할 수 있다. 그러나 실제 시스템을 구현하여 운영하기 위해 충분히 고려되어야 하는 요소이기도 하다. 따라서 구현 측면에서, 향후 이러한 요소를 최대한 개선할 수 있는 연구가 이루어져야 할 것이다.

V. 결론 및 향후 과제

이 논문에서는 기존 메타데이터 레지스트리의 확장성과 활용성 향상을 위한 확장 방안을 제안하였다. 제안하는 확장 방안에 대한 개념과 확장된 모델의 특성을 기술하였으며, 확장 모델을 구현하기 위한 블록체인 기반의 구현 모델을 제시하였다. 또한 비교 평가 결과를 통해 제안 모델의 장점을 보였다. 제안 모델은 여러 가지 측면에서 많은 장점을 지니는 반면, 몇 가지 단점도 지닌다. 그러나 이러한 단점이 제안 모델의 궁극적인 목적을 고려했을 때 결정적인 단점으로 단정하기 어렵다는 점을 보였고 아울러 실질적인 운영을 위해 개선이 요구된다는 점을 기술하였다.

이 논문에서는 확장 방안과 이를 구체화하기 위해 요구되는 개념 모델, 가능한 구현 모델 등을 기술하였다. 향후에는 실질적인 시스템 구현을 통해 제안 모델의 기능성 및 성능 등에 대한 검증 연구가 요구된다. 또한 제안 모델이 지니는 단점을 개선할 수 있는 구현 레벨의 시스템 모델에 대한 연구가 필요하며, 보다 객관적인 평가 결과를 보일 수 있도록 가능한 평가 요소에 대한 정량 평가 연구가 요구된다.

References

- [1] The Network Development and MARC Standards Office (the Library of Congress) and the Standards and the Support Office (the Library and Archives Canada), "MAchine-Readable Cataloging (MARC)", <http://www.loc.gov/marc/>. [accessed: Nov. 21, 2019]
- [2] The Dublin Core Metadata Initiative, "The Dublin Core Metadata Registry", <http://dublincore.org/dcregistry/>. [accessed: Nov. 21, 2019]
- [3] The International Organization for Standardization, "ISO/IEC 11179-3:2013 Information technology - Metadata registries (MDR) - Part 3: Registry metamodel and basic attributes", Feb. 2013.
- [4] Australian Institute of Health and Welfare, "Metadata Online Registry (METeOR)", <http://meteor.aihw.gov.au/>. [accessed: Nov. 21, 2019]
- [5] US National Cancer Institute, "Cancer Data Standards Repository (caDSR)", <https://wiki.nci.nih.gov/display/caDSR/caDSR+Wiki>. [accessed: Nov. 21, 2019]
- [6] Common Education Data Standards, "The Status of State Data Dictionaries", <https://ceds.ed.gov/>. [accessed: Nov. 30, 2019]
- [7] U. S. Department of Homeland Security (DHS) and U.S. Department of Justice (DOJ), "US National Information Exchange Model (NIEM)", <http://www.niem.gov/>. [accessed: Nov. 21, 2019]
- [8] Canadian Institute for Health Information (CIHI), "CIHI Data Dictionary", <https://www.cihi.ca/>. [accessed: Nov. 21, 2019]
- [9] Agency for Healthcare Research and Quality (AHRQ), "US Health Information Knowledgebase (USHIK)", <https://ushik.ahrq.gov/>. [accessed: Nov. 21, 2019]
- [10] US National Institutes of Health (NIH), National Library of Medicine (NLM), "NIH-NLM Common Data Elements (CDE) Repository", <https://cde.nlm.nih.gov/cde/search>. [accessed: Nov. 21, 2019]
- [11] US Department of Justice, "Global Justice XML Data Model (GJXDM)", <https://it.ojp.gov/initiatives/gjxdm>. [accessed: Nov. 21, 2019]
- [12] D. Jeong, H. Jeong, J. Kim, K. H. Jeon, and S. Shin, "Development of the SQL/MM Standard for Metadata Registries", Journal of the Korea Society of Computer and Information, Vol. 15, No. 9, pp. 9-18, Sep. 2010.

- [13] S. Lee, D. Jeong, H. Jung, and D. K. Baik, "Design and Implementation of Sensor Registry Data Model for IoT Environment", KIPS Transactions on Software and Data Engineering (KTSDE), Vol. 5, No. 5, pp. 221-230, May 2016.
- [14] H. Jung, D. Jeong, S. Lee, B. W. On, and D. K. Baik, "A Network Coverage Information-Based Sensor Registry System for IoT Environments", Sensors, Vol. 16, No. 8, Paper No. 1154, Jul. 2016.
- [15] J. Davies, J. Gibbons, S. Harris, and C. Crichton, "The CancerGrid experience: Metadata-based model-driven engineering for clinical trials", Science of Computer Programming, Vol. 89, Part B, pp. 126-143, Sep. 2014.
- [16] International Organization for Standardization, "11179-6:2015 Information Technology - Metadata Registries (MDR) - Part 6: Registration", Aug. 2015.
- [17] Wikipedia, "Collective Intelligence", https://en.wikipedia.org/wiki/Collective_intelligence. [accessed: Nov. 21, 2019]
- [18] Wikipedia, "Blockchain", <https://en.wikipedia.org/wiki/Blockchain>. [accessed: Nov. 29, 2019]

저자소개

정 동 원 (Dongwon Jeong)



1997년 2월 : 군산대학교
컴퓨터과학과(이학사)
1999년 2월 : 충북대학교
전자계학과(이학석사)
2004년 2월 : 고려대학교
컴퓨터학과(이학박사)
2005년 4월 ~ 현재 : 군산대학교

교수(현 소프트웨어융합공학과)

관심분야 : 데이터베이스, 데이터표준화, 시맨틱 서비스,
빅데이터, 사물인터넷, 지능형 융합 서비스