

기록물 생산기관명 접근점 제어 온톨로지 기술에 관한 연구

A Study on the Description of Archives Name by Controlled Access Point in Ontology

강현민 (Hyen Min Kang)*

초 록

기록물 생산기관에 부여된 행정표준기관코드(Standard Administration Code) 식별기호체계를 이용하여 동일 정체성과 동일 신원을 지닌 기록물 최종 생산기관명을 유일한 우선어 형식(preferred form)으로 정의하였다. 또한 동일 정체성과 동일 신원을 지닌 생산기관명에 대해 다양한 형식의 언어와 문자로 알려진 이형형식의 접근점 중에 대표 이름형식 접근점을 선정하고 온톨로지로 기술하였다. 이로써 과거부터 현재에 이르기까지 동일 정체성과 동일 신원의 생산기관이 생산한 모든 기록물에 대해 식별·접근할 수 있도록 하였다. 또한 '출처주의 원칙'과 '원질서 존중의 원칙'이라는 기본적인 기록물 관리 원칙을 온톨로지 환경에서도 준수함으로써, 기록물에 대한 관리방식 개선과 이용자들에게 기록물 접근의 편의성 및 확장된 검색결과를 제공할 수 있도록 하였다.

ABSTRACT

This study defined the name of records producing institution as the unique preferred form of access point which has same identification and same entity by using Standard Administration Code, and also described the name of records producing institution which has various name form as formal-name form of access point, which has same identification and same entity. This study make us be able to identify and access to all of the records that institution, has same identification and same entity, has produced. And the mechanic, I designed by ontology would make reinforce 'the principle of provenance' and 'respect for original order' and make user satisfy in useability of archives and expanded retrieval results.

키워드: 행정표준기관코드, 접근점, 무전거, 비제어형 접근점, 우선어 접근점, 대표 이름형식
access point, non-authority control, non-controlled access point,
preferred access point

* 행정안전부 국가기록원 사서주사(xianmin1970@gmail.com)

■ 논문접수일자: 2018년 8월 19일 ■ 최초심사일자: 2018년 9월 15일 ■ 게재확정일자: 2018년 9월 21일
■ 정보관리학회지, 35(3), 147-164, 2018. [http://dx.doi.org/10.3743/KOSIM.2018.35.3.147]

1. 서론

1841년 프랑스는 풍존중(respect des fonds)의 기본원칙을 공식화하고 출처주의(principle of provenance)와 원질서 존중(respect for original order)을 강조하는 기록물 정리의 원칙을 수립하였다. 현대기록 관리도 국가중앙부처 단위 등에서 생산한 기록물들은 이 원칙들이 존중되면서 정리·기술을 수행하고 있다.

생산기관 중심의 기록물 정리원칙은 생산기관이 생성, 폐지되고 다양한 형식의 이름을 갖고 복잡하게 변천하는 것과 동시에, 기록물생산의 폭발적 증가와 더불어 전자적 기록물 관리환경으로의 급속한 전환으로 기록관리 업무는 새로운 도전을 맞게 되었다. 동일한 정체성을 지닌 특정 생산기관명이 유기체적 생애주기 변천과정을 거치면서 다양한 언어와 문자를 갖는 형식으로 수많은 기록물을 생산하는 과정에서 출처주의의 집중(collocation) 기능과 원질서 존중의 무결성(integrity)을 유지하는 데 많은 어려움을 겪고 있다.

또한 국제표준기록전거레코드 ISAAR(CPF)(International Standard Archival Authority Record for Corporate Bodies, Persons and Families)와 국제표준기술규칙인 ISAD(G)에는 동일 생산기관을 기술하는 서술적 성격이 강한 상이한 요소정의가 중복 존재하며, 이들 간 상호연계성도 부족하여 생산기관의 정보를 획득하는데 많은 어려움이 있다. Bearman과 Lytle

(1985)은 하나의 서지레코드 안에 기록물 자체에 대한 정보와 출처에 대한 정보가 혼재되는 양상을 ‘출처주의 원칙 적용의 역기능’이라 표현한 바 있다(설문원, 2002, pp. 40-41. 재인용).

본 연구의 목적은 기존 관계형 데이터베이스 저장메커니즘과, HTML/XML 기반의 웹 환경과는 달리 RDF/OWL 기반의 시멘틱 웹 환경에서 술어논리에 기반 한 RDF 트리플 저장구조와 기술(description), DL¹⁾ 수준의 개체 표현력을 지닌 언어를 활용하여 특정 기관생산명의 이형 이름 간의 관계를 온톨로지 기술 함으로써, 동일 신원의 이형 이름 형식을 일관되게 제어하여 함으로써 출처주의의 원칙과 원질서 존중의 원칙을 보다 강화할 수 있는 방법을 제시하는 것이다.

또한 기존 전통적인 단체명의 전거형식, 기본표목 등 기존 전통적인 저록(著錄)²⁾생산방식을 대신하여 RDF/OWL 기반의 무전거 비제어형 접근점 즉, 모든 상이한 이름형식을 대등한 접근점으로 인정하여 각 기관이 자유롭게 자신의 기관명에 대해 다양한 언어와 문자로 된 우선어 접근점(preferred entry access point)으로 집중(collocation)되는 새로운 개념의 연결장치(linking device)를 제시하는 것이다. 이 연결장치를 시멘틱 웹 온톨로지 언어로 기술하여 시멘틱 웹 상에서 타 영역(domain)에서도 동일한 이름공간(name space) 속에서 광역적이며 전역적으로 참조될 수 있고, 생산기관명(단체명)의 다양한 유기체적 변화관계

1) OWL 언어에는 온톨로지 표현력에 따라 OWL Lite < OWL DL < OWL Full 3종이 있다.

2) 저록(著錄)은 일반적으로 특정 기술 대상에 대해 특정한 마크업 언어로 기술(description)한 메타데이터 영역에서의 산출물로서 카드목록시대에서 main entry card, MARC 형식의 목록시대에서는 서지레코드 1건 등을 가리킨다. 일반적으로 ‘레코드’라는 용어와 혼용해서 사용된다.

를 통일성, 일관성, 재활용성, 상호운용성, 공유성을 유지하며 기술하는 방안을 제시하는 것이다.

연구방법은 먼저 생산기관 전거레코드와 행정표준기관코드체계를 소개하고 전거의 개념의 문제점과 한계를 지적하였다. 이어 무전거 비제어형 접근점과 이형 접근점을 온톨로지 환경에서 이름공간의 IRI³⁾와 결합하여 이에 대한 이론적 목적과 의의를 설명하였다. 먼저 전거 및 표목의 개념을 배제하고 현재 활동 중인 생산기관의 이름을 무전거 비제어형 우선어로 선정하여 그 행정표준기관코드 식별기호체계와 결합하여 RDF/OWL을 사용하여 다양한 공리⁴⁾를 정의한 온톨로지를 기술하였다. 그리고 동일 정체성 및 동일 신원을 가진 특정 생산기관의 이름이 유기체적 생애주기과정을 거치면서 생성된 언어, 문자, 형식상에서 다양한 이형의 이름 형식 가운데 하나를 대표이형 이름 형식으로 선정하여 상이한 이름형식들이 온톨로지 기반에서는 시멘틱 웹 리즈너가 동일한 이름형식 또는 이름 인스턴스로 인식할 수 있도록 기술하였다. 이때 특정 생산기관으로 '행정안전부'를 선정하였으며, 온톨로지 저작도구인 프로티지(protégé)를 사용하여 클래스, 객체속성, 인스턴스 정의와 공리 및 특질 등을 정의하여 온톨로지를 기술하였다.

이번 연구의 의의는 무전거 비제어형 접근점 클래스에서 우선어 형식의 선정기준을 현재 활동 중인 생산기관명으로 한 점과 과거 다양한

형식의 이름들을 RDF/OWL 객체속성으로 연결하여 모든 형식의 이름들이 동일한 정체성을 갖고 있음을 온톨로지 기술한 것이다. 또한 동일 정체성과 동일 신원을 갖고 있는 특정 생산기관명의 다양한 유기체적 생애주기에 기반한 이형 형식의 이름 접근점과 표준행정기관코드와 결합하여 다양한 이형이름형식에서 대표 이름형식의 접근점을 선정함으로써 동일 신원의 대표 이름형식을 제시하고 다른 이형 형식의 접근점과도 동일 신원임을 광역적으로 공유할 수 있도록 하였다. 이러한 무전거 비제어형 우선어 형식의 접근점과 동일 신원의 대표 이름형식의 접근점을 행정표준기관코드와 결합하여 RDF/OWL 기반 시멘틱 웹 온톨로지 환경에서 유일성, 공유성, 일관성, 광역적 식별성을 지닌 IRI로 정의하였다. 동일 신원의 기록물 생산기관에 대해 일관된 출처주의 원칙과 원질서 존중의 원칙을 지속적으로 보장하고 이 뼈대니즘을 더욱 더 공고히 했다는 점이다. 본 연구에서 기록물 생산기관명은 특정기간동안 존재했던 특정 기관의 다양한 이름 표현형식을 말한다. 즉 '특정기간의 생몰년도'를 가진 조직은 동일신원의 동일정체성을 갖는 특정 기관이라고 지칭할 수 있다. 여기에 통시적으로 특정기관이 다양한 성격과 기능이 변경되는 것에 대한 논점은 본 연구범위에서 제외시켰다.

3) 시멘틱 웹 상에서 특정 개체(개념, 대상, 장소, 인물, 시간 등)의 고유한 이름(International Resource Identifier) 문자열로서 유일성, 광역성, 상호호환성, 공유성을 갖는 문자열을 지칭한다.

4) 공리(axiom)란 사전적 의미로 '자명한 이치'를 말한다. RDF의 기술논리(predicate)로서 주어, 목적어, 술어로 정의되어 특정한 자명한 사실을 RDF 트리플로 표현한 집합을 말한다.

2. 생산기관 접근점 현황

2.1 생산기관명 전거레코드

기록물 생산기관명의 유기체적 생애주기 변천과정을 추적하기 위해서는 유일성이 확보된 코드명과 코드 값이 해당 기관에게 부여되어 있어야 한다. 이는 특정생산기관명의 변천 과정을 정보시스템으로 식별·관리하기 위한 기본 전제조건이기도 하다. 특히 특정기관명에 1:1 대응차수로 할당된 행정표준기관코드 식별기호체

계는 특정기관명의 식별성, 유일성, 광역성, 범용성, 공유성이 확보되어야만 그 정체성과 동일신원을 표현하는 접근점 간 연결을 위한 중요한 핵심 연결장치가 될 수 있다. 또한 생산기관명의 다양한 변화와 표현 형식에 대한 식별과 추적을 위해서는 유기체적 생애주기 변천과정을 관리하기 위한 최소한의 요건으로서 광역적으로 유일하게 식별하는 기호체계(notation) 구축 또한 기본 전제조건이 된다.

〈표 1〉은 국가기록원 전거저록(authority entry, 전거레코드) 중 '행정자치부'에 대한 전

〈표 1〉 국가기록원 전거저록 작성 사례

영역	구분	내용
식별	기관유형	중앙행정기관
	기관코드	1740000
	기관명	행정자치부(OG0121226)
	기타 기관명	Ministry of Government Administration and Home Affairs, 行政自治部
기술	존립기간	2014.11.19 ~ 현재 (존재)
	설치근거	정부조직법 [시행 2014.11.19.] [법률 제12844호, 2014.11.19., 일부개정]
	기관연혁	행정자치부는 1948년 7월에 신설된 내무부와 총무처로 시작되었다. 총무처는 1995년 2월 국무원 사무국으로 되었다가, 1961년 10월 내각사무처로 개편되었다. 1963년 12월 내각사무처가 폐지되고 총무처로 부활한 후, 1998년 2월 총무처와 내무부가 통합되어 행정자치부가 설치되었다. 이후 2008년 2월 행정안전부로 개편되었다가 2013년 3월 안전행정부로 명칭이 바뀌었다. 2014년 11월 '정부조직법 [시행 2014.11.19.] [법률 제12844호, 2014.11.19., 일부개정]'에 따라 안전행정부에서 행정자치부로 명칭이 바뀌었다.
	기관장	제1대 장관 정중섭 (2014.11.19 ~)
	주기사항	
통제	참고정보원	정부조직법 [시행 2014.11.19.] [법률 제12844호, 2014.11.19., 일부개정] 정부조직법 [시행 1948.7.17.] [법률 제1호, 1948.7.17., 제정] 행정자치부 (www.mogaha.go.kr) (2015년 6월 현재) 정부조직관리정보시스템 (org.mogaha.go.kr) (2015년 6월 현재) 행정자치부정부기구도 (org.mogaha.go.kr) (2015년 6월 현재) 국가법령정보센터 (www.law.go.kr) (2015년 6월 현재) 행정표준코드관리시스템 (code.mogaha.go.kr) (2015년 6월 현재)
	최초기술일	2015.11.05
	최종기술일	2015.11.23
	이전기관	안전행정부
	하위기관	

(출처: 국가기록원 2007)

거저록이다. 이 저록은 ISAAR(CPF)를 인용표준으로 삼아 국가기록원이 원내표준으로 제정한 것이다. 물론 ISAD(G)를 인용표준으로 삼아 국가기록원이 원내표준으로 제정한 '연구기록물 기술규칙'에 의해 작성된 기술저록(descriptive entry) 내에도 생산기관명 '행정자치부'에 대한 기술요소가 포함되어 있다. 그런데 이용자 관점에서는 특정 생산기관명을 식별하고 확인하기 위해서는 전거레코드와 기록물 기술서 내의 생산기관 정보 내용을 읽고, 이용자 스스로 다양한 참고정보원과 연구를 통해 동일기관에 대하여 확인하고 이해하고 식별하여야 한다. 즉, 해당 기관이 다른 이름 형식을 가진 동일기관인지 아니면 동일 이름 형식을 가진 다른 기관인지 이용자가 직접 확인하고 판단하여야 한다.

또한 <표 1>의 전거저록은 식별영역, 기관연혁, 참고정보원 등의 요소를 통해 특정 생산기관의 변천연혁을 서술형식, 열거형식 또는 나열형식으로 기술하고 있다. 이러한 전거저록의 각 영역과 기술요소들의 전거정보 기술형식은 구조적이고 체계적으로 타 시스템과 효율적으로 연동이 가능한 정보의 구조가 아니다. 이는 기술저록과 전거저록 간의 연계성을 저하시키고 특정기관의 전거저록 간 생산기관명의 변천연혁을 아날로그 방식으로 추적할 수밖에 없는 장애요인으로 작용하고 있다.

<표 1>에서 식별영역에 동일한 정체성을 갖고 있는 특정기관인 행정자치부에 대한 고유식별기호체계가 행정표준기관코드(1740000)와 구기록물관리기관코드(OG0121226)가 복수로 할당되어 있어 기호체계의 유일성과 대표성이 없다. 또한 특정 시대의 특정기관명인 '행정자치부'에

대한 단편적인 기관정보만 제공해 줄 뿐, 이 기관이 과거 어떤 기관이었으며 어떠한 과정을 거쳐 오늘날에 이르렀는지는 오직 제시된 기관 연혁을 통해 이해하고 판단할 수밖에 없다. 즉, 전거저록 간 또는 기술저록 간 대표 접근점과 연결 접근점이 정보시스템을 통해 생산기관명 변천맥락 접근점이 연결되어 있질 않아 이용자는 동일 정체성을 가진 동일 신원의 기관임에도 불구하고 그 변천이력을 식별하거나 추적하기가 어려운 환경에 처해 있다. 이는 이용자들로 하여금 특정 생산기관이 생산한 기록물 전체에 대해 기록의 생산 맥락의 이해와 식별·접근 등의 활동을 저해하는 근본적인 원인으로 작용하고 있으며, 또한 이용자들에게 특정 생산기관명의 변천연혁에 대한 통시적 연결성을 저하시켜 출처주의 원칙과 원질서 존중의 원칙이라는 기본적인 검색도구로서의 기능을 저해하고 있다.

2.2 행정표준기관코드

현재 행정안전부에서 운영하고 있는 행정표준코드시스템은 행정기관 간 행정정보의 원활한 공동이용을 도모하기 위해 각급 기관의 중요 행정업무에 필요한 행정코드를 표준화하여 정해진 절차에 따라 제정, 고시한 행정코드관리시스템을 말한다. 이때 행정코드란 각급기관의 행정업무에서 분류할 수 있는 업무내용(기관, 직위, 일반, 위임 등)을 일정 기준에 의하여 간소하게 처리할 수 있도록 단순화하고, 유일성을 보장하기 위해 코드 값과 코드 값의 의미를 1:1로 정의하여 각종 정보시스템 등에 공통적으로 활용하는 코드를 말한다(행정자치부, 2008). 2018년 7월 현재 276종의 행정표준코드가 제정되어

있으며, 코드값 '156'은 코드명 '수수료 구분', 코드값 '276'은 코드명 '자연재해저감시설'로, 특히 코드값 '1'은 생산기관명 식별기호체계라 할 수 있는 코드명 '기관'으로 정의되어 있다. 이와 같이 행정개념의 통일성과 표준화를 위해 다양한 행정영역으로 세분하여 코드명과 코드값이 정의되어 있다.

한편 행정표준기관코드는 행정표준코드에 포함된 '기관' 관련 코드로서 7자리의 숫자기호체계를 사용하여 식별하고 있다. 이 행정표준기관코드는 기록물생산기관으로서 '생산기관'을 유일하게 식별하는 중요 식별기호체계이다. <표 2>는 행정표준코드 중 행정표준기관코드 정의 사례를 정리한 것이다. 코드 값 '1'과 코드명 '기관'의 행정표준기관코드 대구분은 코드 값 '01'인 코드명 '국가행정기관'으로, 중구분은 코드

값 '01'로서 코드명 '중앙행정기관 및 이에 준하는 기관'으로, 소구분은 코드 값 '01' 코드명 '부처', 코드값 '02' 코드명 '처', 코드 값 '03'은 코드명 '청' 등과 같은 체제를 이용한다.

<표 3>은 2018년7월 현재 조직 130,668개의 조직 및 세부조직에 부여된 34,919개의 기관 식별기호가 부여된 현황을 보여 주고 있다.

행정표준기관코드는 기관의 변천을 존재, 폐지 두 가지 유형으로 나누고 있다. 존폐년도를 기준으로 새롭게 생성된 기관에는 새로운 행정표준기관코드가 부여되고, 폐지된 기관은 존폐여부의 상태 값을 '폐지'로 변경하고, 폐지기관의 기록물은 생산기관으로 기록물관리시스템을 통해 '가상이관'이라는 절차를 거쳐 새로운 행정표준코드 식별기호 아래 모든 기록물을 이동(이관)시킨다. 이 과정에서 과거에 생산한 전

<표 2> 행정표준코드 중 행정표준기관코드 정의 사례

구분	코드값	코드명	단계	코드값	코드명
행정 표준 코드	1	기관 (행정표준 기관코드)	대구분	01	국가행정기관
			중구분	01	중앙행정기관 및 이에 준하는 기관
			소구분	01	부처
				02	처
				03	청

<표 3> 2018년 행정표준기관코드 현황

기관유형	구축내용	
	기관수	조직수
계	1,416	130,668
중앙행정기관	249	63,467
위원회	88	1,753
지방자치단체	46	36,760
교육청	40	8,602
대학	351	626
공사·공단	642	19,460

(출처: 국가기록원 2018)

자·비전자기록물들은 최종적으로 생성·존재하고 있는 기관아래 모든 기록물이 채수록 된다.

그 동안 기록관리 분야에서는 기타 구기록물관리기관 코드, 국가기록원 자체관리기관 코드, 행정표준기관 코드, 국가기록원 전거레코드지침의 단체기호 코드 등 4종의 체계를 가지고 생산기관을 식별관리하고 있었으나, 이 다원적 코드 체계가 2007년 행정표준기관코드로 통합됨으로써 기관과 코드 간 1:1 매칭에 근거하는 식별기호체계를 운영하기 시작하였다.

그러나 행정표준기관코드는 그 식별기호 부여방식이 불규칙적이고 임의적이어서 조직 변천과정에서 발생하는 다양한 변화와 이력들이 기관코드 식별기호와 시스템에 반영되지 못해 기록물 생산기관을 식별하고 변천과정을 추적하는데 많은 어려움이 있다. 예를 들어 <표 4>에서 보듯이 ‘행정안전부’라는 동일 정체성·동일 신원의 기관명임에도 불구하고 행정안전부(1741000), 행정자치부(1740000), 안전행정부(1312000), 행정안전부(1311000), 행정자치부(1310000), 내무부(9902310), 총무처(9900497) 등 서로 다른 식별기호와, 생몰년도 및 존폐여부 정보만 확인할 수 있다.

결국 이용자들은 행정안전부(1741000), 행정

자치부(1740000), 안전행정부(1312000), 행정안전부(1311000), 행정자치부(1310000), 내무부(9902310), 총무처(9900497)에서 생산한 기록물들을 파편적으로 각각의 생산기관명을 식별·추적해서 기록물을 검색·접근할 수밖에 없다. 이용자들은 동일 신원 기관명인 “행정안전부”에 대해 행정안전부(1741000), 행정자치부(1740000), 안전행정부(1312000), 행정안전부(1311000), 행정자치부(1310000), 내무부(9902310), 총무처(9900497)를 각 각 분리된 기관으로 인식할 수밖에 없고, 각 기관 간 상관관계(전신, 후신 등)나 어떤 기관명이 먼저 또는 나중에 설치되었는지 시대적 배열 관계도 알 수 없다. 오로지, 해당 생산기관이 “존재”, “폐지”되었다는 존폐정보만 확인 할 수 있다.

<표 5>를 보면 동일 신원인 ‘국가기록원’의 행정표준기관코드 식별기호체계도 동일한 문제점이 있다. 시간적으로 생성일자과 폐지일자 값이 존재하고 현재 최종적으로 존재하는 ‘국가기록원’의 마지막 또는 최종 코드값은 ‘1741050’로 ‘행정안전부 국가기록원’이라는 것을 알 수 있다. 문제는 <표 4>에서 각 행정표준기관코드 간의 시대적 연결(linking)이 없고 동일 신원명인 ‘국가기록원’에서 ‘행정안전부 국가기록원’과

<표 4> 동일 신원의 ‘행정안전부’에 대한 다양한 형식의 생산기관명

동일신원기관명	연번	행정표준기관코드	최상위기관명	생성일자	폐지일자	존폐여부
행정안전부	1	1741000	행정안전부	20170726		존재
	2	1740000	행정자치부	20141119	20170726	폐지
	3	1312000	안전행정부	20130323	20141119	폐지
	4	1311000	행정안전부	20080229	20130323	폐지
	5	1310000	행정자치부	19980228	20080229	폐지
	6	9902310	내무부	19481104	19980228	폐지
	7	9900497	총무처	19481104	19980228	폐지

〈표 5〉 국가기록원의 생산기관명 변경 연혁

동일 신원 기관명	연번	행정표준 기관코드	행정표준기관명	생성 일자	폐지 일자	존폐 여부
국가기록원	1	1741050	행정안전부 국가기록원	20170726		존재
	2	1740083	행정자치부 국가기록원	20141119	20170726	폐지
	3	1312125	안전행정부 국가기록원	20130323	20141119	폐지
	4	1311153	행정안전부 국가기록원	20080229	20130323	폐지
	5	1310377	행정자치부 국가기록원	20040524	20080229	폐지
	6	1310091	행정자치부 정부기록보존소	19980228	20040524	폐지
	7	9900604	총무처 기록보존소	null	19980228	폐지

‘행정자치부 국가기록원’과 ‘행정자치부 정부기록보존소’ 등 다양한 형식의 동일 신원명 역시 시대적 연결 관리(linking management)가 이뤄지지 않고 있다.

코드값과 코드명의 체계적 관리 부재의 문제는 더 나아가 이용자를 각 행정표준기관코드나 다양한 동일 신원명에 대한 각각 비맥락적 검색 과정을 요구하거나, 생산기관에 대한 이용자의 낮은 사전 인지정도를 통해 파편적으로 기록물에 접근할 수밖에 없는 환경적 제약을 초래하고 있다. 동일 신원명의 다양한 이름형식을 사전 인지와 각 생산기관 식별기호체계 인식을 기반으로 하는 이용자 검색은 동일 신원명인 ‘행정안전부 국가기록원’에 대한 단면적인 기록물 일부에만 접근 할 수밖에 없다.

이처럼 행정기관코드 식별기호체계 간의 전후 관계의 연결이 끊겨 있고 동일 신원명의 다양한 이름형식에 대한 이용자 제시 과정이 미흡한 상태에서는 이용자들로 하여금 동일 신원명의 특정 생산기관의 모든 기록물을 통시적으로 제공 받을 수 없으며, 이용자들에게 그 기관 기록물 생산의 맥락정보와 출처정보를 제대로 활용하지 못하는 정보이용환경을 초래하고 있다.

3. 무전거 비제어형 우선어 접근점 온톨로지 기술

3.1 무전거 비제어형 우선어 형식 접근점

2003년 한국목록규칙(KCR) 제4판에서 표목(heading)의 개념과 용어를 없애고 저록 간 연결기능을 강화 하기 위해 접근점(access point)이라는 용어로 대체되었다. 또한 전자적 목록환경의 도래로 전거제어의 개념이 약화되고, KCR 4판에서도 기본표목 개념의 배제로 전거(典據, authority)라는 용어도 삭제되고 ‘무전거’라는 개념이 도입되었다. 전통적인 카드목록시대에서 기계가독형 목록과 XML 기반의 목록생산 방식이 확산되면서 전거표목, 전거제어 및 표목의 개념은 무의미해지고 ‘무전거 비제어형 접근점’이라는 새로운 개념적 틀 속에서 저록을 작성하게 되었다.

전통적인 카드목록 시대의 도서관 영역에서 전거 표목은 동일 신원에 대한 다양한 형식의 문자열 중 접근의 대표성을 갖는 형식을 전거표목을 선정하여, 모든 형식의 표목들을 보라참조, 도보라 참조라는 연결장치를 통해 전거표목으

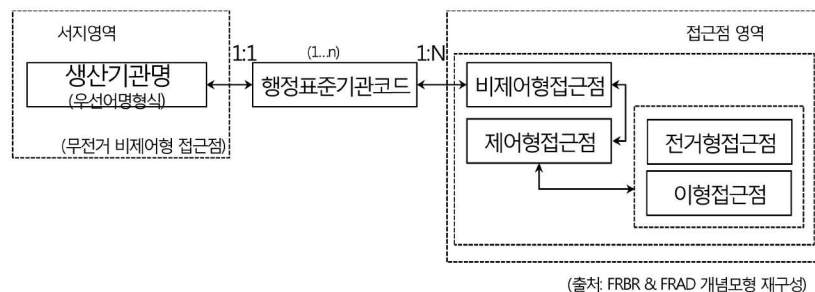
로 집중하도록 하는 것이었다. 그런데 정보시스템 환경에서는 표목의 개념을 접근점의 개념으로 대체·확장함으로써 전거의 개념과 표목의 개념이라는 대표성과 경직된 용어형식 목록규칙이 개정되고 다양한 형식의 용어들이 접근점으로 사용할 수 있도록 개정되었다. 이후 모든 용어들이 목록작성자에 의해 제어되었거나 제어되지 않았어도 특정 문헌 또는 기록물에 대해 모두 대등한 접근점을 허용하기 시작하였다.

〈그림 1〉의 접근점 영역에서는 제어형 접근점과 비제어형 접근점이 존재하고 제어형 접근점에는 전거형 접근점과 이형 접근점이 있다. 이 접근점 영역에 있는 모든 접근점들을 서지영역에서 접근 가능한 대등한 접근점 지위를 부여하면, 언어와 문자, 사회통념상의 습관 등 상대적으로 제어되지 않은 접근점도, 행정표준기관코드를 부여함으로써 서지영역에서 무전거 비제어형 접근점인 우선어명(preferred name) 형식으로 특정 생산기관의 이름으로 사용될 수 있다. 즉, 이 우선어명 형식은 단지 접근점으로만 활용되기 위한 형식이지 그 대표성, 규칙성과 규범성은 배제된 개념의 접근점이다.

3.2 동일 신원 대표 이름형식 접근점

동일 신원명을 표현하는 대표 이름형식은 이용자의 언어습관과 생산기관의 정한 정책, 일반적으로 자주 사용되는 이름형식으로 정해지기도 한다. 즉, 대표 이름형식은 특정기관의 다양한 이름형식에서 사용빈도, 사회적 통념, 대내외적인 이름 인지도 등을 고려하여 정한다. 따라서 3.1절 〈그림 1〉에서 이형 접근점, 전거형 접근점, 비제어형 접근점, 제어형 접근점 모두가 대표 이름 형식의 후보 접근점이 될 수 있다.

이처럼 접근점에 대한 제어기능이 약화된 이유는 전통적 카드목록시대와 기계가독형 목록형식 시대를 거치면서 관계형 데이터베이스 및 XML 기반 정보통신 환경이나 시멘틱 웹 환경의 도래와 관련있다. 각 각의 접근점은 모두 동일한 가치를 갖고 있는 일반적인 키워드 중 하나일뿐이고 수많은 동일 신원명의 이형 형식 간의 우선순위, 계층적 관계, 전후 관계, 종속관계 등의 복잡한 관계유형이 사라졌다. 이에 이용자 중심의 접근점 형식을 대표 형식으로 자유롭게 부여함으로써 이용자는 접근점 간 구조적 관계를 인지할 필요가 없어지고 정보이용의 편의성이 높아졌다. 다만, 대표 이름 형식



〈그림 1〉 무전거 비제어형 우선어명 접근점과 행정표준기관코드 간 연결모형

과 다양한 이형 형식의 이름 간에는 모두 동일 신원명임을 나타내기 위해 접근점 간 상호 연결이 사전에 정의되어야만 한다.

4. 생산기관명 접근점 제어 온톨로지 기술

4.1 행정표준기관코드와 접근점 간 연결

시멘틱 웹 온톨로지 환경에서는 특정 개체, 개념, 대상, 인물, 사건, 장소, 관계 등은 반드시 유일하고 고유한 문자열로 표현되어야 한다. 이를 위해 일반적으로 IRI 형식으로 이뤄진 문자열과 식별기호체계를 결합하여 시멘틱 웹 상에서 활용할 수 있는 고유한 문자열을 생성한다. 본 고에서는 RDF/OWL 온톨로지 스키마 내에 이름공간 'http://kanghm.corporatebody.org#'을 부여하고 생산기관명의 고유 식별기호체계이자 접근점인 '행정표준기관코드(SAC)'를 결합시켜 특정 생산기관을 지시하는 유일하고 고유한 문자열을 생성한다. 관계형 데이터베이스 정보시스템환경에서는 무전거 비제어형 접근점은 해당 시스템 내에서만 유일성과 고유성을 유지하고 있지만 온톨로지 환경에서 접근점을 표현하면 이 접근점은 어떠한 플랫폼이나 시멘틱 웹 환경에서 광역적이고, 전역적이며, 공유성과 상호호환성을 지닌 범용적 접근점으로 활용될 수 있다.

시멘틱 웹 환경에서 IRI 또는 QNAME 접두사와 행정표준기관코드를 결합시켜 동일한 이름공간 내에서 생산기관명의 유일성을 확보할 수 있다. <그림 2>에서 보는 바와 같이 이름공간의 문자열을 'http://kanghm.corporatebody.org#'로 정의하였으나 길고 복잡한 이름공간을 간략하게 표현하기 위해 접두사(QNAME)로써 지칭하는 'cb:'로 정의하고, 그 뒤에 행정표준기관코드를 결합시킨다. 즉 이름공간 http://kanghm.corporatebody.org#에 행정표준기관코드(9902310)를 결합시켜 특정단체명(내무부)을 지칭하는 유일한 IRI 문자열 식별기호체계를 생성한다.

<그림 2>는 이름공간 http://kanghm.corporatebody.org#를 QNAME 접두사인 'cb:'와 행정표준기관코드를 결합한 IRI 생성 사례이다. 생성된 '내무부' IRI 식별기호체계는 유일성, 범용성, 광역성을 지닌 문자열로서 항구적으로 특정기관 즉 '내무부'를 식별하는 문자열 기호체계로 사용된다. 예를 들어 행정표준기관코드 클래스는 우선어형 접근점 클래스의 인스턴스 갖는다는 의미의 객체속성을 간략하게 'cb:hasSAC'의 형식으로 나타낼 수 있다.

이렇게 생성된 특정기관의 IRI식별기호체계는 기존 행정표준기관코드가 갖고 있던 시스템 종속성과 식별체계의 지역성, 비공유성, 비재활용성이라는 제약에서 벗어나 시멘틱 웹 공간에서 범용적으로 자유롭게 공유되고 재사용될 수 있는 식별기호체계로 전환된다.

행정표준기관 이름공간		행정표준기관코드		QNAME 접두어 IRI
http://kanghm.corporatebody#	+	9902310	=	cb:9902310

<그림 2> 행정표준기관 이름공간과 행정표준기관코드를 활용한 IRI 생성

4.2 무전거 비제어형 우선어 형식 공리 정의

시멘틱 웹 상에서 모든 인스턴스들은 유일한 식별기호 문자열로 나타내어져야 한다. 이는 다양한 영역의 온톨로지 내의 인스턴스들과 의미나 개념의 충돌을 방지하기 위해서이다. <그림 3>은 서지저록과 접근점저록 상호 간의 관계를 표현한 것으로, 동일 신원명인 '행정자치부'를 인스턴스로 갖는 서지저록의 무전거 비제어형 우선어 형식 클래스(정의역)와 행정표준기관코드 클래스(치역)의 인스턴스를 객체속성(서술어)을 사용하여 연결한 온톨로지 도식이다. 행정표준코드 클래스 내에는 동일 정체성과 동일 신원을 가지며 다양한 이름 형식을 갖는 여러 인스턴스들을 치역으로 정의한 것이고, 행정표준코드 클래스의 인스턴스들 중 현재 존재하고 활동 중인 '행정안전부(1741000)'를 무전거 우선어형 접근점 클래스 정의역의 우선어 접근점 인스턴스로 정의한 것이다.

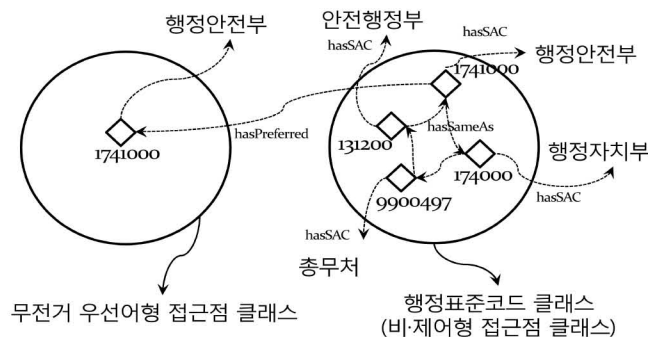
행정표준기관코드 클래스 내에서 각 식별기호들은 동일한 정체성과 신원임을 정의하기 위해 'hasSameAs'라는 객체속성을 정의하였으며

이어 모든 인스턴스들은 SPO 술어논리에 의한 RDF트리플로 저장된다. 이는 행정안전부, 안전행정부, 총무처, 행정자치부 등이 모두 대등한 우선어형 접근점임을 의미하며 이 중 현재 활동 중인 인스턴스 '행정안전부(1741000)'를 객체속성 'hasPreferred'를 사용하여 우선어형 접근점으로 기술하였다. 이러한 기술을 통해 이용자들은 모든 인스턴스들 중 그 어느 하나의 인스턴스 명만 알아도 동일 정체성·동일 신원의 '행정자치부'에 대한 모든 과거의 명과 현재의 명에 포함된 기록물을 탐색할 수 있게 된다.

<표 6>은 SPO 술어논리를 통한 RDF/OWL 기술 내용을 요약정리 한 것이다. 정의역은 주어 클래스이며 치역은 목적어 클래스이고 이들 간의 관계를 정의한 도구는 서술어 영역인 객체속성이다.

4.2.1 owl:minQualifiedCardinality

우선어형 접근점 클래스와 행정표준기관코드 클래스 중 우선어형 접근점 인스턴스는 행정표준기관코드 클래스 내 여러 인스턴스들은 1:n 대응차수를 갖는다. 이때 모든 인스턴스들은 반드시 하나의 행정표준기관코드(hasSAC)



<그림 3> 무전거 비제어형 우선어 형식 접근점 온톨로지 개념도

〈표 6〉 동일 정체성 · 동일 신원의 인스턴스 간 우선어 형식 정의

주어(S, 우선어 형식)	서술어(P, 객체속성)	목적어(O, 동일신원)
행정안전부 (1741000)	hasSAC hasSameAs	행정안전부(1741000)
		안전행정부(1312000)
		행정자치부(1740000)
		총무처(9900497)

를 가져야 한다. 왜냐하면 우선어형 접근점 클래스의 인스턴스는 행정표준기관코드 클래스 인스턴스 중 통시적으로 여러 생산기관명을 가지고 있고 이중 현재 존재하여 활동 중인 특정한 이름형식 하나를 우선어 형식의 인스턴스로 반드시 하나를 가져야하기 때문이다.

이 우선어 형식은 우선어형 접근점 클래스에서 유일하게 식별하는 기본형식 코드로서 행정표준기관코드 클래스의 인스턴스와 관계를 맺어간다. 따라서 우선어형 접근점 클래스→행정표준기관코드 클래스 관계에서 출현회수 공리를 인스턴스 특질인 owl:minQualifiedCardinality를 사용하여 '1'로 지정하였다. 〈그림 4〉는 인스턴스의 출현회수가 반드시 하나이어야 한다는 것을 RDF/OWL 구문으로 표현한 것이다.

4.2.2 클래스 간 서로 소(疎) 공리 정의 우선어형 접근점 클래스(cb:PreferredClass)

의 인스턴스와 행정표준기관코드 클래스(cb:SAC)의 인스턴스들은 상호 배타적이어야 한다. 즉, 특정 클래스의 인스턴스는 타 클래스의 인스턴스가 될 수 없다. 〈그림 5〉에서 보듯이 이를 공리정의하기 위해서 'DisjointClasses' OWL/XML 요소를 사용하였다.

이 공리정의는 우선어형 접근점 클래스에는 우선어형 접근점 클래스 인스턴스가 행정표준기관코드 클래스에는 행정표준기관코드 인스턴스들만 갖도록 제약한다. 특정 클래스의 특질과 같은 특질을 갖는 인스턴스들을 동일한 특정 클래스 소속하에 있게 함으로써 특정 클래스는 동일하고 단일한 특질을 갖는 인스턴스들만 포함할 수 있도록 명세적으로 정의한다.

4.2.3 정의역 · 치역 객체속성 정의

RDF/OWL은 술어논리(predicate logic)를 기반으로 자원을 기술하는데 '주어(S)-술어(P)-

```

<owl:ObjectProperty rdf:about="#hasSAC">
  <rdfs:range rdf:resource="#PreferredNameClass"/>
  <rdfs:domain rdf:resource="#SAC"/>
  <rdfs:domain>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#hasSAC"/>
      <owl:onClass rdf:resource="#SAC"/>
      <owl:minQualifiedCardinality rdf:datatype="
        &xsd;nonNegativeInteger">1</owl:minQualifiedCardinality>
    </owl:Restriction>
  </rdfs:domain>
</owl:ObjectProperty>

```

〈그림 4〉 행정표준코드 클래스 내 인스턴스의 우선어 형식 접근점과의 대응차수 정의

```

<!-- http://kanghm.corporatebody.org #PreferredNameClass -->
<owl:Class rdf:about="http://kanghm.corporatebody.org #PreferredNameClass">
  <owl:disjointWith rdf:resource="http://kanghm.corporatebody.org #SAC"/>
</owl:Class>

```

〈그림 5〉 우선어형 접근점 클래스와 행정표준기관코드 클래스 간 서로 소(疎) 공리 정의

```

<!-- http://kanghm.corporatebody.org #hasSAC -->
<owl:ObjectProperty rdf:about="http://kanghm.corporatebody.org #hasSAC">
  <rdfs:domain rdf:resource="http://kanghm.corporatebody.org #PreferredNameClass"/>
  <rdfs:range rdf:resource="http://kanghm.corporatebody.org #SAC"/>
</owl:ObjectProperty>

```

〈그림 6〉 객체속성에 대한 정의역과 치역 정의 OWL/XML 구문

목적어(O)' 형식의 트리플 구조를 갖는다. 즉, 술어인 객체속성을 중심으로 주어인 정의역, 목적어인 치역이라는 구조를 갖는다. 〈그림 6〉은 객체속성 'cb:hasSAC'를 술어로 하고 우선어형 접근점 클래스의 인스턴스들을 주어인 정의역으로 정의하고, 행정표준기관코드 클래스의 인스턴스들을 술어인 치역으로 선언한 구문이다.

이로써 객체속성이 주어와 목적어로 가질 수 있는 인스턴스의 종류들이 명세적으로 선언되어 이에 따른 RDF 트리플(예: cb:PreferredNameClass, cb:hasSAC, cb:SAC)로 표현할 수 있다. 즉, '행정안전부는 행정표준기관코드 1741000을 우선어형 접근점을 갖는다'는 구문 작성과 기타 사례에 대해서도 지속적인 RDF 트리플을 생성할 수 있다.

4.2.4 객체속성: 인스턴스 대응차수 제약

생산기관 변천연혁을 보면 일정기간에 존재하는 특정 생산기관은 오직 한 개의 행정표준기관코드를 부여를 받아 동일 신원을 유지한다. 특정한 모든 기간 내 행정표준기관코드가 서로

다르면 이는 완전히 서로 다른 신원의 독립된 생산기관이며 이 생산기관은 모든 기간 내 다른 기관과 상호 배타적으로 존재한다.

〈그림 7〉과 같이 cb:hasSAC 객체속성을 정의한 후 그 특질로 인스턴스 대응차수 1이라는 제약을 선언함으로써, 특정 행정표준기관명 클래스의 인스턴스는 오직 유일한 하나의 행정표준기관코드 클래스의 인스턴스만을 갖는다는 공리를 정의할 수 있다. 다시 말해 특정 생산기관은 오직 하나의 행정표준기관코드를 갖는 것을 명세적으로 선언한 것이다.

4.2.5 객체속성: 인스턴스 출현빈도 제약

객체속성 cb:hasSAC의 목적어인 cb:SAC 클래스 내에는 적어도 하나 이상의 인스턴스를 갖는다는 것을 공리로 정의해야 한다. 왜냐하면 행정표준기관코드 클래스 내의 인스턴스가 적어도 하나 이상은 존재해야 하기 때문이다. 즉 우선어형 접근점 클래스와 결합하는 행정표준기관코드 클래스의 인스턴스는 최소한 하나 이상 존재(출현빈도 1..n)해야 객체속성 cb:hasSAC으로 연결할 수 있다.

```

<!-- http://kanghm.corporatebody.org#hasSAC -->

<rdfs:domain>
  <owl:Restriction>
    <owl:onProperty rdf:resource="http://kanghm.corporatebody.org#hasSAC"/>
    <owl:qualifiedCardinality
      rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger">1</owl:qualifiedCardinality>
    <owl:onClass rdf:resource="http://kanghm.corporatebody.org#SAC"/>
  </owl:Restriction>
</rdfs:domain>

```

〈그림 7〉 행정표준기관코드 클래스의 인스턴스 대응차수 제약
(대응차수 1:1)

```

<!-- http://kanghm.corporatebody.org#hasSAC -->

<rdfs:domain>
  <owl:Restriction>
    <owl:onProperty rdf:resource="http://kanghm.corporatebody.org#hasSAC"/>
    <owl:someValuesFrom rdf:resource="http://kanghm.corporatebody.org#SAC"/>
  </owl:Restriction>
</rdfs:domain>

```

〈그림 8〉 행정표준기관코드 클래스의 인스턴스 출현빈도 제약
(출현빈도 1...n)

〈그림 8〉 구문은 최소한 하나 이상의 다양한 이형형식의 이름을 가진 생산기관명 클래스의 인스턴스가 객체속성 cb:hasSAC를 통해 하나의 우선어형 접근점 클래스의 인스턴스와 연결된다는 것을 선언한 것이다.

4.3 동일신원의 대표 이름형식 접근점 공리정의

하나의 생산기관명은 오직 하나의 이름 형식으로만 표현되지 않는다. 언어, 문자, 언어습관, 언론용어, 기관명 운영 정책 등에 따라 다양한 이형 형식이 이름이 존재한다. 행정안전부, 행안부, MOIS, 行政安全部, Ministry of the Interior and Safety 등은 모두 동일 정체성·동일 신원을 가리키는 접근점 표현형식들이다. 이들 중 어느 특정하나를 사용빈도, 인지 및 인식도,

편리성, 인증의 습관 등 다양한 방안을 고려하여 대표 이름형식을 편의적, 실용주의적 사고에 근거하여 하나를 선정한다. 또한 이 다양한 이름형식과 대표 이름형식 접근점 간 공고한 연결을 위해 객체속성, 클래스, 인스턴스 및 특질 선언 등을 통하여 온톨로지로 기술한다.

〈표 7〉은 동일 정체성·동일 신원의 치역 인스턴스들을 객체속성 hasSAC, hasVariousName, hasTopWord를 사용하여 정의역 클래스 내 대표 이름형식 접근점 인스턴스와 연결을 선언한 것을 표로 정리한 것이다. 여기서 대표 이름형식 접근점인 ‘행정안전부(1741000)’는 시대, 장소, 사건과는 상관없이 동일 신원의 목적어(치역) 인스턴스들을 형식적으로 대표하는 접근점 역할을 할뿐이다. 이 두 클래스 간 인스턴스 간에는 전거나 표목의 개념은 배제되었기 때문이다.

〈표 7〉 동일 정체성 · 동일 신원의 인스턴스 간 우선어 형식 정의

주어(S, 대표 이름형식)	서술어(P, 객체속성)	목적어(O, 동일신원)
행정안전부 (1741000)	hasSAC hasVariousName hasTopWord	MOI(1741000)
		행정안전부(1741000)
		행정자치부(1740000)
		MOPAS(1740000)
		行政自治部(1740000)
		안행부(1312000)
		안전행정부(1312000)

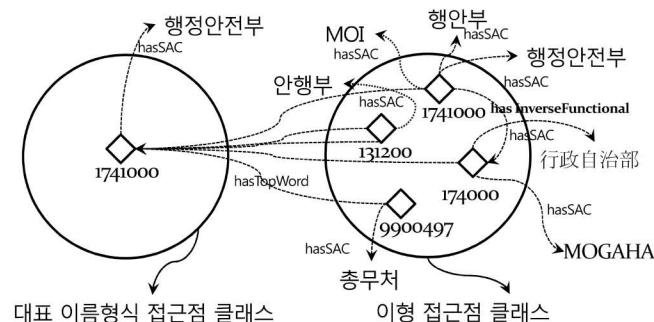
〈그림 9〉는 동일 정체성 · 동일 신원을 갖는 다양한 이형형식 접근점 클래스의 이름들은 대표 이름형식 접근점 클래스 내 인스턴스들과 객체속성 hasTopWord를 술어로 사용하여 상호 트리플을 구성하는 온톨로지 도식이다. 이형 접근점 클래스에 존재하는 다양한 형식의 접근점들은 객체속성 hasTopWord를 통해 유일한 대표 이름 형식으로 통일되어 연결되어 온톨로지 상에서 어떠한 이형 이름형식들 간뿐만 아니라 대표 이름형식과도 온톨로지로 정의됨에 따라 모든 접근점이 대등하게 서로 연결된다.

4.3.1 인스턴스 간 역함수 공리

이형 접근점 클래스 내에는 다양한 형식의

이름 인스턴스가 존재한다. 각 인스턴스는 객체속성 hasSAC를 통해 오직 하나의 대표 이름 형식의 접근점 인스턴스와 연결된다. 이때 복수의 이형형식의 인스턴스들이 동일한 대표 이름 형식의 접근점 클래스의 특정 인스턴스 1개와 동시에 연결된다면, 이 복수의 이형형식의 생산기관명 인스턴스들은 동일한 신원의 대표 이름 형식을 가리킨다는 것을 의미하므로 이를 RDF/XML 구문으로 나타내면 〈그림 10〉과 같다.

〈그림 10〉에서 만일 치역인 이형 접근점 클래스의 인스턴스인 ‘행정안전부’와 ‘行政自治部’가 대표 이름형식 접근점 클래스 내 동일한 인스턴스와 연결된다면 ‘행정안전부’와 ‘行政自治部’는 동일한 정체성과 동일한 신원의 생



〈그림 9〉 동일신원의 대표 이름형식 접근점 정의

```

<!-- http://kanghm.corporatebody.org #hasSAC -->
<owl:ObjectProperty rdf:about="http://kanghm.corporatebody.org #hasSAC">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#InverseFunctionalProperty"/>
  <rdfs:domain rdf:resource="http://kanghm.corporatebody.org #TopWord"/>
  <rdfs:range rdf:resource="http://kanghm.corporatebody.org #SCA"/>
</owl:ObjectProperty>

```

〈그림 10〉 행정표준기관코드 클래스와 생산기관명 클래스 간 역함수 관계 정의

산기관임을 온톨로지 리즈너⁵⁾는 추론할 수 있다. 즉, ‘행정안전부’와 ‘行政自治部’는 역함수 관계로 동일 클래스 내에서 서로 동일한 인스턴스를 정의한 것이다. 이것은 역함수 관계 공리의 중요 특질인데, 대표 이름형식 접근점 클래스 내 특정 인스턴스가 유일한 접근점이고 이형 접근점 클래스 내 인스턴스 사이에 역함수 특질이 정의되었다면, 이 다양한 이형식 이름들은 모두 동일한 신원의 생산기관임을 암시한다는 것이다. 이처럼 이형 접근점 클래스 내에서 동일신원을 갖는 모든 인스턴스들에게 역함수 특질을 부여하면 그 특질이 부여된 접근

점들은 동일한 정체성과 신원을 갖는 인스턴스로 시멘틱 웹의 추론기에서는 암시적 추론을 한다.

4.3.2 우선어명 접근점 대응차수 및 출현빈도 제약

객체속성 cb:hasPreferred는 〈그림 11〉에서 보는 바와 같이 정의역을 행정표준기관코드 클래스로, 치역을 생산기관명 클래스를 갖는다. 우선어명 접근점이란 앞서 언급했듯이 전거형 접근점과는 달리 전거제어(authority control)의 개념이 배제되고, 저록을 작성하는 기관의

```

<!-- http://kanghm.corporatebody.org #hasPreferred -->
<rdfs:domain>
<owl:Restriction>
  <owl:onProperty rdf:resource="http://kanghm.corporatebody.org #hasPreferred"/>
  <owl:qualifiedCardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger">1</owl:qualifiedCardinality>
  <owl:onClass rdf:resource="http://kanghm.corporatebody.org #PreferredClass"/>
</owl:Restriction>
</rdfs:domain>
<rdfs:range rdf:resource="http://kanghm.corporatebody.org #PreferredClass"/>
<rdfs:range>
<owl:Restriction>
  <owl:onProperty rdf:resource="http://kanghm.corporatebody.org #hasPreferred"/>
  <owl:someValuesFrom rdf:resource="http://kanghm.corporatebody.org #SAC"/>
</owl:Restriction>
</rdfs:range>
</owl:ObjectProperty>

```

〈그림 11〉 행정표준기관코드 클래스와 객체속성 우선어명 접근점 공리 정의

5) 리즈너(reseaner)란 owl로 정의한 온톨로지에 대해 문법상, 구문상, 의미상의 오류와 모순이 있는지 판별하여 온톨로지의 논리성과 완결성을 체크하는 일종의 엔진이다.

정책, 언어, 문자 등 다양하고 대등한 이형 접근점 중 합리적이고 적절한 이름 형식을 선정한 가변적 성격의 접근점이다. 모든 이형들은 하나의 행정표준기관코드에 연결되며 이중 우선어명 접근점은 가장 널리 알려진 이름의 생산기관명의 표현(representation) 형식을 활용할 뿐이다. 우선어명 접근점은 생산기관명 클래스의 인스턴스이며 행정표준기관코드 클래스의 인스턴스와 반드시 함수형 대응차수 1:1과 출현빈도 1...n으로 객체속성을 통해 연결된다.

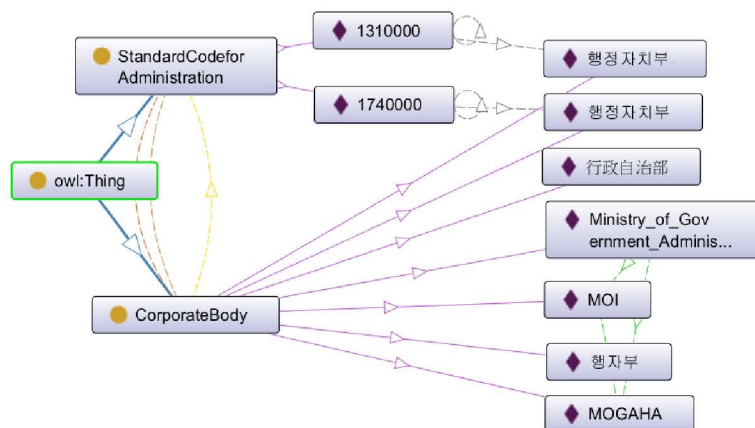
또한 객체속성 `cb:hasPreferred`의 치역에 해당하는 `cb:SAC` 클래스 내에는 적어도 하나 이상의 인스턴스가 반드시 존재해야 한다. 이때 정의역 `cb:PreferredClass`는 대응차수 1로서 객체속성 `cb:hasPreferred`에 대해 반드시 하나의 치역 클래스의 인스턴스만으로 우선형 접근점과 연결되어야 한다.

〈그림 12〉는 앞서 무전거 비제어형 우선어 형식 접근점과 동일신원 대표 이름형식 접근점에 대해 각각의 객체속성과 특질을 부여하여 기술한 온톨로지 시각화 결과화면이다. 모든 이형 형

식의 접근점들이 행정표준기관코드와 결합하여 온톨로지 상에서 유일한 식별기호체계인 IRI로 표현되어 이들 간의 우선어 형식과 대표 이름형식을 정의·기술하였다.

5. 결 론

과거 전통적인 관계형 데이터베이스 메커니즘 속에서는 특정 식별기호체계 즉 행정표준기관코드 체계는 시스템 종속적이며 지역적으로만 사용할 수 있는 구조였다. 하지만 이를 시맨틱 웹 온톨로지 환경에서 RDF/OWL의 객체속성과 클래스, 인스턴스 정의 및 객체속성 특질 및 공리정의를 통해 하나의 접근점 표현 형식인 IRI로 전환되어 모든 접근점을 온톨로지 상에서 대등한 접근점으로 정의할 수 있었다. 또한 이 IRI 접근점 간에 역함수 관계, 대응차수, 출현빈도, 동일성 등 RDF/OWL의 공리정의 및 기술을 통해 동일 정체성과 동일 신원을 가진 특정 생산기관명을 연결하여 RDF 트리



〈그림 12〉 무전거 비제어형 우선어 접근점 온톨로지 시각화

플로 표현하여 저장하면, 이 저장소에 정의되어 있는 접근점 간 관계가 온톨로지 환경에서 공유성, 재활용성, 범용성, 상호호환성 등의 특성을 갖고 모든 영역에서 접근점간의 통일된 관계 관리가 가능해진다.

이는 전통적인 출처주의 원칙과 원질서 준중의 원칙이라는 기록관리 기본실무를 달성하는데 도움이 될 것으로 판단된다. 또한 전거의 개념과 표목의 개념을 배제하고 모든 접근점을 대등한 인스턴스 개체로 인정하여 이들 간의 우선어 형식과 대표 이름형식을 자유롭게 선정

하였고, 그럼에도 불구하고 각 생산기관명의 변천과는 상관없이 동일한 정체성과 동일한 신원을 가진 모든 변천 과정에서 특정 생산기관이 생산한 모든 기록물을 일관되게 통합적으로 통시적으로 추적·식별할 수 있는 온톨로지 기반을 마련하였다.

향후의 연구과제로는 현재 기록관리 현장에서 이론적·실무적 적용이 어려운 생산기관 변천연혁을 기록물 생산기관의 유기체적 생애주기에 기반한 조직 변천 유형과 연혁을 온톨로지 관점에서 기술하는 것이다.

참 고 문 헌

- 국가기록원 (2007). 전거레코드 작성 지침. 대전: 국가기록원 표준협력과.
 행정자치부 (2008). 행정표준코드 운영 매뉴얼. 서울: 행정자치부.
 Bearman, D. A., & Lytle, R. (1985). The power of the principle of provenance. *Archivaria*, 21, 14-27.
 재인용: 설문원 (2002). 기록물을 위한 단체 전거레코드 연구. *한국기록관리학회지*, 2(2), 39-68.

• 국문 참고문헌에 대한 영문 표기
 (English translation of references written in Korean)

- NAK (2007). The description rules of authority record. DaeJeon: The Division of Standard and Cooperation of the NAK.
 MOGAHA (2008). The manual of standard administration code. Seoul: Ministry of Government Administration and Home Affairs.