

구조적 학술용어사전 구축에 있어서 지역명의 개념적 특성에 따른 관계 속성 기술에 관한 연구

A Study on Describing Relational Properties of Terms in Geographical Categories According to Conceptual Characteristics for Construction of Structured Glossary

임보람, 성균관대학교 대학원 문헌정보학과, byim@skku.edu

Bolam Yim, Dept. of Library & Information Science, Graduate School of Sungkyunkwan University

본 연구는 지역명 범주에 속하는 용어들의 개념적 특성을 분석하고, 이를 토대로 다른 범주와의 관련도를 파악하여 지역명 범주 용어들을 중심으로 관계 속성들 사이의 논리적 연관성을 부여할 수 있는 모형 도출에 기반이 되는 기초 연구이다. 지역명 범주 용어 중 국가명에 한정하여 분석한 결과, 국가명 개념 속성 중심으로는 계층 구조 관계의 지역명 범주 용어들끼리 연관이 높으며, 전체 범주 용어들의 개념 속성 중심으로는 지역명 범주 용어가 지리적 위치로서의 의미로 주로 쓰이나, 행위의 주체 또는 객체의 의미나 시대의 개념으로도 많이 활용됨을 알 수 있었다. 국가명이 참조되는 개념 속성과 연관되어 활용되는 관계 속성의 경우의 일부는 참조하는 주요 개념 범주와 연관 관계를 토대로 논리적 의미 관계를 생각해볼 수 있는 것으로 나타났다.

1. 서론

유사한 속성을 가진 학술 용어들을 동일 개념으로 범주화하여 이들을 범주별로 분류하고, 각 범주들의 개념들이 공통적으로 지니고 있는 속성들을 체계화하여 학술 용어 하나하나의 의미를 용어가 속한 개념의 속성에 따라 정의하는 사전은 구조적 학술용어사전이라 한다(고영만 외 2013). 이렇게 구축되어진 사전 데이터베이스에서는 관련성이 있는 속성들에 대한 연관관계 설정과 용어 간 관계술어 추출 및 추론규칙 생성을 할 수 있으며, 이를 통해 의미적 연관 검색을 가능하게 하는 지식베이스로 발전시키는 것이 가능하다(정보관리연구소 2012a; 박성은, 신유정 2013).

본 연구는 구조적 학술용어사전에 구축된

지역명 범주의 용어들을 중심으로 지역명의 개념 속성을 분석하고 이를 토대로 지역명 용어들을 중심으로 관계 속성 유형의 상세화를 통한 관계 속성들 사이의 논리적 연관성을 부여할 수 있는 모형 도출에 기반이 되는 기초 연구이다. 지역명 범주 용어는 특정 주제 분야나 개념 범주에 한정되어 활용되지 않고, 인문사회분야 전반에 걸쳐 다른 범주의 학술 용어의 개념 속성과 관계 속성에서 다양한 의미로 활용되기 때문에 향후에 추론 규칙 생성을 위한 분석 자료로 적합하다. 본 연구의 분석 결과는 국가명 범주의 용어만을 대상으로 한정하기 때문에 분석 결과를 지역명 범주의 모든 용어로 일반화시키는 점에는 한계가 있다.

2. 개념 정의

2.1 개념 범주

유사한 속성을 가지는 학술용어를 동일한 상위 개념으로 범주화한 것을 개념 범주라고 하며, 구조적 학술용어사전에 포함된 개념들을 집단화하기 위한 값을 말한다. 또한, 동일한 개념 범주에 속해 있는 모든 용어들은 동일한 개념 속성을 갖는 것을 원칙으로 한다(박성은, 신유정 2013; 정보관리연구소 2012b).

2.2 개념 속성

개념 속성은 특정 클래스, 즉 특정한 개념 범주가 가지는 속성들의 집합을 의미하며, 특정 개념은 해당 클래스의 속성을 상속받는다. 개념 속성 값의 유형에는 text와 object가 있다. text는 개념 속성값을 용어사전 구축자가 직접 입력하는 유형이며 object는 입력되는 값의 일관성을 유지하기 위하여 용어사전 데이터베이스에서 개념 속성값을 조회하여 입력하는 유형을 말한다(정보관리연구소 2012b).

2.3 관계 속성

관계 속성이란 두 용어에 동등관계(USE/UF), 계층관계(BT/NT), 연관관계(RT) 등의 관계를 설정한 후 이를 더 세분하여 그 근거를 제시하기 위한 수단이다. 용어사전 구축자는 일관성있고 체계적인 구축 작업을 위하여 기본 관계 유형별로 세분화되고 정형화된 관계 속성의 제시 유형을 참조하여 관계 속성을 구축하며 필요에 따라 제시되지 않은 관계 속성값을 자유롭게 입력이 가능하다(정보관리연구소 2012b).

3. 연구 대상과 분석 방법

3.1 연구 대상

연구대상은 성균관대학교 정보관리연구소에서 구축하는 인문사회분야 구조적 학술용어사전인 STNet 데이터베이스에 2014년 7월 16일 까지 구축된 용어들 중 지역명 범주의 국가명 클래스에 속한 314개의 용어들을 표본으로 한다. 국가명 용어로 한정된 이유는 다른 범주의 용어들에 비해 분류에 있어서 비교적 오류가 적어 더욱 정확한 분석이 가능하기 때문이다.

3.2 분석 방법

국가명 용어와 관련된 개념 속성과 연관 개념 범주를 알아보기 위해 국가명 범주의 개념 속성과 개념 속성값의 개념 범주를 살펴보고, 국가명을 개념 속성값으로 참조하고 있는 다른 범주의 용어들의 어떤 개념 속성에서 참조하고 있는지 알아보았다. 이를 토대로 개념 속성과 관련된 관계 속성에서도 국가명이 어느 정도로 연관 관계를 이루고 있는지를 분석하였다.

4. 분석 내용

4.1 국가명 개념 속성

국가명 개념속성 유형 값이 object인 속성에서 참조하는 개념속성 값은 어느 개념범주와 관련되어 있는지 살펴보면 <표 1>과 같다.

<표 1> 국가명 개념속성과 관련된 개념 범주

개념범주 개념속성	국가명	각국어	도시 구동명	대륙명	기타	총합계
인접국	492		2		28	522
언어		226			60	286
위치	24			202	47	273

수도	2		214		12	228
원지명	8				10	18
동맹국						0
총합계	526	226	216	202	157	1,327

* 기타는 클래스가 미분류된 용어 또는 구축자의 판단 오류로 클래스가 명백히 잘못 분류된 경우를 나타낸다.

국가명 개념 속성 값이 속해 있는 클래스를 살펴보면 국가명, 각국어, 도시구동명, 대륙명으로 크게 나누어져 있음을 알 수 있다. 특히 국가명, 도시구동명, 대륙명은 지역명 범주안에 있는 클래스로서, 관련 클래스의 대부분을 차지하고 있음을 보아 국가명은 지역명 범주 용어와 개념적으로 관련도가 높음을 알 수 있다. 또한 도시구동명, 대륙명은 국가명을 중심으로 지역명 범주에서 계층 구조 관계를 이루고 있는 속성들 이므로 도시구동명과 대륙명 또한 서로 논리적으로 연관 관계를 이룬다.

4.2 국가명이 참조되는 주요 개념 속성

모든 범주 용어의 개념 속성에서 국가명이 개념 속성값으로 참조되는 경우에 주로 어떤 범주의 개념 속성에서 참조되는지 개념 속성에 관련된 관계 속성별로 분류하면 <표 2>와 같다.

<표 2>에서 알 수 있듯이, 다양한 범주의 용어의 개념 속성에서 국가명이 사용되었는데, 관련 관계 속성으로 살펴보면, 'hasLocation' 관계로 보아 지리적 위치개념인 개념 속성으로 국가명이 많이 참조됨을 알 수 있다. 그러나 지리적 위치 개념 이외에도 'isProducedBy', 'hasAgent', 'hasPatient', 'hasCreator'와 같은 관계 속성에 해당하는 개념 속성, 즉 행위의 주체 또는 객체의 의미로도 많이 참조된 것을 알 수 있다.

예를 들어 사건명 범주 용어인 '한일회담'의 개념 속성 중 '주관자'에 국가명인 '일본'이 참조된 경우는 국가명이 지리적 위치 개념이 아닌 정부의 의미로 사용된 점을 알 수 있다.

<표 2> 국가명이 참조되는 개념 속성과 관련 관계 속성

관련 관계 속성	관련 범주	개념 속성	참조된 국가명 수
hasLocation	인명	국적	1,641
		출생지	1,062
	지명 (도시구동명, 국가명)	위치	1,302
		인접국	492
		원지명	8
	이론	해당국가	846
	사건명	장소	230
	기념물명	소재지	209
	상품명	생산국	139
	조약/협약	관련국가	90
	정책(제도)명	시행국가(지역)	24
	자연물	분포지	7
	지명	수도	2
isProducedBy	사건명	관련자	93
		주관자	82
	기관/단체명	설립자	9
상품명	제조자	8	
hasAgent	관계/상호작용	주체자	23
	행위, 변화/변동	행위자	15
hasEra	이론, 인명	시대	121
		기념물명	제작시대
hasPatient	특성	대상자	35
		행위, 변화/변동	행위 대상자
hasCreator	기념물명	제작자	6
관련 관계속성 없음	기념물명	소유자(처)	72
		관리자(처)	4
	인간(사회집단)	민족/인종	10

이외에도 지역명이 시대의 의미로 쓰인 점을 'hasEra' 관계 속성과 관련된 인명 또는 이론의 시대 개념 속성에서 많이 쓰임을 알 수 있다. 이는 국가명이 '조선시대', '고려시대' 같이 정권의 의미로 해석될 수 있다.

4.3 국가명이 참조되는 관계 속성 유형

<표 3> 국가명이 참조되는 관계속성

관계 속성	참조된 국가명 수	참조하는 주요 범주
hasLocation	463	각국어, 지명 범주, 행위 등 전반적인 범주
isProducedBy	14	사건명
hasEra	6	제도, 조직, 유형(장르) 등
hasAgent	5	사건명

* <표 3>의 관계 속성은 <표 2>와 연관된 관계 속성만 나타내었다.

<표 3>과 같이 전반적인 범주의 용어들이 ‘hasLocation’ 관계 속성을 사용하여 국가명을 참조하였는데, 주로 용어가 해당하거나 관련된 지역이라는 의미로 많이 활용되었다. 예로 사건명 범주와 국가명이 hasLocation 관계로 연결이 되었을 경우 연결된 국가명은 장소, 즉 지리적 공간을 의미한다. 그러나 ‘isProducedBy’ 관계 유형으로 연결된 경우에는 ‘사건명’의 ‘주관자(isProducedBy)’는 ‘국가명’인데, 이 경우에 우리는 국가명이 지리적 공간이 아닌 정부의 의미로도 쓰였음을 알 수 있다. 따라서 사건명과 국가명이 어떤 관계 속성으로 연결되느냐에 따라서 국가명이 지리적 공간을 의미할 수도 있고 행위의 주체, 즉 정부를 의미할 수도 있다. 이와 관련하여 ‘사건명의 주관자(isProducedBy)’와 대상자(hasPatient)가 국가명일 경우, 주관자와 대상자는 해당 사건명과 관련하여 서로 영향 관계에 놓인 국가일 가능성이 있으며, 이 때 국가는 정부의 의미를 가진다.’라는 논리적 의미 관계도 생각해 볼 수 있다.

5. 결론

본 연구는 인문사회분야 구조적학술용어사전에 구축된 지역명 범주 용어 중 국가명에

한정하여 관련된 개념적 특성을 분석하고 이를 토대로 지역명 용어들을 중심으로 관계 속성 유형을 분석해보았다. 지역명 용어가 지리적 위치로서의 의미로 주로 쓰이긴 하나, 정부와 같이 행위의 주체나 객체의 의미나 시대의 개념으로도 많이 활용됨을 알 수 있었다. 이에 따라 활용되는 관계 속성도 다양해지는데, 개념 속성과 연관되어 활용되는 관계 속성 경우의 일부는 참조하는 주요 개념 범주와 연관 관계를 토대로 논리적 의미 관계를 생각해볼 수 있는 것으로 나타났다. 본 연구 결과는 구조적 학술용어사전의 지식베이스 구축에 있어 공리와 추론 규칙 생성의 기초 연구로서 의의가 있다. 향후 지역명처럼 전반에 걸쳐 활용되는 범주의 용어를 대상으로 후속 연구가 이루어진다면 더욱 실용적인 구조적 학술용어사전의 지식베이스 구축에 도움이 될 것이다.

참고문헌

- 고영만 외. 2013. 인문학 및 사회과학분야 국내 학술논문의 저자키워드 출현빈도와 피인용횟수의 상관관계 연구. 『정보관리학회지』, 30(2): 227-243.
- 박성은, 신유정. 2013. 구조적 학술용어사전 구축에 있어서 인문사회분야 복합명사의 한자어 접미사 형태분석에 의한 용어의 개념화에 관한 연구. 『한국정보관리학회 학술대회 논문집』, 20(-): 123-128.
- 정보관리연구소. 2012a. 2012년도 기초연구 지원인문사회(토대연구) 연구계획서.
- 정보관리연구소. 2012b. 구조적 용어사전 구축지침.