

KARMA를 활용한 ISAD(G)와 CIDOC CRM 연계에 관한 탐색적 연구*

An Exploratory Study on Linking ISAD(G) and CIDOC CRM Using KARMA

박 지 영 (Zi-young Park)**

목 차

- | | |
|----------------------|--------------------------|
| 1. 서론 | 3. 기록물 기술정보의 재구조화와 연계 |
| 2. 이론적 배경 | 3.1 ISAD(G) 기반 기술 정보의 분석 |
| 2.1 데이터 모형의 유형과 특성 | 3.2 CIDOC CRM 온톨로지 맵핑 |
| 2.2 도메인 온톨로지의 확장 | 3.3 시범적 맵핑 결과 및 시사점 |
| 2.3 연계 플랫폼으로서의 KARMA | 4. 결론 |

<초 록>

기록물 기술은 기록관리 전문가의 창작과 편집 과정이며, 그 결과로 생산된 기록 레코드는 이용자가 기록을 효과적으로 이용하기 위한 밑거름이 된다. 그리고 기록물 기술을 위해 보급되는 다양한 기술 표준은 기록 레코드의 내용과 구조를 규정하는 중요한 도구에 해당된다. 본 연구에서는 기록물 기술표준의 특징을 데이터 모형의 구조 측면에서 분석하고, 기록물 기술표준인 ISAD(G)와 문화자원 분야의 도메인 온톨로지인 CIDOC CRM을 시범적으로 연계하였다. 연계에 앞서 ISAD(G)를 데이터 구조 측면에서 분석했으며, 분석 결과를 CIDOC CRM에 맵핑하였고, 연계 플랫폼으로는 KARMA를 활용하였다. 그 결과 온톨로지 기반의 맵핑 방식에 대한 이해와 이벤트 기반 CIDOC CRM에 대한 인식 제고 및 CRM 연계온톨로지의 개발, ISAD(G)의 기존 기술요소 값에 대한 재구조화가 필요하다는 판단을 하였다.

주제어: 기록물 기술표준, 도메인 온톨로지, 메타데이터, 국제기록물기술규칙(ISAD(G)),
박물관 분야 개념모형(CIDOC CRM), 시맨틱웹, 온톨로지 맵핑

<ABSTRACT>

Archival description is considered as a creation and curation process, and the results of the descriptive records can be used for archival information service. Therefore, various archival descriptive standards provide essential guidelines for establishing a semantic and synthetic structure of the archival records. In this study, the structural aspects of the archival descriptive standards were analyzed and an experimental mapping between General International Standard Archival Description (ISAD(G)), the archival standard, and CIDOC Conceptual Reference Model (CIDOC CRM), the domain ontology of cultural heritage field was performed. The data structure of ISAD(G) is examined in advance and mapping was performed using Karma as a platform. It was thus concluded that there is a need to understand the ontology-based mapping method and the event-focused domain ontology. Moreover, developing a CIDOC CRM-compatible archival ontology and restructuring the legacy ISAD(G) are needed.

Keywords: standards for archival description, domain ontology, metadata, ISAD(G), CIDOC CRM, semantic web, ontology mapping

* 본 연구는 한성대학교 교내학술연구비 지원 과제임.

** 한성대학교 응용인문학부 부교수(zgpark@hansung.ac.kr)

■ 접수일: 2018년 4월 30일 ■ 최초심사일: 2018년 5월 8일 ■ 게재확정일: 2018년 5월 24일

■ 한국기록관리학회지 18(2), 189-214, 2018. <<http://dx.doi.org/10.14404/JKSARM.2018.18.2.189>>

1. 서론

기록물 관리자는 기록과 이용자 간의 중개자의 역할뿐 아니라 기록관리 업무를 수행하기 위한 메타정보의 생산자의 역할도 수행한다. 중개자를 '큐레이터'로 생산자를 '크리에이터'로 구분할 수는 있지만, 정보량이 방대해 질수록 중개자로서의 큐레이터와 생산자로서의 크리에이터의 경계는 모호해 진다(Bhaskar, 2016). 기록물 관리자가 기록정보를 관리하기 위해 메타정보를 생산한다는 점에 주목한다면, 우리는 모든 창작자들이 갖는 고민을 기록물 관리자도 하고 있다는 것을 짐작할 수 있다. 즉, 어떻게 해야 좋은 메타정보를 만들어 활용할 수 있는가에 대한 고민이다. 그 결과로서 기록관리 분야에서는 기록물을 정리하고 기술하는데 필요한 모범사례와 가이드라인, 기술표준 등을 개발하여 적용해 왔다. 그리고 이러한 노력은 개별 기관 차원의 내부 지침에서 국제적인 합의를 통한 표준 제정에 이르기까지 다양하게 나타났다. 현재 기록물 기술규칙으로 널리 보급된 국제기록기술표준(General international standard archival description, ISAD(G))도 개별 기록물 관리기관에서 인벤토리 형식으로 작성하던 기록물 기술 정보를 표준화하고 공유하기 위해 제정되었다.

또한 기록전문가들은 기록관리 환경의 변화를 수용하여 메타정보의 내용과 구조를 주기적으로 개선해 왔다. ISAD(G)의 후속 표준으로 개발된 국제기록전거레코드기술표준(International Standard Archival Authority Record for Corporate Bodies, Persons and Families, ISAAR(CPF))은 기록물 메타데이터가 관계형 데이터

베이스와 같이 다중 레코드를 지원할 수 있는 구조로 저장되고 활용되는 정보환경을 수용한 결과이다(박지영, 2016; 2017b). 그리고 최근에 ICA는 분산된 기존의 표준을 통합하고 시맨틱 웹 환경에 적합한 기록물 메타정보를 생산하고 공유하기 위해 'RiC-CM'과 'RiC-O'를 개발하고 있다(ICA, 2016; ICA EGAD, 2016). RiC 모형의 개발도 기록관리 및 정보관리 환경의 변화에 적극적으로 대응하기 위한 변화의 일환이다.

그런데 기술표준의 변화와 함께 자주 언급되는 '개체-관계 구조의 개념모형'이나 '트리플 구조의 온톨로지'와 같은 개념이나 메타정보의 구조는 전통적인 기술방식을 적용하고 있는 기관의 담당자들에게 매우 생소하게 느껴질 수 있다. 특히 문화유산 기관 간의 협력을 위한 데이터 통합을 측면을 고려하면, 기술표준의 변화 범위는 기록물 관리기관의 경계를 넘어서기도 한다. 따라서 본 연구에서는 우리에게 친숙한 단일 레코드 형식의 기록물 기술정보가 시맨틱 웹 환경에서 어떻게 재구조화될 수 있는지, 기존의 표준이 변경된다면 기존의 방식으로 생산된 메타정보가 변환되는 과정은 어떠한지를 실험적으로 제시하고자 하였다. 시맨틱 웹 환경에 가능한 메타정보는 어떤 특징을 가지고 있는지를 전통적인 형식의 데이터에서 출발해서 변형하는 과정에서 접근한다면, 새로운 방식으로 메타정보를 생산하고 활용하기 위한 이해를 높이는 데 부담감을 줄일 수 있을 것이다. 또한 기존의 방식에서 변화하는 것과 변하지 않는 것을 구분해서 판단하는 데에도 도움이 될 것으로 기대할 수 있다.

이를 위해 시맨틱 웹 환경에서 상이한 여러 데이터를 통합할 수 있는 플랫폼인 KARMA를

활용하였다. 이를 위해 본 연구에서는 데이터 모형의 유형을 4가지로 구분하여 살펴보고, 단일 레코드 모형인 ISAD(G)를 그래프 기반의 모형인 CIDOC-CRM과 시범적으로 연계하고자 한다. 기록물 기술표준과 CIDOC-CRM을 연계하는 것 자체가 목적이려면, ISAD(G)보다는 EAD (Encoded Archival Description)나 RiC-CM을 활용하는 것이 더 적합할 것이다. 그러나 현재 EAD기반 기술은 우리의 기록 환경에 널리 보급되지 않았고, RiC-CM도 초안만 공개된 시점이므로 ISAD(G)를 선정하였다. 그리고 데이터의 변환과 연계에는 시맨틱웹 데이터 통합 플랫폼인 KARMA를 활용하였다. KARMA는 그래픽 인터페이스를 통해 기존 데이터의 변환과 맵핑 과정을 직접 확인할 수 있고, 변환된 데이터를 웹에 발행할 수 있는 형식으로 내보내는 기능도 지원하기 때문에 연계플랫폼으로 적합하다고 판단하였다.

2. 이론적 배경

2.1 데이터 모형의 유형과 특성

이 절에서는 메타 정보의 구조적 측면을 데이

터 모형의 특성을 통해 살펴보고자 한다. Hooland & Verborgh(2014)는 정보의 구조적인 표현 방식을 4가지로 구분하여 설명하였는데, 이 절에서는 이를 기록물 기술 정보의 예시와 함께 정리하였다.

2.1.1 단일 레코드 형식의 구조

단일 레코드 형식의 데이터(tabular data)는 매우 친숙하며 데이터 간의 관계를 직관적으로 이해할 수 있는 장점이 있다. 이 형식에서는 가로 행과 세로 열이 끊임없이 반복되며 특정 값을 표시하는데, 데이터 간의 경계는 줄 바꿈이나 쉼표 등을 통해 표기할 수도 있다. 일반적으로 맨 윗줄은 각 열의 제목을 나타내며, 동일한 열에 속한 칸에는 동일한 정보가 입력된다. 이 때, 각 열에 입력되는 데이터 간의 관계는 암묵적으로 나타나며, 해당 레코드의 구조를 이해할 수 있는 내부 관계자들의 배경지식을 통해 식별된다. 단일 레코드 형식의 데이터 구조는 단일 개체 중심의 기록물 기술 표준에서도 흔히 볼 수 있다. ICA의 기술표준 중 하나인 ISAD(G)의 경우에도 ‘기록’이라는 하나의 개체를 각각의 기술영역과 기술요소로 구분하여 표현하도록 규정하고 있는데, 기술요소와 그 값은 <그림 1>의 예시와 같이 단일 레코드 형식으로 입력될 수 있다.

식별영역	참조코드	AG29/S2	
	제목	교류협력	
	생산시기	1977~2015	
	기록물 유형별 수량	공문서 55권, 정부간행물 205권	
	기술계층	現기술계층	기록물계열
		상위계층	통일부 기록물군
하위계층			



참조코드	제목	생산시기	수량	계층
AG29/S2	교류협력	1977-2015	공문서 55권, 정부간행물 205권	기록물 계열
...
...
...

<그림 1> 기록물 기술정보를 표 형식으로 구조화한 예시
(기록물 예시 출처: 국가기록원 웹사이트 <www.archives.go.kr>)

2.1.2 다중 레코드의 관계형 구조

관계형 모델(relational model)은 단일 레코드를 연결하여 확장한 다중 레코드 형식에 해당된다. 하나의 표 안에 다른 표가 포함되어 있는 것과 같이 개별 레코드가 다른 레코드와 연결될 수 있다. 연결키 역할을 하는 열을 통해 여러 개의 레코드가 참조될 수 있는데, 여러 개의 레코드를 필요에 따라 조합해서 활용하거나 연계된 인터페이스를 확인하기 위해서는 별도의 관리시스템이 필요하다. 즉, 다중 레코드 구조의 데이터는 저장된 데이터 그 자체가 아니라, 여러 개의 분산된 테이블을 이용자가 원하는 인터페이스로 조합해서 처리하고 보여줄 수 있는 시스템과 관리자가 반드시 필요하다. ICA에서는 ISAD(G)의 후속표준으로서 ISAAR(CPF)를 발표하고, 이 두 표준을 연계하여 기술하는 방안을 제안하였다(ICA, 2004). 그 결과 ISAD(G) 기반의 단일 레코드 형식의 메타 정보와 ISAAR(CPF)를 바탕으로 구축한 메타정보는 정규화되어 <그림 2>와 같이 별도의 레코드로 분리되었다.

2.1.3 중첩된 계층형 레코드 구조

중첩된 계층형 구조는 HTML이나 XML과

같은 마크업 언어(meta-markup languages)로 된 레코드의 특징으로서 계층관계는 상위 태그 아래에 하위 태그를 다시 삽입하고 들여쓰으로써 간단하게 표현할 수 있다. 이러한 구조에서는 단순한 데이터 요소는 하위 태그를 삽입하지 않고 직접 기술하고, 더 복잡한 데이터 요소는 하위 태그를 여러 번 삽입하여 상세히 기술할 수 있다. 앞서 살펴본 다중 레코드 구조(관계형 구조)에서는 하나의 레코드에 속한 기술요소를 더 상세히 구조화시키고 싶을 때 다른 레코드를 추가하여 연결했었다. 그러나 마크업 언어로 된 레코드에서는 단일 레코드 내에서 하위 계층의 태그를 추가하여 특정 기술요소만을 상세히 표현할 수 있다. <그림 3>은 EAC(Encoded Archival Context) 기반 XML 형식으로 표현된 '박경리' 전거 정보의 일부분이다(박지영, 2014: RAMP Websites, 2018). 이러한 구조는 다계층 기술에 적합하며, 기록물군부터 아이템까지 한 레코드에 상세히 기술할 수 있는 장점도 있다. 앞서 살펴본 바와 같이 단일레코드에 ISAD(G) 기반 메타정보를 기술한다는 것은 실제로는 기록물군 중심의 기술만 가능하다는 의미이다. ISAD(G)는 긴 두



<그림 2> 기록물 기술정보와 전거 정보를 관계형으로 구조화한 예시 (기록물 예시 출처: 국가기록원 웹사이트 <www.archives.go.kr>)

루마리 형식의 매체에 기록물 전체를 기술하는 인벤토리 기술방식을 표준으로 발전시킨 셈이다. 따라서 ISAD(G)를 바탕으로 기술된 기록물군 레코드 자체를 EAD로 변환만 한다면, 마찬가지로 기록물군 중심의 EAD 레코드가 생성된다. 그러나 EAD 구조를 활용한다면 한 레코드에 기록물군과 시리즈, 파일, 개별자료를 별도로 기술하는 방법도 가능하다.

```
<?xml version="1.0" encoding="UTF-8"?>
<eac-cpf xsi:schemaLocation="urn:isbn:1-931666-33-4 http://eac.staatsbibliothek-berlin.de/schema/cpf.xsd"
xsi:link="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="urn:isbn:1-931666-33-4">
  <control>
    <cpfDescription>
      <identity>
        <entityType>person</entityType>
        <nameEntry xml:lang="en" scriptCode="Latin">
          <part>Park, Kyung-ni</part>
        </nameEntry>
      </identity>
      <description>
        <dates>
          <dateRange>
            <fromDate standardDate="19261028">19261028</fromDate>
            <toDate standardDate="20080505">20080505</toDate>
          </dateRange>
          <levelDates>
            <occupation>
              <term></term>
            </occupation>
            <localDescription localType="gender">
              <term>female</term>
            </localDescription>
            <localDescription>
              <term></term>
            </localDescription>
            <biogHist>
              <term></term>
            </biogHist>
          </description>
          <relations>
            <cpfDescription>
              <eac-cpf>
```

<그림 3> EAC-CPF 형식의 전거레코드 예시
(출처: RAMP 웹사이트
(<https://tools.wmflabs.org/ramp/all-items.php>))

2.1.4 그래프형 데이터

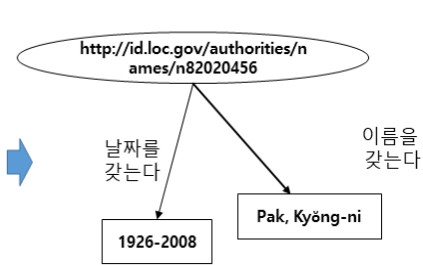
그래프형 데이터는 논리적인 측면의 구조로서 계층형 데이터와 같이 대개 XML 기반의 레

코드로 작성된다. 그러나 자원기술구조(Resource Description Framework, RDF)와 같이 각 태그가 주어-술어-목적어 구조를 이루는 트리플로 표현되며, 노드와 아크를 이용하여 데이터 간의 관계를 상세히 표현할 수 있다. <그림 4>는 미국국회도서관에서 제공하는 '박경리'의 전거데이터로서 RDF/XML 문서인데, 계층형 속성을 지니며 하나의 문서에 모든 트리플이 입력된다는 점에서는 단일 레코드의 일종이다. 그러나 그래프형의 트리플 레코드는 다른 트리플 레코드와 연계되어 트리플 문서 클라우드를 이룰 수 있다는 점에서 확장성이 크다. 대량의 트리플 문서가 연계된 링크드 데이터(linked data) 집합의 대표적인 사례로는 LOD 클라우드가 있다(Insight Centre for Data Analytics, 2018).

2.2 도메인 온톨로지의 확장

전통적으로 박물관이나 기록관, 도서관과 같은 문화유산 관리기관에서는 개별 기관이 소장한 문화자원의 특성에 특화된 메타 정보의 구조화 방식을 채택하여 발전시켜왔다. 기록물 관리

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  <ma:PersonName rdf:about="http://id.loc.gov/authorities/names/n82020456"
  xmlns:ma="http://www.loc.gov/mads/rdf/v1#">
  <rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1#Authority"/>
  <ma:authoritativeLabel xml:lang="en">Pak, Kyŏng-ni, 1926-2008</ma:authoritativeLabel>
  <ma:elementList rdf:parseType="Collection">
    <ma:FullNameElement>
      <ma:elementValue xml:lang="en">Pak, Kyŏng-ni,</ma:elementValue>
    </ma:FullNameElement>
    <ma:DateNameElement>
      <ma:elementValue xml:lang="en">1926-2008</ma:elementValue>
    </ma:DateNameElement>
  </ma:elementList>
```

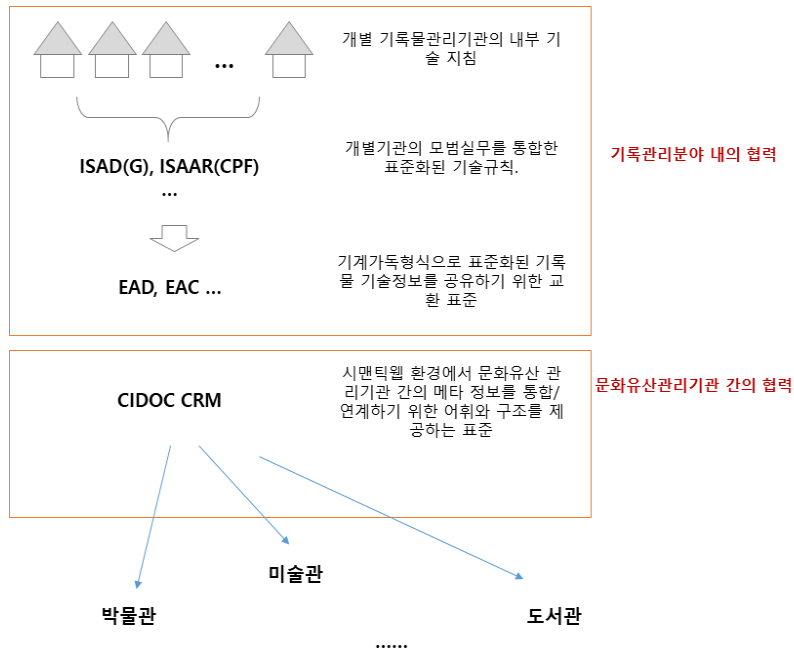


<그림 4> RDF/XML (MADS and SKOS) 방식으로 표현된 그래프형 데이터 예시
(전거데이터 출처: 미국국회도서관 (<http://id.loc.gov/authorities/names/n82020456.html>))

기관에서도 소장 기록에 대한 관리와 서비스에 필요한 정보를 별도로 구축해 왔으며, 정보의 교환과 공유 환경이 발달할수록 복수의 기록물 관리기관이 협력하여 기록물에 대한 메타정보를 표준화하고 교환하는 협력체계가 확장되었다. 이를 간략히 도식화하면, <그림 5>와 같은데, 개별 기관에서 상이하게 관리하던 기록에 대한 기술정보가 ISAD(G)를 비롯한 기술지침의 제정 이후 표준화된 기술영역과 요소로 구축되어 왔다(ICA, 1999). 그리고 정보환경의 변화로 기록물 기술정보의 공유와 교환에 대한 요구가 커지자 기계가독형식으로 기록정보나 전자정보를 교환하기 위해 XML 기반의 마크업 언어로 표현된 EAD(Encoded Archival Description)와 EAC(Encoded Archival Context)도 개발되었다(박지영, 김태수, 2007; Library of Congress,

2018).

물론 기록물 자체나 생산자에 대한 기술 요소와 값을 표현하는데 EAD와 EAC와 같은 인코딩 표준을 적용하면, ISAD(G)나 ISAAR(CPF)에 없던 새로운 기술 정보가 추가되는 것은 아니다. ISAD(G)나 ISAAR(CPF)가 어떤 내용을 기술할 것인지를 제시하는 내용표준이었기 때문에, 동일한 내용을 어떤 그릇에 담아 운반하고 공유할 지에 대한 고민을 EAD와 EAC에 담은 것이다. 기계가독형식으로 데이터가 이동한다는 것은 메타정보의 복제와 공유가 매우 촉진된다는 것을 의미한다. 메타정보가 네트워크를 통해 시간과 거리의 제약을 받지 않고 빠르게 전송될 수 있으며, 전달받은 정보의 구조와 내용을 기계가 바로 해독할 수 있기 때문이다.



<그림 5> 기록물 기술정보의 표준화와 협력체계의 범위

웹이 온라인 네트워크에서 의미까지 이해하고 처리할 수 있는 시맨틱웹에서는 메타정보의 공유 범위가 더욱 확장될 수 있다. EAD나 EAC가 기록물 관리기관 간의 협력을 주요 범위로 하여 개발되었다고 한다면, CIDOC CRM은 기록관뿐 아니라 박물관이나 도서관, 미술관 등 가능한 모든 문화유산관리기관 간의 협력까지 가능하도록 확장된 개념이다(박지영, 2008; 현문수, 2015). 이제 기록관에서 구축한 기록물 메타정보와 박물관에서 구축한 메타정보, 도서관에서 구축한 메타정보가 CIDOC CRM이라는 표준화된 데이터 모형으로 변환될 수 있다. 도서관과 박물관의 데이터를 연계하기 위한 CIDOC CRM의 플러그인 온톨로지인 FRBRoo가 개발된 것도 주목할 만하다(박지영, 2017c; 현문수, 2014). 기록물 관리기관의 정보 협력망이 개별 기록관에서 기록관리 분야 전반으로 확장되었을 뿐 아니라, 유관 기관까지도 가능한 환경이 갖추어 진 것이다.

그런데 메타정보의 협력에는 정책적인 측면과 정보 인프라 외에 중요한 조건이 갖추어 져야 한다. 즉, 통합하고 연계하고자 하는 메타정보의 내용과 구조 간의 합의가 필요하다. 분야마다 동일한 용어를 서로 다른 의미로 사용하거나, 기술요소 간의 관계를 상이하게 정의한다면 체계적인 연계가 불가능할 것이다. 이에 메타정보가 표현되는 구조의 유형과 유형별 특징 및 차이점을 이해할 필요가 있다. Stasinopoulou et al.(2007a; 2007b)은 상이한 유형의 데이터를 통합하기 위해 EAD의 기술요소와 하위요소를 CIDOC-CRM에 맵핑하는 프로젝트를 수행하였다. 맵핑 과정에서는 대상(object) 중심으로 정보를 기술하는 기록관리 분야의 메타데이터에

이벤트 중심으로 구조화된 CIDOC CRM을 적용하는 것을 이슈로 꼽았다. 예를 들어, CIDOC CRM에서는 오직 '활동'(activity)만이 '행위자'(actor)와 '일자'(date) 속성을 지닐 수 있었기 때문에, '문서'(document)'에 행위자와 날짜를 부여하는 EAD의 기술구조를 정확히 맵핑할 수는 없었다. 생산자의 경우에도 CIDOC CRM에서는 '생산'이라는 이벤트를 통해 기록과 행위를 연결해야 한다. Hong et al.(2005)은 중국의 문화유산 자원 정보를 통합하는데 필요한 중간 매개 온톨로지로서 CIDOC CRM을 분석하였다. 논문 본문에서는 기존의 메타데이터가 CIDOC CRM에 맵핑되는 과정이 상세히 제시되지는 않았지만, 관계형 데이터베이스에 저장된 속성 값을 CIDOC CRM의 클래스와 속성과 연계하는 예시를 확인할 수 있었다.

2.3 연계 플랫폼으로서의 KARMA

2.3.1 KARMA 개요

데이터 처리와 시각화를 위한 많은 응용프로그램이 개발되고 있지만, 대부분의 프로그램은 특정 형식으로만 구조화된 데이터를 처리하도록 설계되어 있다. 비정형 데이터를 다루는 연구의 경우에도 처리 과정의 상당 부분은 비정형 데이터에서 불필요한 정보를 걸러내는데 할애하게 된다. 데이터의 시각화와 연계를 통해 원 정보에서 부가가치를 높이는 연구도 활발히 진행되고 있는데, 이 경우에도 데이터의 재구조화에 많은 비용과 시간이 소요된다(Knoblock & Szekely, 2013; 2015). 이에 KARMA는 데이터의 통합 플랫폼으로서 상이한 출처의 데이터를 구조적으로 연계하는데 필요한 2가지 중요한 기

능을 제공한다. 첫 번째는 KARMA에 반입된 복수의 데이터를 동일한 인터페이스로 배포하는 통합(integration) 기능이고, 두 번째는 원래의 데이터 구조를 공통의 도메인 온톨로지에 적합하도록 정렬하는 재구조화(restructuring) 기능이다. 특히 재구조화 과정에서는 기존 데이터의 구조와 의미를 분석하고, 그 결과를 도메인 온톨로지의 클래스와 속성, 관계 등으로 변환할 수 있어야 한다. 기존의 데이터가 도메인 온톨로지에 맞도록 재구조화되는 순간 데이터 통합은 자연스럽게 이루어진다고 볼 수 있다. 따라서 본 연구에서도 기존의 데이터가 도메인 온톨로지에 맞도록 재구조화되는 과정에 주목하였다. 이 점에서 KARMA는 단일 파일이나 관계형 데이터베이스 등에 스프레드시트 형식으로 표현되어 있는 데이터가 시맨틱 데이터 모형에 적합하도록 변환되는 과정을 확인하고 그 결과를 공유하는 데에도 활용될 수 있다. 도서관이나 기록관에서 생산된 메타 정보가 대부분 표 형식으로 구조화되고 활용되고 있는데, 기존의 메타데이터를 시맨틱웹 환경에서 연계 가능한 방식으로 재구조화되는 과정을 단계별로 살펴볼 수 있는 것이다. KARMA가 오픈소스로 개방되어 있고, 설치와 이용을 위한 가이드라인이 잘 정비되어 있다는 점은 다양한 데이터를 대상으로 KARMA의 활용 범위가 확대되는 중요한 요인으로 간주된다.

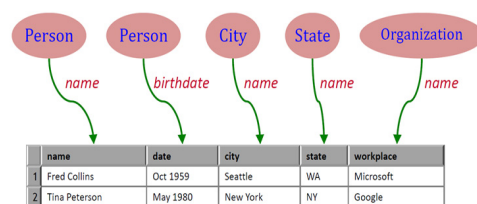
2.3.2 기존 데이터의 재구조화

KARMA는 도메인 온톨로지를 중심으로 데이터를 통합한다. 이 과정에서 이루어지는 기존 데이터의 재구조화 과정은 전통적인 방식의 메타정보를 시맨틱웹 환경에 적합하도록 변환

하는데 시사점을 줄 수 있다(Knoblock, 2017).

1) 의미유형의 선정

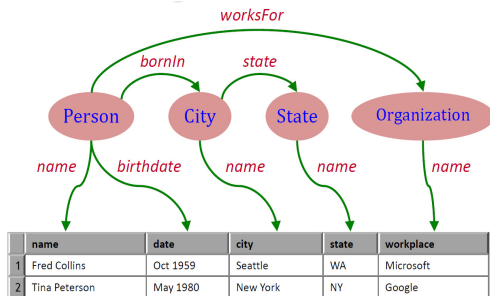
KARMA에서 기존의 데이터를 시맨틱 데이터로 변환하기 위해서는 우선 개별 컬럼이 갖는 의미유형을 선정해야 하는데, 이 구분은 이후 온톨로지와 같은 모델에서 클래스와 속성을 분석하는 데이터로도 활용된다. <그림 6>의 사례를 보면, 'name'은 인명으로 해석하여 'Person'이라는 의미유형의 'name'으로 분석하였고, 'date'는 'Person'의 'birthdate'로 분석하였다. 이제 두 컬럼이 특정 인물의 이름과 생년월을 표현한다는 것이 명시적으로 분석되었다. 다음 컬럼인 'city'는 인물이 아니라 도시명이 값으로 들어가므로 'City' 유형의 'name', 다음 컬럼은 각각 'State' 유형의 'name', 'Organization' 유형의 'name'으로 분석되었다. 단, 의미유형은 고정된 단일 값이 아니며, 데이터를 변환하고 통합하기 위해 설정한 뼈대가 되는 기본 도메인 모형에 따라 달라질 수 있다.



<그림 6> 제목줄 컬럼의 의미유형 구분 (Knoblock, 2017, p. 14)

2) 의미유형별 요소 간의 관계

의미유형을 분석하여 컬럼명을 클래스와 속성으로 변환한 후에는 개별 클래스 간의 관계를 밝혀주어야 한다. <그림 7>의 사례를 보면,



<그림 7> 유형별 요소 간의 관계 설정 (Knoblock, 2017, p. 15)

‘Person’과 ‘City’, ‘State’, ‘Organization’ 간의 관계가 나타나 있다. 표 형식에서 생략되어 있던 관계를 명시적으로 표현해 줌으로써 데이터의 작성자나 작성기관 외의 제3자가 이해하고 연계하기 용이한 메타정보를 구축할 수 있게 된다. 이 때 원 정보원에서 데이터 유형을 식별해 내고 관계를 설정하는 것은 데이터의 통합과 연계를 위한 플랫폼에서 선정한 도메인 온톨로지를 기준으로 하게 된다. 동일한 온톨로

지를 바탕으로 분석되고 변환된 데이터는 해당 온톨로지를 공유하는 모든 데이터와 효과적으로 통합되고 연계될 수 있다. KARMA는 통합과 연계의 바탕이 되는 모델을 시맨틱웹 환경에 적합한 모델로 선정한다.

2.3.3 원 데이터와 도메인 모형의 연계

KARMA는 원 데이터를 반입하여 설정된 도메인 모형에 맞도록 변환할 수 있는 기능을 제공한다. <그림 8>과 같이 각각 반입된 상이한 기관의 문화자원 기술 정보는 하나의 도메인 모형을 기준으로 변환된다. 따라서 공통의 도메인 모형에 따라 변환된 여러 기관의 상이한 메타정보가 통합되는 것이다. KARMA는 구축된 모형을 다시 웹을 통해 발행할 수 있도록 지원하며, 그래픽 유저 인터페이스를 통해 맵핑 과정과 결과를 모두 확인할 수 있다(Szekely, 2015: University of Southern California, 2016).

도메인 모형과의 맵핑 정보 (온톨로지 맵핑)

반입된 원 정보의 기술요소 값 (제목줄)

반입된 원 정보의 데이터의 값

<그림 8> KARMA의 원 데이터와 도메인 모형의 맵핑 화면(Knoblock, 2017, p. 20)

3. 기록물 기술정보의 재구조화와 연계

3.1 ISAD(G) 기반 기술 정보의 분석

이 절은 CIDOC CRM과의 맵핑을 위한 ISAD(G)의 사진 분석 단계로서 기록물 기술정보를 온톨로지 관점에서 분석하였는데, 이 과정에서 ISAD(G) 기술요소와 RiC-CM 요소의 비교 결과도 참고하였다(박지영, 2017a, pp. 114-115). 가독성을 위해 클래스와 속성명은 국문으로 표기하였다. 전체 분석 결과는 <부록 1>에서 확인할 수 있으며, 분석결과를 온톨로지 모델링과 관련하여 영역별 이슈로 정리하면 다음과 같다.

3.1.1 식별영역의 분석

식별영역의 기술대상은 기술 단위(unit of description)로서 기록에 해당되는데, 전체 기록물 집합에서 기록물군이나 시리즈와 같이 기술을 위해 정리한 기록 전체 또는 부분 집합을 의미한다. 식별영역의 참조코드는 기술 단위를 식별하는 대표적인 값으로써 관련된 기술을 연계하기 위한 중요한 연계점이 될 수 있다. 따라서 참조코드에 URI를 부여하여 데이터 타입 속성이 아닌 오브젝트 타입 속성으로 변환해야 한다. 제목이나 일자는 모두 레이블 값으로 취급될 수 있는데, 일자의 경우 단일 날짜보다는 대개 기술 단위에 속한 기록의 전체 생산 기간 등을 나타내는 기간 범위로 표현된다. 기술계층은 기술요소값인 인스턴스를 유형화하여 미리 구조화할 수 있다. 기술단위의 제목을 모두 미리 예상해서 입력해 둘 수는 없지만, 기술계층은 대개 3~5계층 내외로 구성되며 계층

이름도 미리 지정되어 있기 때문이다. 기술단위의 규모와 유형은 물리적 특성과 논리적 특성이 모두 포함될 수 있어 텍스트 형식으로 서술되는 주기로 취급할 수 있다. 단, 매체 정보는 유형별로 구조화하여 기술계층과 같이 미리 인스턴스를 유형화할 수도 있다.

3.1.2 배경영역의 분석

배경영역에서는 기술단위의 생산과 관리에 관련된 개인이나 단체, 기록의 이관 이력 등에 대한 정보를 기술한다. 생산자는 행위자로서 개인이나 단체에 해당되는데, 온톨로지 변환을 위해서는 텍스트로 입력되는 레이블뿐 아니라 식별기호인 URI를 부가할 필요가 있다. 배경영역에서는 행정연혁/개인이력과 수집/이관의 직접적 출처와 같이 속성의 도메인이 기술단위가 아니라 행위자인 개인이나 단체가 중심이 되는 기술요소가 있다. 기록물 이력 정보나 수집/이관의 직접적 출처도 생산자나 기증자와 같이 기록물의 맥락을 제공하는 배경 정보에 속한다. 그런데 이와 같은 이력이나 출처 정보도 개인이나 단체, 날짜 등이 모두 제공되지만 구조화되지 않은 텍스트 형식으로 포괄적으로 기술되어 있다. 기록물 기술과 관련된 서술 정보의 일부를 식별기호나 레이블, 날짜 형식 등으로 구조화할 수 있다면 기술 정보 간의 연계가 더욱 용이할 것이다.

3.1.3 내용과 구조영역의 분석

기술 단위의 내용과 정리체계, 평가 결과에 대한 부분을 기술하는 영역은 기록의 검색과 관리에 중요한 단서를 제공한다. 범위와 내용 부분에서는 기술 단위와 관련된 시대나 지리 정보,

매체, 주제 정보 등이 제공되는데 이러한 유형은 온톨로지를 통해 효과적으로 연계될 수 있는 대표적인 유형의 정보이다. 그러나 ISAD(G) 기술요소의 특성상 풍부한 정보를 제공하기 위해 모두 서술식 텍스트를 입력하도록 되어 있는데, 온톨로지에서는 모두 주기로 취급되며 속성을 통해 개별 주기의 내용을 개별화하는 방안을 고려할 수 있다. 이 영역의 정보에서 URI를 부여하거나 값을 유형화하여 연계할 수 있는 일부 값을 분리하고, 나머지를 서술형으로 표현한다면 더욱 효과적인 온톨로지 모델링이 가능할 것이다.

3.1.4 접근과 이용환경 영역의 분석

접근과 이용환경에서는 기술단위의 이용과 검색을 제한하거나 지원하는 관련 규정이나 조건, 도구 등을 기술한다. 이 영역에서도 각 기술요소는 서술식으로 기술되어 텍스트 형태이므로 주기로 모델링된다. 그러나 자료의 이용과 관련된 법규나 이외 규정, 자료의 복제 조건 등은 대부분 미리 규정되어 있는 것이므로 인스턴스 값을 유형화하여 레이블과 같이 사용할 수 있다. 자료의 언어와 같은 요소도 언어 코드 값 등으로 구조화된 형식도 도입할 수 있다. 기록물 기술레코드의 특성상 서술식 텍스트가 필요하다면 해당 요소 값에서 필요한 일부 값만 추출하여 URI나 별도의 코드 값을 부여하는 것이 온톨로지 환경에서는 더 바람직하다. 특히 검색 도구는 인벤토리나 목차, 색인과 같은 다른 검색도구를 가리키는 내용이 기술되어 있다. 기록물 자체뿐 아니라 기록물 기술레코드나 별도의 인벤토리, 색인 등도 별도의 식별기호를 부여하여 식별되고 연계되도록 구조화하면 이용자의

검색 편의를 향상시킬 수 있다.

3.1.5 관련자료 영역의 분석

관련자료 영역은 특정 기술 단위와 관련된 다른 기술 단위의 위치나 존재여부를 제시한다. 기술 단위에 속한 기록이 원본일 경우에는 사본의 위치를, 사본일 경우에는 원본의 위치를 제시한다. 기술 단위를 바탕으로 발행한 출판물이 있다면 출판물 정보를 제시하고, 이 외에 어떤 관련성이라도 제시할 필요가 있다면 관련 기술단위 요소에 기재할 수 있다. 단, 기술 단위의 유형이 '원본'과 '사본'으로 구분되므로 모델링에서도 이를 구분할 수 있어야 한다. 그런데 관련자료 영역에서도 모든 기술요소 값이 서술형으로 제시되어 있어 텍스트 형태의 데이터 타입으로만 모델링되었다. URI와 같은 식별기호가 부여된 기술단위나 출판물은 오브젝트 타입으로 별도 관리할 필요가 있다. 관련자료 영역은 기술 대상 자체뿐 아니라 그것과 연관된 다른 정보에 대한 포괄적인 지식이 있어야만 기술할 수 있는 요소이다. 관련자료 영역의 정보를 활용한다면 기록물의 활용도도 제고할 수 있을 것이다.

3.1.6 추가설명 영역과 기술통제영역의 분석

추가설명 영역은 가장 포괄적인 부가 설명을 그 기술 내용으로 삼고 있어 전형적인 주기에 속한다. 기술통제영역은 기술요소를 속성으로 변환할 때, 도메인이 모두 기술 단위가 아니라 기록물 레코드라는 점이 특징이다. ISAD(G)에서 표현된 개인이나 단체가 대부분 기록 자체의 생산이나 소장 등에 관계되어 있다면, 기술통제영역은 기록물의 기술 과정에서 추가한

주기로 구성되어 있다. 기록 관리에 있어서 기록물 기술은 핵심적인 부분이며, 정리와 기술의 결과로 생산되는 기록레코드는 메타 기록이 된다. 따라서 기술통제영역에 포함된 생산자와 일자, 관련 규정 등의 정보는 메타기록의 식별 영역, 배경영역, 내용과 구조 영역 등으로 더욱 세분화될 수도 있다. RiC-CM 초안에 대한 검토 의견 중에는 아키비스트와 기록레코드에 관한 개체와 속성을 명확히 표현하지 못한 점에 대한 비판도 있었다. 기술통제영역은 온톨로지 모델링에서도 새로운 클래스와 속성으로 구조화되어야 하는 부분이다.

3.2 CIDOC CRM 온톨로지 맵핑

이 절에서는 앞 절에서의 분석결과를 바탕으로 ISAD(G)의 기술요소를 CIDOC CRM의 클래스와 속성에 맵핑하였다. 이 과정에서는 CIDOC CRM 온톨로지 명세서와 더불어 EAD와 CIDOC CRM 맵핑에 관련된 선행 연구도 참조하였다 (Bountouri & Gergatsoulis, 2011; ICOM/CIDOC, 2017; ISO, 2014; Stasinopoulou T. et al., 2007a; 2007b).

3.2.1 CIDOC CRM 맵핑 시의 고려사항

CIDOC CRM은 문화유산관리 분야의 메타 데이터를 연계하고 통합하기 위한 큰 우산과 같은 역할을 하는 상위의 개념체계이므로, 그 틀 안에서 ISAD(G)가 어떻게 자리 잡을 수 있는지를 중점적으로 분석하였다. 전체 분석 결과는 <부록 2>에서 확인할 수 있다. 또한 기록물 기술에 있어서 가장 핵심적인 개념은 기록이며, ISAD(G)에서는 기술 대상을 '기술 단위'로

명시하고 있다. 따라서 CIDOC CRM과의 맵핑을 위해서는 이 '기술 단위'가 CRM의 어느 클래스에 속하는지 먼저 식별해야 한다. 그리고 기록의 생산기관이나 기록물 레코드의 작성자와 같은 관련 요소들이 별도의 클래스로 정의되는지 속성을 통해 세분화되는지 판단해야 한다.

이와 관련하여 Stasinopoulou T. et al.(2007b)은 CIDOC path를 통해 EAD와 CIDOC CRM의 맵핑 방식을 일종의 경로와 같이 표현하였다. 1:1 맵핑이 어려운 메타데이터와 온톨로지의 맵핑으로 유용한 방식이라고 판단했기 때문이다. 그 결과, 기록의 물리적 측면과 정보적 측면, 그리고 기록을 기술한 기록레코드는 CIDOC CRM의 클래스에 다음과 같이 맵핑될 수 있다. 단 E22, E73, E33, E31 클래스는 모두 E71 Man-Made Thing의 하위 클래스이다. 따라서 <부록 2>에서 기술 단위인 기록을 맵핑할 때는 E71 클래스를 기재하였으며, ISAD(G)와 CIDOC CRM의 유사한 맵핑 유형에도 이를 활용하였다. 단, 본 연구에서 기술 단위와 기록레코드를 모델링할 때는 구분을 위하여 E71 클래스명 뒤에 1과 2를 추가하였다.

- 기술 단위(Unit of description)
 - ① 기록의 물리적 측면 - 인간 활동의 결과 생성된 물리적 객체
 - E22 Man-Made Object
 - ② 기록의 정보적 측면 - 물리적 특성과 독립적으로 인간의 기억을 담은 객체
 - E73 Information Object, E33 Linguistic Object
- 기록레코드(Archival record) - 기록물을 기술한 텍스트 형태의 객체

→ E31 Document, E33 Linguistic Object

이와 같은 맵핑 정보를 바탕으로 “기록물 기술 정보는 다른 기록물 기술정보의 일부분이며, 이 기술정보에는 기록물에 관한 사항이 문서화되어 있다. 그리고 기록물은 특정 언어 형식의 어떤 정보를 담고 있으며, 언어 형식의 정보는 제목과 같은 명칭으로 언급된다.”를 CIDOC CRM의 클래스와 속성을 이용해 표현하면 다음과 같다.

- [E31, E33] - P106 is composed of/forms part of - [E31, E33]
→ ‘기록레코드’는 다른 ‘기록레코드’의 일부분이다.
- [E31, E33] - P70 documents/is documented in - [E22]
→ ‘기록레코드’에는 ‘기술 단위’(기록)에 관한 사항이 문서화되어 있다.
- [E22] - P128 carries/is carried by - [E73, E33]
→ ‘기록’은 특정 ‘언어’ 형식의 어떤 ‘정보’를 담고 있다.
- [E33] - P67 refers to/is referred to by - [E41 Appellation]
→ ‘언어’ 형식의 객체는 어떤 ‘명칭’으로 언급된다.

3.2.2 식별 영역의 맵핑

식별영역은 참조코드와 같은 식별코드와 제목, 일자, 기술계층, 기술단위의 규모와 유형이 속해 있다. 우선 식별코드는 E42 Identifier와 연계할 수 있는데 CRM의 특성상 ‘식별자를 부여’하는 이벤트 클래스도 함께 사용하며, 제목

은 일종의 레이블로서 명칭에 해당되므로 E35 Title 클래스와 맵핑할 수 있다. 일자는 E52 Time-span 클래스와 맵핑되는데, 시간대의 표기방식에 따라 E49 Time Appellation 클래스와 함께 사용한다. 기술계층의 경우에는 E55 Type이라는 특수한 클래스를 사용할 수 있다. 기록물군이나 시리즈 등으로 구성되는 기록물 기술계층이나 정리체계에 관한 사항은 CIDOC CRM과 같은 상위 개념 온톨로지에서 구체적으로 명시되어 있지 않기 때문이다. 대신 CIDOC CRM에서는 추상적인 CRM의 개체에 개별 분야에서 사용하는 특정 어휘를 인스턴스로 입력할 수 있도록 E55 Type을 정의하였다(ICOM/CIDOC, 2017). 예를 들면, CIDOC CRM에서는 “John Smith”라는 발행자명을 다음과 같이 표현할 수 있다(Stasinopoulou T. et al., 2007b).

- [E33 Linguistic Object]-P94 performed-[E65 Creation Event]
- [E65 Creation Event]-P14 performed-[E39 Actor]-P2 has type/is type of-[E55 Type = “Publisher”]
- [E39 Actor]-P131 is identified by/identifies-[E82 Actor Appellation] = “John Smith”

이와 같은 방식으로 기록레코드에서 기술 단위가 속한 기술계층을 맵핑한다면 다음과 같이 표현할 수 있다.

- [E7 Activity (E87 Curation Activity)] - P19 was intended use of/was made for - [E71 Man-Made Thing(E31, E33)]

- 이벤트(큐레이션 행위-기록물 정리기술 등)로 기술단위가 생성됨
- [E71 Man-Made Thing(E31, E33)] - P103 was intended for/was intention of - [E55 Type="fonds"]
- 기술단위는 'fonds'이라는 유형을 대상으로 함

기술단위의 규모와 유형 요소는 E26 Physical Feature 클래스를 맵핑할 수 있는데, 물리적 특성 외에 논리적 특성까지 포함하는 경우에는 E73 Information Object, E33 Linguistic Object와 관련된 속성을 추가로 적용할 수 있다.

3.2.3 배경 영역 및 내용과 구조 영역, 추가 설명 영역의 맵핑

배경영역은 생산자명을 E39 행위자 클래스에 맵핑할 수 있다. 그런데 CIDOC CRM에서는 생산자를 '생산'이라는 이벤트와 연계하여 표현해야 한다. 또한 생산자는 그 자체의 식별기호와 명칭으로 구분되어 맵핑이 되므로, 아래와 같이 이벤트를 나타내는 E12 Production Event와 생산자의 명칭을 나타내는 E82 Actor Appellation 클래스를 활용하여 맵핑이 이루어진다.

- [E22] - P108 has produced/was produced by - [E12 Production Event]
- '기록물'은 '생산'이라는 이벤트를 통해 생산된다.
- [E12] - P14 carried out by/performed - [E39 Actor]
- '생산' 이벤트는 '행위자'를 통해 수행된다.

- [E39 Actor] - P131 is identified by/identifies - [E82 Actor Appellation]
- '행위자'는 '명칭'을 통해 식별된다.

그런데 생산자명을 제외한 행정연혁/개인이 력이나 기록물 이력, 수집/이관의 직접적 출처와 같은 요소는 모두 서술식 텍스트로 구성되어 있고 관련 행위자나 사진, 날짜 등이 모두 함께 기술되어 있다. 이와 같은 특징은 내용과 구조 영역 및 추가설명 영역에서도 동일하게 나타난다. 따라서 모두 주기에 해당되는 E62 String 클래스와 맵핑되거나, 요소 값을 모두 분해하여 다시 구조화시켜야만 특정 클래스와 직접적으로 맵핑할 수 있다. 이러한 서술식 텍스트는 ISAD(G)의 기술요소값의 많은 부분을 차지하고 있다. 사람이 육안으로 모두 읽고 이해하는 데에는 서술식 텍스트가 풍부한 편이 바람직하나, 일부 정보는 재구조화하여 기계도 함께 해석할 수 있도록 변환할 것을 고려해야 할 것이다.

3.2.4 접근과 이용환경 영역의 맵핑

접근과 이용환경 영역의 기술요소 중에서 접근환경과 이용환경, 물리적 특성과 기술적 요구조건은 서술식 텍스트로 여러 내용을 전달하고 있어 E62 String으로 포괄적으로 처리된다. 이와 같이 클래스 자체로 구체적인 모델링이 어려운 경우에는 속성 값을 이용할 수 있는데, 예를 들면 P3 has note 속성에 하위속성을 추가할 수 있다. 추가된 하위 속성은 속성 기호에 소수점 아래 번호를 추가하는데, 예를 들면 P3.1 has note about access control과 같다. 이와 같이 속성을 세분화하여 추가하면 동일한 문자열이라도 관계를 통해 차별화할 수 있다. 검색도

구 요소도 마찬가지로 E62 String과 맵핑되는데, 인벤토리나 색인 등의 검색도구는 서술식으로 기술하는 것 이외에 기록레코드와 같이 별도의 식별기호를 부여하여 링크 정보를 제공하는 것이 바람직할 것이다. 자료의 언어 요소는 E56 Language 클래스와 맵핑될 수 있는데, 이때는 속성 P72 has language/is language of 의 도메인이 E33 Linguistic Object로 구체화된다.

3.2.5 관련자료 영역의 맵핑

관련자료 영역은 사본의 경우 원본의 위치, 원본의 경우 사본의 위치, 이외 관련 출판물이나 기술 단위를 기재하므로 온톨로지 형식으로 메타 정보를 연계할 때 활용도가 높은 기술요소이다. 그러나 관련자료 영역의 요소 값도 서술식으로 입력되어 E62 String 클래스와 P3 has note 속성으로 연계할 수 있다. 관련자료 영역에서는 기술요소값이 대부분 다른 기록이나 출판물이 되므로 식별기호를 통해 오브젝트 타입 속성으로 연계할 필요가 있다.

3.2.6 기술통제 영역의 맵핑

기술통제 영역은 다른 영역과 달리 기술 단위인 기록 자체보다 기록레코드를 대상으로 하는 요소가 포함되어 있다. 기술담당자의 경우 생산자와 같이 E39 Actor로 맵핑이 되며, 관련 이벤트가 생산이 아니라 E87 Curation Activity란 점이 다르고, E82 Actor Appellation에서도 차별화된다. 규칙과 협약도 접근환경이나 이용 환경과 달리 기록레코드 작성을 위한 기술규칙이 주로 해당되는데, 서술식 텍스트로 기재되어 있어 E62 String과 맵핑된다. 기술일자는 식별

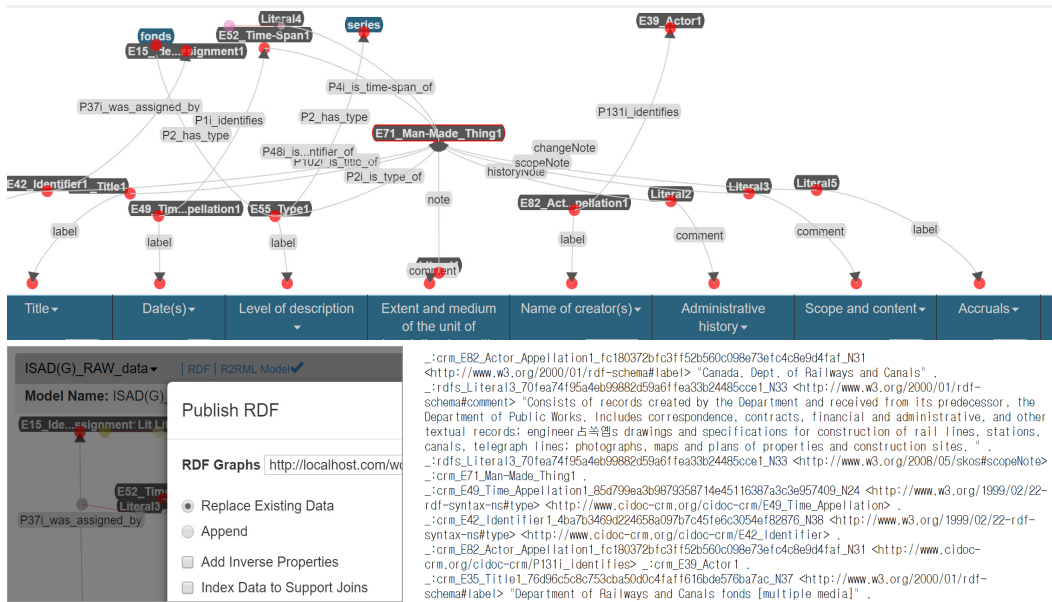
영역의 일자와 같이 E52 Time-span 및 E49 Time Appellation 클래스와 맵핑된다.

3.3 시범적 맵핑 결과 및 시사점

ISAD(G)와 CIDOC CRM을 분석하여 작성된 맵핑 정보를 KARMA 플랫폼에 반영하면 ISAD(G)의 기술요소가 CIDOC CRM의 클래스와 속성으로 재구조화되어 시각화되고 트리플 데이터로 발행된다. 시각화된 맵핑 화면과 RDF로 발행된 파일의 일부는 <그림 9>와 같으며, 맵핑 과정과 결과를 통해 기록물 기술 정보의 확장과 문화유산 관리기관 간의 데이터 통합을 위한 시사점을 도출할 수 있었다.

3.3.1 기술요소별 크로스 맵핑에서 온톨로지 기반 맵핑으로의 변화

문화자원 관리 분야에서 CIDOC CRM이 더 많이 보급 될수록 우리는 기술요소만을 1:1로 대응시키는 맵핑 방식의 한계를 체감할 것이다. 사실 복수의 기관이 작성한 상이한 기록물 기술 정보를 통합하는 것은 문화자원 관리기관이 언제나 지니고 있는 숙제이다. 그래서 각 분야에서 발행하는 표준 문서의 말미에 관련 표준과의 맵핑 정보가 수록되기도 했다. 예를 들면 ISAD(G) 식별영역의 '제목' 기술요소와 더블린코어의 'Title' 기술요소를 맵핑한 것이다. 이런 맵핑표를 요소 간 '크로스워크'(Crosswalk)라고도 불렀다. 메타데이터의 유형이 다양해질수록 맵핑표에는 더 많은 메타데이터가 일렬로 늘어서게 되었고, 이웃한 기술요소와 연계될 요소가 없는 경우에는 빈 칸을 두기도 했다. 그러나 앞으로는 두 세 개의 메타정보만을 연결하



〈그림 9〉 KARMA를 이용한 데이터 통합과 발행 예시 화면

고자 할 경우에도 개별 메타데이터 표준이 지닌 어휘와 구문을 모두 이해해야 할 것이다. 그리고 동일한 기술요소의 경우에도 식별기호가 입력되어 있는지 레이블이나 긴 설명이 입력되어 있는지 구분해야 할 것이다. 물론 과거에 빈번히 수행했던 크로스워크도 쉬운 작업이 아닐 때가 많았다. 그러나 앞으로는 단일 레코드에서 다중 레코드, 중첩된 계층형 레코드, 그래프형 트리플 레코드까지 다양한 데이터 모형이 동시에 존재할 것이다. 그리고 그 중에서도 가장 일반화된 개념을 상세히 풀어놓아야 하는 온톨로지가 시맨틱웹 환경에서 가장 적합한 맵핑 기준이 될 것이다. 온톨로지 기반 맵핑은 ISAD(G)와 CIDOC CRM의 경우와 같이 기존 구조의 분해와 재배치를 필요로 한다.

3.3.2 도메인 온톨로지 기반 맵핑과 연계 온톨로지의 개발

우리가 속한 분야를 흔히 ‘문화유산’이나 ‘문화자원’ 관리 분야라고 하는데, 현재 이 분야에서 가장 포괄적인 도메인 온톨로지는 ICOM/CIDOC의 CRM이다. 그런데 이 온톨로지는 문화자원과 행위자 사이에 ‘이벤트’를 삽입해서 표현하도록 안내하고 있다. 행위자가 기록을 생산할 때는 제작이나 발행과 같은 이벤트가 필요하고, 심지어 행위자의 명칭을 입력할 때도 이름을 선언하는 이벤트가 필요할 정도이다. 아키비스트가 기록레코드를 작성할 때는 큐레이션과 같은 이벤트가 필요하다. CIDOC CRM에서 이벤트 개념을 중시한 것은 문화자원 자체가 갖는 시간의 속성 때문이다. 문화자원은 오랜 시간에 걸쳐 여러 행위의 결과로 가치가 올라가는데, 이것은 문화자원 관리의 신뢰성을 나타내 주는

기준이 될 수도 있다. 그래서 CIDOC CRM에서는 일시적인(temporal) 것과 지속적인(persistent) 것도 구분한다.

그러나 시간이나 이벤트와 같은 개념은 연혁이나 이력정보를 중시하는 기록관리 분야의 기술표준에서도 주로 암묵적으로 간주되어 왔다. 주로 기록이라는 기술 대상 그 자체나 모든 행위가 끝난 뒤의 결과를 주요 기술 대상으로 지정했기 때문이다. 게다가 CIDOC CRM에서는 각 사건과 그 행위로 나타난 결과를 물리적인 것과 개념적인 것으로 다시 구분하는데, 기록관리 기술표준의 개별 기술요소와 요소 값에서 이 둘을 명확히 분리하기는 어렵다. 따라서 맵핑을 위한 분석과정에서 기존 기술표준을 모두 분해하게 된다. 이 과정에서 더욱 발전된다면, 박물관과 도서관 분야에서 FRBR^{oo}를 협력하여 개발했듯이, CIDOC CRM의 특징을 수용하면서도 기록관리 분야의 주요 개념을 반영한 연계 온톨로지로서 이어질 수 있을 것이다. 현재 ICA EGAD에서 개발 중인 RiC-O 온톨로지가 CRM과의 매개 역할도 할 수 있을지는 추이를 지켜볼 필요가 있다.

3.3.3 서술식 데이터의 재구조화

CIDOC-CRM과 맵핑 정보를 구축하면서 가장 어려웠던 점은 행위자나 날짜, 기록 및 출판물과 같은 개별 개체 정보가 서술식으로만 표현되어 있는 경우였다. 긴 텍스트로 이루어진 기술요소값은 인벤토리와 같은 검색도구에서 기록레코드를 브라우징하는 데에는 바람직하다. 그러나 온톨로지와 같이 작은 단위의 정보가 촘촘히 연결되어야 하는 구조에서는 제한이 생긴다. 전통적으로 기록물을 기술할 때는 참조코드

를 레코드의 배열용으로 활용하고 제목이나 일자는 기록의 내용을 가장 빠르게 식별하게 해주는 역할을 했었다. 이외 기술 정보는 기록의 식별이나 기록물의 대체물 역할도 했다. 그런데 앞으로는 배열 및 식별 기능과 더불어 ‘연계’ 기능이 더욱 강화될 것이고, 연계 대상은 유사한 기능을 수행하는 유형별 기록물 관리기관보다 넓은 문화자원 관리기관 전체가 될 것이다. 기존 ISAD(G)의 기술요소값을 재구조화할 때는 서술형으로 남겨둘 부분과 개체명으로 선정할 것의 기준을 정책적인 판단과 함께 정해야 한다. 기존 값에 있는 모든 서술형 정보를 간략화시킬 필요는 없다. 그러나 기록물이나 기록레코드, 행위자, 출판물 등과 같은 개체명이 기록 정보의 검색과 연계에 중요하다고 판단된다면 별도의 식별기호를 부여하여 식별기호를 지닌 다른 개체와 명확히 연계될 수 있어야 한다.

4. 결론

기록관리 분야 내에서도 상이한 기관에 소장된 기록과 그 기록에 대한 기술 정보를 연계하는 데에는 많은 노력이 필요하다. 그러나 이제는 기록관리 분야뿐 아니라 더 넓은 개념인 문화자원 관리기관이라는 우산 속에 박물관이나 미술관, 도서관과 함께 메타 정보를 공유하고, 이용자에게 통합 서비스를 제공해야 한다는 요구가 커지고 있다. 이에 본 연구에서는 문화자원 관리를 위한 도메인 온톨로지인 CIDOC CRM을 가장 널리 보급된 기록물 기술표준인 ISAD(G)와 시범적으로 맵핑하였다. ISAD(G)는 매우 친숙한 표준이지만 이것을 온톨로지 형식이나 관계가

명시된 그래프형으로 분석하는 과정은 낯설었다. ISAD(G)를 분석하는 과정에서 풍부한 서술 정보의 강점을 느끼기도 했지만, ISAD(G)에 내재된 기록물 기술 관행도 함께 알 수 있었다. ISAD(G)의 기술요소를 요소 값을 참조해 가며 CIDOC CRM의 클래스와 속성에 맵핑하는 것은 더욱 까다로운 작업이었다. 이 과정에서 선행연구들이 참고가 되었는데, 특히 EAD를 CIDOC CRM에 연계한 연구가 많은 도움이 되었다. 그리고 작성한 맵핑 정보에 따라 KARMA에서 실제로 기술요소와 온톨로지의 맵핑 작업을 수행해 보았다. KARMA는 맵핑과정을 시각적으로 확인할 수 있다는 장점이 있었지만, 연결 노드의 수가 증가할수록 시각화의 가독성이 떨어졌다. 그러나 KARMA는 온톨로지를 바탕으로 데이터를 통합하고, 통합된 데이터를 다시 RDF 형식으로 발행해 주는 편리한 기능을 제공해 주었다.

그런데 CIDOC CRM을 바탕으로 맵핑을 진행할수록 기록물 기술표준의 특성과 온톨로지의 장점을 조화시킬 수 있는 연계 온톨로지가 필요하다는 판단이 들었다. 상이한 메타데이터의 기술요소값을 1:1로 맵핑하는 방식보다는 온톨로지 기반으로 맵핑하는 방식이 정확하고 상세했고, 이를 위해선 반드시 매개가 되는 온톨로지가 필요했다. 그런데 CIDOC CRM 자체로는 기록관리에 필요한 상세한 어휘나 세부 속성이 부족한 면이 있었기 때문이다. 앞으로 FRBRoo와 같은 연계 온톨로지 구축을 위한 협력이 강화되고 기록물 기술 정보를 온톨로지 명세에 적합하도록 표현하기 위한 ISAD(G) 기술요소값의 재구조화가 진행된다면, 기록물 기술 정보의 활용도가 제고될 것이다. 그리고 그 과정에서 필요한 요소를 추가하고 다른 분야의 장점도 수용하는 분야 간 교류를 통한 혜택도 얻을 수 있을 것이다.

참 고 문 헌

- 국가기록원 (2018). 기록물 검색. 검색일자: 2018. 4. 30.
<http://www.archives.go.kr/next/search/viewDescClassMain.do>
- 박지영 (2008). 문화유산 자원 통합 활용을 위한 CRM 기반 FRBR 응용 온톨로지 적용에 관한 연구. 한국비블리아학회지, 19(2), 45-62.
- 박지영 (2014). RAMP를 활용한 EAC 기반 전거레코드의 연계 및 공유 관한 연구. 한국기록관리학회지, 14(2), 61-82.
- 박지영 (2016). 차세대 기록물 기술표준에 관한 연구. 한국기록관리학회지, 16(1), 223-245.
- 박지영 (2017a). ISAD(G)에서 RiC-CM으로의 전환에 관한 연구. 한국기록관리학회지, 17(1), 93-115.
- 박지영 (2017b). 디지털 시대의 기록물 기술과 접근. 한국기록관리학회지, 17(4), 87-107.
- 박지영 (2017c). FRBRoo 분석을 통한 FRBR 개념모형의 확장과 개선. 정보관리학회지, 34(4), 201-225.

- 박지영, 김태수 (2007). EAD를 이용한 기록정보의 기술 및 활용: 4월 혁명 연구반 컬렉션을 중심으로. 지식처리연구, 8(1/2), 17-57.
- 현문수 (2014). FRBRoo/CIDOC CRM 기반의 로컬리티 정보자원 구조화 연구. 한국비블리아학회지, 25(4), 265-290.
- 현문수 (2015). CIDOC-CRM을 이용한 지역기록의 공간 기반 구조화. 한국기록관리학회지, 15(2), 83-101.
- Bhaskar, M. (2016). Curation: The power of selection in a world of excess. Piatkus. (최윤영 역. 2016. 큐레이션: 과감히 덜어내는 힘. 예담아카이브).
- Bountouri, L. & Gergatsoulis, M. (2011). Mapping Encoded Archival Description to CIDOC CRM. Advances in Water Resources - ADV WATER RESOUR.
- Hong B., Hongzhe L., Jiehua Y., & Hongwei X. (2005). An Ontology-Based Semantic Integration for Digital Museums. in Advances in Web-Age Information Management.
- Hooland, S. V. & Verborgh, R. (2014). Linked Data for Libraries, Archives, and Museums. ALA. ICA (1999). General International Standard Archival Description. 2nd edition.
- ICA (2004). International Standard Archival Authority Records - Corporate Bodies, Persons, and Families. 2nd edition.
- ICA (2016). "ICA International Congress 2016."
- ICA EAGD. (2016). Record In Contexts: A Conceptual Model For Archival Description. Consultation Draft v.0.1.
- ICOM/CIDOC (2017). Definition of the CIDOC Conceptual Reference Model. ver. 6.2.3. Produced by the ICOM/CIDOC Documentation Standards Group, Continued by the CIDOC CRM Special Interest Group. (Editorial Status: In Progress since [12/10/2017]).
- Insight Centre for Data Analytics (2018). The Linked Open Data Cloud. Retrieved April 30, 2018, from <http://lod-cloud.net>
- ISO (2014). Information and documentation - A reference ontology for the interchange of cultural heritage information. ISO 21127:2014(E).
- Knoblock, C & Szekeley, P. (2013). Semantics for Big Data Integration and Analysis. 2013 AAAI Fall Symposium Series.
- Knoblock, C. & Szekeley, P. (2015). Exploiting Semantics for Big Data Integration. AI Magazine. 36. 25-38.
- Knoblock, C. (2017). Aligning and Integrating Data in Karma. ISWC 2017 Tutorial. Constructing Domain-specific Knowledge Graphs (KGC). Session I: Web Information Extraction, Wrappers and Information Integration using Karma. Vienna, Austria. 2017. 10. 21-25.

- <http://usc-isi-i2.github.io/slides/part-1-2.pdf>
- Library of Congress (2018). Eancoded Archival Description Official sites. Retrieved April 30, 2018, from <https://www.loc.gov/ead/index.html>
- RAMP Websites (2018). Remixing Archival Metadata Project. Retrieved April 30, 2018, from <https://tools.wmflabs.org/ramp/>
- Retrieved April 30, 2018, from <https://www.ica.org/en/international-congress-2016>
- Stasinopoulou T. et al. (2007a). WP5-Task 5.5 "Ontology Driven Interoperability", technical report of DELOS.
- Stasinopoulou T. et al. (2007b). Ontology-Based Metadata Integration in the Cultural Heritage Domain. In: Goh D.H.L., Cao T.H., Sølvsberg I.T., Rasmussen E. (eds) Asian Digital Libraries. Looking Back 10 Years and Forging New Frontiers. ICADL 2007. Lecture Notes in Computer Science, vol 4822. Springer, Berlin, Heidelberg.
- Szekely, P. (2015). Using Karma to map museum data to the CIDOC CRM ontology [video recording]. Retrieved April 30, 2018, from https://www.youtube.com/watch?v=h3_yiBhAJIc
- University of Southern California. (2016). Karma: A Data Integration Tool. Retrieved April 30, 2018, from <http://usc-isi-i2.github.io/karma/>

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- Bhaskar, M. (2016). Curation: The power of selection in a world of excess. Piatkus. (Translated by Choi, Youn young. 2016. Yeamoon Archive).
- Hyun, Moonsoo (2014). A Study on the Conceptualization of Information Resources for Localities Based on the FRBRoo/CIDOC CRM. Korean Biblia Society for Library and Information Science, 25(4), 265-290.
- Hyun, Moonsoo (2015). A Space-Based Approach to Organizing Local Records Using CIDOC CRM: A Case of the Port of Busan. Journal of Korean Society of Archives and Records Management, 15(2), 83-101.
- National Archives of Korea (2018). Archival Search menu on Website. Retrieved April 30, 2018, from <http://www.archives.go.kr/next/search/viewDescClassMain.do>
- Park, Zi-young & Kim, Tae-soo (2007). Implementation of EAD 2002 for 4.19 collection. Journal of Knowledge Processing and Management, 8(1/2), 17-57.
- Park, Ziyong (2008). A Study on the Application of a CRM-based FRBR Ontology for Cultural

- Heritage Information: Based on the FRBRoo(object-oriented FRBR). Korean Biblia Society for Library and Information Science, 19(2), 45-62.
- Park, Ziyong (2014). Linking and Sharing EAC Authority Records Using RAMP: Focusing on the Records of "Park, Kyung-ni". Journal of Korean Society of Archives and Records Management, 14(2), 61-82.
- Park, Ziyong (2016). Analyzing the Next-generation Archival Description Standard: "Record in Context" of ICA EGAD. Journal of Korean Society of Archives and Records Management, 16(1), 223-245.
- Park, Ziyong (2017a). Transition of Archival Description from ISAD(G) to Record in Context Conceptual Model. Journal of Korean Society of Archives and Records Management, 17(1), 93-115.
- Park, Ziyong (2017b). Archival Description and Access in Digital Age that Focuses on the Practices of The National Archives. Journal of Korean Society of Archives and Records Management, 17(4), 87-107.
- Park, Ziyong (2017c). Expanding and Improving FRBR Conceptual Model through FRBRoo. Journal of the Korean Society for Information Management, 34(4), 201-225.

[부록 1] ISAD(G) 제2관의 온톨로지 맵핑을 위한 클래스와 속성 분석(안)

ISAD(G) 2 관		클래스와 속성 분석
영역	기술요소	
식별영역	참조코드 (Reference code(s))	<ul style="list-style-type: none"> • 관련 클래스: 식별기호 • 속성: 식별기호를 갖는다/~의 식별기호이다 • 속성의 도메인(domain): 기술 단위 • 속성의 범위(range): 레이블(식별기호)
	제목 (Title)	<ul style="list-style-type: none"> • 관련 클래스: 제목 • 속성: 제목을 갖는다/~의 제목이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 레이블(제목)
	일자 (Dates)	<ul style="list-style-type: none"> • 관련 클래스: 날짜 • 속성: 날짜를 갖는다/~의 날짜이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 레이블(날짜)
	기술계층 (Level of description)	<ul style="list-style-type: none"> • 관련 클래스: 기술계층 • 속성: 기술계층에 속한다/~의 기술계층이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 레이블(기록물 기술계층)
	기술단위의 규모와 유형 (Extent and medium of the unit)	<ul style="list-style-type: none"> • 관련 클래스: 물리적 특성, 논리적 특성 • 속성: 특성을 갖는다/~의 특성이다 • 속성의 도메인: 기술단위 • 속성의 범위: 주기(물리적 특성, 논리적 특성 관련)
배경영역	생산자명 (Name of creator)	<ul style="list-style-type: none"> • 관련 클래스: 행위자(또는 개인 및 단체) • 속성: 생산하다/~에 의해 생산되다 • 속성의 도메인: 기술단위 • 속성의 범위: 레이블(행위자)
	행정연혁/개인이력 (Administrative/Biological history)	<ul style="list-style-type: none"> • 관련 클래스: 연혁, 이력 • 속성: 연혁/이력을 지닌다/~의 연혁/이력이다 • 속성의 도메인: 행위자 • 속성의 범위: 주기(연혁, 이력 관련)
	기록물 이력 (Archival history)	<ul style="list-style-type: none"> • 관련 클래스: 이력 • 속성: 이력을 지닌다/~의 이력이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 주기(이력 관련)
	수집/이관의 직접적 출처 (Immediate source of acquisition)	<ul style="list-style-type: none"> • 관련 클래스: 개인 또는 단체 (출처) • 속성: 수집(이관)하다/~로 부터 수집(이관)되다 • 속성유형: 데이터 타입 • 속성의 도메인: 개인 또는 단체 • 속성의 범위: 주기(기술단위의 수집/이관 관련)
내용과 구조영역	범위와 내용 (Scope and content)	<ul style="list-style-type: none"> • 관련 클래스: 주기 • 속성: 주기를 갖는다/~의 주기이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 주기
	평가, 폐기, 처리일정 정보 (Appraisal, destruction and scheduling)	<ul style="list-style-type: none"> • 관련 클래스: 평가, 폐기, 처리일정 • 속성: 관리 평가 정보를 갖는다/~의 평가 정보이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 주기(평가, 폐기, 처리일정 관련)

ISAD(G) 2 판		클래스와 속성 분석
영역	기술요소	
	추가수집 예상 기록물 (Accruals)	<ul style="list-style-type: none"> • 관련 클래스: 수집/이관 • 속성: 수집 정보를 갖는다/~의 수집 정보이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 주기(수집 일정 관련)
	정리체계(System of arrangement)	<ul style="list-style-type: none"> • 관련 클래스: 정리/분류 체계 • 속성: 정리/분류 체계를 갖는다/~의 정리/분류 체계이다. • 속성의 도메인: 기술 단위 • 속성의 범위: 주기(정리/분류 체계 관련)
접근과 이용환경영역	접근환경 (Conditions governing access)	<ul style="list-style-type: none"> • 관련 클래스: 이용 조건 • 속성: 이용 조건을 갖는다/~의 이용 조건이 있다 • 속성의 도메인: 기술단위 • 속성의 범위: 주기(이용 조건 관련)
	이용환경 (Conditions governing reproduction)	<ul style="list-style-type: none"> • 관련 클래스: 복제 조건 • 속성: 복제 조건을 갖는다/~의 복제 조건이 있다 • 속성의 도메인: 기술 단위 • 속성의 범위: 주기(복제 조건 관련)
	자료의 언어 (Language/scripts of material)	<ul style="list-style-type: none"> • 관련 클래스: 언어/문자 • 속성: 언어/문자를 사용하다/~의 언어/문자가 이용되다 • 속성의 도메인: 기술 단위 • 속성의 범위: 주기(언어/문자 관련)
	물리적 특성과 기술적 요구조건 (Physical characteristics and technical requirements)	<ul style="list-style-type: none"> • 관련 클래스: 물리적/기술적 특성 • 속성: 물리적/기술적 요건이 필요하다/~이용을 위한 물리적/기술적 요건이다 • 속성의 도메인: 기술 단위의 이용 • 속성의 범위: 주기(물리적/기술적 특성 관련)
	검색도구 (Finding aids)	<ul style="list-style-type: none"> • 관련 클래스: 검색도구 • 속성: 검색도구를 갖는다/~의 검색도구이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 주기(검색도구 관련)
관련자료영역	원본의 존재와 위치 (Existence and location of originals)	<ul style="list-style-type: none"> • 관련 클래스: 원본 • 속성: 관련자료가 있다/~의 관련자료이다 • 속성의 도메인: 기술 단위 (유형: 원본) • 속성의 범위: 주기(기술 단위 -유형: 사본 관련)
	사본의 존재와 위치 (Existence and location of copies)	<ul style="list-style-type: none"> • 관련 클래스: 사본 • 속성: 관련 자료가 있다/~의 관련 자료이다 • 속성의 도메인: 기술 단위 (유형: 사본) • 속성의 범위: 주기(기술 단위 -유형: 원본 관련)
	관련 기술단위 (Related units of description)	<ul style="list-style-type: none"> • 관련 클래스: 기술 단위 • 속성: 관련 자료가 있다/~의 관련 자료이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 주기(기술 단위 관련)
	출판물 설명 (Publication note)	<ul style="list-style-type: none"> • 관련 클래스: 출판물 • 속성: 관련 출판물이 있다/~의 관련 출판물이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 주기(출판물 관련)
추가설명영역	추가설명 (Note)	<ul style="list-style-type: none"> • 관련 클래스: 주기 • 속성: 주기가 있다/~의 주기이다 • 속성의 도메인: 기술 단위 • 속성의 범위: 주기

ISAD(G) 2 판		클래스와 속성 분석
영역	기술요소	
기술통제영역	기술담당자 주기 (Archivist's note)	<ul style="list-style-type: none"> • 관련 클래스: 주기 • 속성: 주기가 있다/~의 주기이다 • 속성의 도메인: 기록물 레코드(기술 정보) • 속성의 범위: 주기(기록물 기술 관련)
	규칙과 협약 (Rules or conventions)	<ul style="list-style-type: none"> • 관련 클래스: 기록물 기술규칙 • 속성: 관련 규정이 있다/~의 관련 규정이다 • 속성의 도메인: 기록물 레코드(기술 정보) • 속성의 범위: 주기(기술규칙 관련)
	기술일자 (Date(s) of descriptions)	<ul style="list-style-type: none"> • 관련 클래스: 날짜 • 속성: 날짜를 갖는다/~의 날짜이다(또는 제정되다, 개정되다) • 속성의 도메인: 기록물 레코드(기술 정보) • 속성의 범위: 주기(제정/개정 일자 관련)

[부록 2] ISAD(G) 제2판의 기술요소와 CIDOC CRM 연계 (안)

ISAD(G) 2 판		관련 속성
영역	기술요소	
식별영역	참조코드	<ul style="list-style-type: none"> • 주요 클래스: E42 Identifier • 관련 속성1: P48 has preferred identifier (is preferred identifier of) [도메인: E1 CRM Entity, 범위: E42 Identifier] • 관련 속성2: P37 assigned (was assigned by) [도메인: E15 Identifier Assignment, 범위: E42 Identifier]
	제목	<ul style="list-style-type: none"> • 주요 클래스: E35 Title • 관련속성: P102 has title (is title of) [도메인: E71 Man-Made Thing, 범위: E35 Title]
	일자	<ul style="list-style-type: none"> • 주요 클래스: E52 Time-span • 관련 속성1: P4 has time-span (is time-span of) [도메인: E2 Temporal Entity, 범위: E52 Time-Span] • 관련 속성2: P78 is identified by (identifies) [도메인: E52 Time-Span, 범위: E49 Time Appellation]
	기술계층	<ul style="list-style-type: none"> • 주요 클래스: E71 Man-Made Thing • 클래스 SN: This class comprises discrete, identifiable man-made items that are documented as single units. • 관련 속성1: P103 was intended for (was intention of) [도메인: E71 Man-Made Thing, 범위: E55 Type] • 관련 속성2: P19 was intended use of (was made for) [도메인: E7 Activity, 범위: E71 Man-Made Thing] • 관련 속성3: P102 - has title (is title of) [도메인: E71 Man-Made Thing, 범위: E35 Title]
	기술단위의 규모와 유형	<ul style="list-style-type: none"> • 주요 클래스: E26 Physical Feature • 관련 속성: P56 - bears feature (is found on) [도메인: E19 Physical Object, 범위: E26 Physical Feature]
배경영역	생산자명	<ul style="list-style-type: none"> • 주요 클래스: E39 Actor • 관련 속성1: P14 carried out by (performed) [도메인: E7 Activity, 범위: E39 Actor] - E7 하위에 E65 Creation 있음 • 관련 속성2: P131 is identified by (identifies) [도메인: E39 Actor, 범위: E82 Actor Appellation] • 관련 속성3: P75 possesses (is possessed by) [도메인: E39 Actor, 범위: E30 Right]
	행정연혁/개인이력 기록물 이력	<ul style="list-style-type: none"> • 주요 클래스: E62 String (E59 Primitive Value의 하위 클래스) • 관련 속성: P3 has note [도메인: E1 CRM Entity, 범위: E62 String] • 관련 이슈1: E7 Activity 하위 클래스에 E8 Acquisition, E9 Move, E10 Transfer of Custody, E11 Modification 등 이력 정보나 출처 정보를 나타낼 수 있는 클래스와 관련 속성들이 있으나, ISAD(G)에서는 서술형 텍스트로만 제공되어 있어 직접적으로 클래스와 속성 연계가 어려움 • 관련 이슈2: 서술형 텍스트의 구조화가 어려울 경우, E62 String 클래스의 하위 클래스를 추가하거나 P3 has note에 하위 속성을 추가하여 각 기술요소를 구분하여 모델링 하는 방안도 고려할 수 있음
내용과 구조영역	수집/이관의 직접적 출처	
	범위와 내용	
	평가, 폐기, 처리일정 정보	
	추가수집 예상 기록물	
접근과 이용환경영역	정리체계	
	접근환경	
	이용환경	
	자료의 언어	<ul style="list-style-type: none"> • 주요 클래스: E56 Language • 관련 속성: P72 has language (is language of) [도메인: E33 Linguistic Object, 범위: E56 Language]
접근과 이용환경영역	물리적 특성과 기술적 요구조건	<ul style="list-style-type: none"> • 주요 클래스: E62 String (E59 Primitive Value의 하위 클래스) • 관련 속성: P3 has note [도메인: E1 CRM Entity, 범위: E62 String]
	검색도구	<ul style="list-style-type: none"> • 주요 클래스: E62 String (E59 Primitive Value의 하위 클래스) • 관련 속성: P3 has note [도메인: E1 CRM Entity, 범위: E62 String]

ISAD(G) 2 관		관련 속성
영역	기술요소	
관련자료영역	원본의 존재와 위치	<ul style="list-style-type: none"> • 주요 클래스: E62 String (E59 Primitive Value의 하위 클래스) • 관련 속성: P3 has note [도메인: E1 CRM Entity, 범위: E62 String] • 관련 이슈: 관련 자료는 모두 E71 Man-Made Thing의 일종으로 모델링하여 데이터 타입이 아닌 오브젝트 타입으로 연결할 필요가 있음
	사본의 존재와 위치	
	관련 기술단위	
	출판물 설명	
추가설명영역	추가설명	<ul style="list-style-type: none"> • 주요 클래스: E62 String (E59 Primitive Value의 하위 클래스) • 관련 속성: P3 has note [도메인: E1 CRM Entity, 범위: E62 String]
기술통제영역	기술담당자	<ul style="list-style-type: none"> • 주요 클래스: E39 Actor • 관련 속성1: P14 carried out by (performed) [도메인: E7 Activity, 범위: E39 Actor] - E7 하위에 E87 Curation Activity 있음 • 관련 속성2: P131 is identified by (identifies) [도메인: E39 Actor, 범위: E82 Actor Appellation] • 관련 속성3: P75 possesses (is possessed by) [도메인: E39 Actor, 범위: E30 Right] • 관련 이슈: E39 클래스에 E7과 같은 이벤트 클래스를 연계하여 기술담당자를 표현함
		규칙과 협약
	기술일자	<ul style="list-style-type: none"> • 주요 클래스: E52 Time-span • 관련 속성1: P4 has time-span (is time-span of) [도메인: E2 Temporal Entity, 범위: E52 Time-Span] • 관련 속성2: P78 is identified by (identifies) [도메인: E52 Time-Span, 범위: E49 Time Appellation]