

온톨로지의 개념간 관계 설정을 위한 AGROVOC 시소러스의 분석에 관한 연구

A Study on the Analysis of AGROVOC for Establishment of
Concept Relationships of Ontology

유 영 준(Yeong-Jun Yoo)*

초 록

이 연구에서는 AGROVOC 시소러스의 개념간 관계를 분석하여 시소러스의 의미 관계의 모호성과 비일관성을 밝히고, 이러한 단점들을 개선한 온톨로지의 개념간 관계를 제시하였다. 개념간 관계 분석의 결과로 온톨로지의 개념간 관계의 핵심 요소인 개념 모형과 의미론적으로 발견된 개념간 관계 유형을 제시하였다. 이 관계들은 부분적으로 추론 기능을 수행할 수 있으며 보다 명확한 의미 관계를 기반으로 하는 지식조직시스템에 적용할 수 있을 것이다. 그리고 시소러스의 개념간 관계 유형을 확장하는데 이용할 수 있는 새로운 관계 유형들을 밝혀내었고, 이 관계 유형들이 법률분야 관련어집과 같은 기존 시소러스에도 활용할 수 있음을 확인하였다.

ABSTRACT

This study uncovered ambiguity and inconsistency of the semantic relationships of the existing thesaurus by analyzing the concept relationships of AGROVOC and proposed the concept relationships of ontology in partially overcoming these limitations. By the results of analyzing the concept relationships, the study proposed conceptual model as most important part of concept relationships of ontology and semantically developed concept relationship types. These relationships partially can perform inferences and must be useful for information knowledge system based on more exact semantic relationships. Also the study found out new relationship types and they will be useful for extension of the concept relationships of existing thesaurus. And these relationship types showed that they were useful for the existing thesaurus as Legal Thesaurus.

키워드: 개념간 관계, 개념간 관계 유형, 개념 모형, 시소러스, 온톨로지, AGROVOC
concept relationships, conceptual model, thesaurus, ontology

* 연세대학교 문헌정보학과 강사(yp6370@hanmail.net)

- 논문접수일자 : 2005년 2월 13일
- 게재확정일자 : 2005년 3월 12일

1. 서론

1. 1 연구의 목적 및 배경

문헌정보학의 학문 대상인 지식은 개념이라는 단위로 이루어지며, 개념간의 관계에 따라 지식은 일정한 구조를 갖는다. 이러한 지식 구조는 개념간의 관계에 의해 구성되는 것으로 분류체계, 시소러스, 온톨로지와 같은 구체적인 형태로 표현될 수 있다.

시소러스나 온톨로지에서 개념간의 관계는 의미관계로 볼 수 있는데, 시소러스에서는 용어간의 개념관계를 명확하게 제시하는 것이 검색효율과 관련해서 매우 중요하다. 그러나 시소러스의 의미관계가 갖는 단점 중의 하나는 검색 대상이 되는 주제 영역 내에서 개념을 식별하는 능력이 떨어진다는 것이다. 예를 들어 일반적으로 이용되고 있는 텍스트 입력방식의 검색엔진의 경우 동형의어를 구분하지 못하는 결과를 낳게 된다. 시소러스도 이와 유사한 문제를 일으킬 수 있는데, 그 이유는 시소러스를 구성하는 용어들이 계층구조로 이루어져 있지만 광범위하고 기초적인 관계들로만 구조화되어 있기 때문이다. 따라서 개념과 용어간의 구별이 일관성 있게 이루어지지 못하고, 인간이 개념과 언어의 연결을 통해 사물을 인지하는 방식을 제대로 반영하지 못하는 결과를 나타내게 되는 것이다.

그리고 대부분의 시소러스는 개념을 유기체, 재료, 과정 등의 범주로 구분하지 않고, 계층관계(BT/NT)와 연관관계(RT) 정도만으로 구분하여 매우 제한적인 개념간의 관계들로 구조화하고 있다. 이러한 개념간 관계로는 이용자

에게 의미론적 정보 검색(meaningful information discovery)을 제공할 수 없으며, 온톨로지에서 필요로 하는 추론 기능을 지원하기는 더 어렵다고 할 수 있다. 따라서 시소러스의 개념간 관계들은 이용자들의 질의를 확장하거나 질의를 보다 구체적으로 표현하기 위해 개념들을 이용하는 지식조직시스템(knowledge organization system)에서는 적용하기 어렵다. 예를 들어 'cow'와 'cow milk'간의 관계가 하위어(narrower term)로 표현되는데, 이러한 방식으로 개념간의 관계를 표현하는 것은 개념간 관계가 모호해질 수 있다. 오히려 <hasComponent>처럼 의미를 명확하게 표현하는 것이 더 바람직하며, 시소러스의 관계 유형만으로는 웹 자원을 검색하거나 의미적으로 기술하는데 적합하지 못하며, 다양하고 심층적인 개념간의 관계들이 필요하다.

결과적으로 용어간의 정확한 의미를 구조화하지 못하는 시소러스에서는 개념간 관계가 일관성을 유지하지 못하고 모호성을 갖는다. 계층관계(BT/NT)에 속하는 사례가 비계층적 관계인 연관관계(RT)로 표현되거나 그 반대의 사례도 있다. 또한 일반적으로 시소러스는 수작업 색인과 질의어 작성 등을 목적으로 설계되어서 모호한 의미 관계로 인하여 자동화 처리(automated processing)에 적용하기에는 부적합하다.

이러한 한계들을 극복하고 웹상에서 효율적으로 활용하기 위해서, 시소러스는 보다 정밀하게 개념간의 관계들을 정의하고 해당 개념을 표현하는 용어를 구조화한 주제영역의 개념들로 조직되어야 한다. 이와 함께 연관된 개념들을 일반화하고 추론 기능이 가능하도록 타당성

있는 규칙과 조건들이 구체적으로 제시되어야 할 필요가 있다.

그리고 온톨로지는 개념 관계를 세분하여 차별화할 수 있는 지식조직시스템이라고 할 수 있으며, 개념과 용어를 분리시킴으로써 해당 주제 영역을 파악하는 인간의 이해 구조를 더 잘 반영한다. 또한 주제 영역 내에서 일관성 있고 명확하게 각각의 개념을 정의하고, 개념간의 관계들을 구조화함으로써 의미구조가 구축된다. 따라서 온톨로지의 개념간 관계는 지식 모형화의 대상이 되는 해당 주제 영역의 특성을 보다 더 분명하게 반영할 수 있다.

또한 기존의 시소러스보다 정확하고 분명한 의미구조를 갖는 온톨로지에서는 구조화된 지식으로부터 새로운 지식을 추론할 수 있다. 이 새로운 지식은 일반화 혹은 상호운용성 규칙을 적용함으로써 도출할 수 있는데, 이러한 점은 적용 범위가 제한되어 있는 시소러스와 같은 지식조직시스템과는 구별된다. 이러한 새롭게 덧붙여지는 지식은 지능적인 정보 처리에 적용될 때 보다 큰 역할을 할 수 있을 것이다.

따라서 이 논문에서는 시소러스의 개념간 관계를 분석함으로써 온톨로지의 개념간 관계의 핵심인 개념 모형과 의미론적으로 보다 발전된 개념간 관계 유형을 제시하고자 한다. 그리고 이러한 개선된 개념간 관계 유형의 제시는 이용자에게 보다 나은 검색 서비스를 제공할 수 있는 지식조직시스템에 필수적인 요소가 될 것이다.

1. 2 연구의 방법

이 논문에서는 AGROVOC 시소러스의 개

념간 관계를 분석하여 온톨로지에서 개념간의 의미 관계를 보다 세밀하고 분명하게 기술할 수 있는 방안을 제시함으로써 시소러스의 개념간의 의미 관계를 어떻게 개선할 것인가를 제시하였다. 먼저 AGROVOC을 분석 대상으로 선택한 이유는 다음과 같다. AGROVOC은 시소러스의 기본적인 개념간 관계로만 구축된 가장 일반적인 시소러스라고 할 수 있고, 이 분석 결과를 다른 시소러스의 개념간 관계를 분석하는데에 기초자료로 활용할 수 있을 것이라고 판단했기 때문이다.

AGROVOC은 농업, 임업, 어업, 식량 및 환경 등과 관련된 주제 영역의 디스크립터 16,500개와 비디스크립터 10,000개로 구성되어 있고, 통제 어휘집에 의한 색인 및 탐색을 목적으로 관련 주제 영역의 문헌과 정보 자원을 기술하기 위해 유엔 산하의 국제식량기구(Food and Agriculture Organization)과 유럽공동체위원회(the Commission of the Europe Communities)에 의해 1980년대 초에 개발되었다. 또한 국제식량기구(FAO)의 공식 언어인 영어, 불어, 스페인어, 중국어, 아랍어, 체코어, 포르투갈어로는 온라인으로도 이용가능하고, 그 외에 다수의 언어로도 번역되었다.

또한 온톨로지의 개념간 관계 유형을 기존의 시소러스인 법률분야 관련어집에 적용하여 온톨로지의 개념간 관계가 시소러스의 개념간 관계에 적용할 수 있는지를 실험해 보았다.

2. 이론적 배경

2. 1 시소러스와 온톨로지의 개념간 관계

시소러스의 개념간 관계와 온톨로지의 개념간 관계를 분석함으로써 시소러스의 개념간 관계에서 논리적이고 합리적인 온톨로지의 개념간 유형을 도출할 수 있는 근거를 살펴보고, 추론 기능처럼 시소러스에서는 수행할 수 없었던 기능이 온톨로지에서는 어떻게 가능할 수 있는지에 대해서도 검토하겠다. 먼저, ERIC(Educational Resources Information Center) 시소러스에

서 정의한 관계들을 온톨로지에서도 정의할 수 있는 관계들의 예와 비교하여 <표 1>에 제시하였다.

온톨로지는 <표 1>과 같은 구체적인 의미 관계를 바탕으로 <그림 1>(Soergel et al. 2004)과 같은 추론 규칙들을 생성할 수 있지만, ERIC(Educational Resources Information Center Thesaurus)은 상위어(BT)와 연관어(RT) 관계만을 정의하고 있으므로, 이들 관계는 추론 규칙을 생성할 수 있을 만큼 의미 관계가 분명하지 못하다.

<표 1> 온톨로지의 개념 관계와 ERIC 시소러스의 개념 관계의 비교

ERIC 시소러스	온톨로지
reading instruction BT instruction RT reading RT literary education	reading instruction <isa> instruction <hasDomain> reading <governedBy> literary education
reading ability BT verbal ability RT reading RT reading attitudes	reading ability <isa> verbal ability <hasDomain> reading <supportedBy> reading attitudes

규칙 1
Instruction in a domain should consider ability in that domain: X shouldConsider Y IF X <isa(type of)> instruction AND X <hasDomain> W AND Y <isa> ability AND Y <hasDomain> W yields: The designer of reading instruction should also consider literary education.
규칙 2
X shouldConsider Z IF X <shouldConsider> Y AND Y <supportedBy> Z yields: The designer of reading instruction should also consider reading attitudes.

<그림 1> 온톨로지의 의미 관계에 의한 추론 규칙

다른 예로, 온톨로지의 개념간 관계와 규칙을 가정할 수 있는데, 이러한 관계들은 AGROVOC 시소러스에서 추출한 것이다. 개념간 관계는 아래 <표 2>와 같고, 개념간 관계에 의해 도출한 추론 규칙은 <그림 2>(Soergel et al. 2004)와 같다. 이러한 추론 규칙에 의해 새로운 개념간 관계를 도출해 낼 수 있는 것이다.

<표 2>에서 'cow milk'와 'milk fat'이 'milk'의 NT 관계로 표현되지만 온톨로지에서는 두 하위어를 <includesSpecific>과 <containsSubstance>를 사용하여 의미 관계를 정확하게 규명하여 표현함으로써 용어간의 의미 관계를 더 정확하게 구조화 하고 있다.

<그림 2>에서는 온톨로지의 추론 기능을 가 능하게 하는 추론 규칙의 예를 보여 주고 있는

데, 온톨로지가 갖는 진술과 규칙에 의해 시스템은 'Cheddar cheese'(체다 치즈)가 'milk fat'(유지방)을 함유하고 있다는 것을 <containsSubstance> 관계에 의해 추론할 수 있다. 그러나 AGROVOC 시소러스는 하위어(NT)/상위어(BT) 관계만을 설정할 수 있어서, 추론 과정에 의한 지능적인 개념간 관계를 구조화할 수 없다.

시소러스를 온톨로지와 같은 구조로 변환시키기 위해서는 시소러스의 의미 관계들을 명확하게 표현해야 하는데, 이때 가장 중요한 것은 일정한 패턴이 있는지를 분석하는 것이다. 이 패턴들을 근거로 규칙을 생성하고 차후에 발생하는 유사한 사례에 적용할 수 있으며, 그 사례 들은 다음과 같다.

<표 2> AGROVOC의 개념간 관계와 온톨로지의 개념간 관계의 비교

AGROVOC	온톨로지
milk NT cow milk NT milk fat	milk <includesSpecific> cow milk <containsSubstance> milk fat
cow NT cow milk	cow <hasComponent> cow milk
Cheddar cheese BT cow milk	Cheddar cheese <madeFrom> cow milk

규칙 1
Part X <mayContainSubstance> Substance Y IF Animal W <hasComponent> Part X AND Animal W <ingests> Substance Y
규칙 2
Food Z <containsSubstance> Substance Y IF Food Z <madeFrom> Part X AND Part X <containsSubstance> Substance Y

<그림 2> 온톨로지의 의미 관계에 의한 추론 규칙

1) 'cow NT cow milk'는 'cow <hasComponent> cow milk'로 변환한다.

2) 'animal <hasComponent> milk 또는 더 일반적인 경우는 'animal <hasComponent> body part 라는 관계를 보편적인 패턴의 예로 인식한다.

3) 위와 같은 패턴에 의해, 'goat NT goat milk'가 'goat <hasComponent> goat milk'라는 것을 시스템은 추론할 수 있다. 그 이유는 염소는 동물이고 염소의 젖은 'milk'라는 단어로 끝나고 따라서 젖의 한 유형이 된다는 것을 알 수 있기 때문이다.

이와 같은 방법을 보다 자동화시키기 위해서는 'animal <hasComponent> body part'와 같은 패턴 리스트(inventory of patterns)를 작성해야 한다. 'cow, goat, sheep, horse, chicken' 등과 같은 'animal' 유형의 개념들과 'skeletal meat part, liver, bone, milk, egg' 등과 같은 'body part'와 같은 유형의 개념들을 명시하는 온톨로지를 구축해야 한다.

이러한 패턴들은 제약 조건(constraints)의 특수한 유형들인데, 시소러스의 개념 관계들을 개선하기 위해 변환해야 할 선택사항들을 제한

하기 위해 사용될 수 있다. 그러한 시소러스의 개념 관계들에 의해서 제약 조건들이 설정되고, 다른 한편으로는 개념들의 엔티티/패시 유형도 시소러스의 개념 관계에 의해 설정된다. <표 3>은 시소러스의 개념 관계에 의해 정해진 제약 조건들의 예를 보여주고 있다(Soergel et al. 2004).

위의 패턴 리스트는 시소러스의 개념간 관계를 보다 분명한 온톨로지의 개념 관계로 변환할 때, 어떤 선택이 가능한지를 제한할 수 있을 것이다. 물론 위의 표에 제시한 개념간 관계가 절대적인 것은 아니다.

2.2 선행연구

시소러스의 개념간 관계에 대한 연구로는 국내의 20개 시소러스의 기본 관계와 부가 관계를 분석한 한상길(1999, 2000)의 연구가 있는데, 기존 시소러스의 개념간 관계의 한계를 극복하고 온라인 검색 환경에 적합한 방향으로 개념간 관계를 확장하는 데 초점을 맞추었다. 그리고 김태수(2001)의 연구에서는 용어정의 를 도입한 시소러스의 개발에 관한 실험적 연

<표 3> 개념간 관계의 제약 조건

시소러스의 개념간 관계	온톨로지 개념간 관계
NT/BT	<hasMember> / <memberOf> <includesSpecific> / <isa> <hasComponent> / <componentOf> <spatiallyIncludes> / <spatiallyIncludedIn> etc.
RT	<similarTo> <growsIn> / <EnvironmentForGrowing> <treatmentFor> / <treatedWith> <hasMember> / <memberOf> etc.

구를 수행하였다. 김태수의 연구가 갖는 의의는 용어 정의를 시소러스에 도입함으로써 디스크립터의 의미 범주를 한정할 수 있고, 동시에 정의에 사용된 용어의 의미관계를 디스크립터 관계구조로 표현할 수 있어서 의미범위와 관계구조의 표준화 가능성을 기대할 수 있다는 점이다. 그리고 용어의 정의를 분석함으로써, 개념이 지닌 본질적 속성을 기본 데이터요소로 하는 용어의 속성과 관계구조에 기반한 용어 정의 시소러스를 구축하기 위한 것이다(김태수 2001). 따라서 용어 정의를 도입한 시소러스가 지향하는 바도 용어의 개념적 속성을 관계구조에 반영하려는 시도라고 할 수 있다. 위의 세 연구와 동일한 맥락에 있는 남영준(2004)의 연구는 기존의 시소러스가 명확하게 표현하지 못한 용어간의 대등/동의 관계를 확장함으로써 시소러스에서 디스크립터간의 의미 관계를 보다 심층적으로 구축하는데 중요한 지침이 될 수 있을 것이다.

그리고 현재 문헌정보학 영역에서는 위에 제시한 시소러스에 관한 연구와 함께 온톨로지에 관한 연구가 함께 이루어지고 있다. 그러나 대부분의 온톨로지에 관한 연구는 새로운 방법론에만 치우침으로써 기존의 문헌정보학의 지적 내용에 대한 분석은 찾아보기가 어렵고, 더욱이 시소러스와 온톨로지의 연관성을 찾으려는 시도는 부족한 것으로 보인다. 그럼에도 불구하고 시소러스와 온톨로지의 연관성에 근거한 정도현(2003)의 연구는 패싯형 시소러스인 AAT(Art and Architecture Thesaurus)의 일부분을 웹 온톨로지 언어(OWL)로 구축한 연구로서, 온톨로지의 공리(axioms)를 이용해 기존의 온라인 시소러스를 통해 얻을 수 없었던 개념간의 새로운 관계를 정의할 수 있는 방안을

살펴 본 연구라고 할 수 있다. 기존의 시소러스에 대한 분석을 기반으로 함으로써 시소러스의 지식 구조를 활용하였고, 시스템 구축 도구로서 웹 온톨로지 언어를 채택한 시스템구축사례로 볼 수 있다. 문헌정보학 영역에서 이미 개념간의 관계를 특정하게 표현하는 패싯 체계에 기반한 시소러스의 지식구조의 장점(손대형 1998)을 최대한 살린 것이 돋보이는 점이라고 할 수 있다. 그리고 조현양과 남영준(2004)의 연구는 OWL(Web Ontology Language) Lite의 어휘를 시소러스의 기본 관계에 대입하여 시소러스와 온톨로지의 호환 가능성을 입증하였고, 다국어 시소러스 구축에 대한 방법론도 제시하였다.

그러나 국내의 연구들이 온톨로지의 추론 시스템과 관련된 부분에 관한 연구를 다루지 못한 반면 Soergel 등(2004)의 연구에서는 시소러스의 개념간 관계를 분석하여 온톨로지의 새로운 개념간 관계 유형을 제시함과 동시에 추론 규칙의 활용 가능성에 대한 시사점을 제시해 주었다.

이렇게 문헌정보학의 지식구조에 대한 의미 있는 분석을 바탕으로 시스템 구축 도구로 웹 온톨로지 언어를 활용하는 연구가 문헌정보학 영역의 축적된 지식과 자원을 제대로 이용할 수 있는 방향이 될 것이다.

3. AGROVOC의 개념 구조

3.1 AGROVOC의 개념간 관계 유형

3.1.1 동의 관계 및 계층 관계

동의 관계는 'USE' 관계와 'USED FOR'(UF)

관계로 일반적으로 두 용어간의 동의성을 표현하고, 두 관계는 역관계이다.

그리고 계층관계의 하위어(NT)를 조건구문으로 설명하면 다음과 같다. X가 Y의 하위어(NT) 일 때, X는 Y보다 의미적으로 하위어이다. 예를 들면 다음과 같다.

milk NT cow milk, grain NT rice

그리고 상위어(BT)는 다음과 같은 조건구문으로 설명할 수 있다. Y가 X의 상위어 일 때, Y는 X보다 포괄적 이다. 예를 들면 다음과 같다.

cow milk BT milk, rice BT grain

결과적으로 BT는 NT의 역관계이다.

3. 1. 2 연관 관계

관련어(RT)는 계층 관계가 아닌 두 용어간의 관계를 표현하기 위한 것으로 가장 모호한 관계라고 할 수 있다.

3. 1. 3 범위 주기

AGROVOC에서도 범위 주기를 갖는 디스크립터들이 있는데, 범위 주기는 용어의 정의, 내력 주기, 색인자나 탐색자에 대한 지시사항이나 간략한 설명 등이다. 범위 주기는 이용자에게 해당 용어에 관한 더 구체적인 내용과 용례를 제공한다.

3. 1. 4 최상위 수준 구조

현재 AGROVOC은 1,500개가 넘는 최상

위어를 가지고 있으나, 최상위어에서 시소러스를 검색하거나 전체 계층을 브라우징하기가 쉽지 않다. 따라서 AGROVOC에서는 AGRIS (International Information System for the Agricultural Sciences and Technology) 범주화 체계를 함께 사용하고 있다.

3. 2 AGROVOC의 개념간 의미 관계의 모호성

3. 2. 1 동의 관계

AGROVOC에서는 USE/UF를 사용해서 동의 관계를 설정한다. 그러나 동의 관계에 의해 유사 동의어 및 매우 특정한 하위 용어들을 연결하는 데에도 이용되고 있으며, 따라서 AGROVOC에서는 동의 관계에 대해 다음과 같이 정의하고 있다:

1. “두 개념이 한 디스크립터에 의해 동의 관계로 식별될 정도로 충분히 유사할 때”
2. “특정 개념과 반의어”
3. “매우 특정한 개념들이 하나의 디스크립터에 의해 포섭될 때”

1번 정의는 의미적으로 밀접하게 관련된 개념을 다루고 있지만, 동의어로 볼 수 없다. 그 예는 아래와 같다.

famine(기아)

UF hunger(배고픔)

2번 정의는 의미 기준에서 정반대편의 개념들을 포함하고 있다. 이 용어들은 반의어를 지시하고 있음에도 동의 관계에 포함되어 있다.

예는 아래와 같다:

hydrophilicity(물과 친한 성질)
 UF hydrophobicity(물을 두려워 하는 성질)

반의어에 관한 의미 관계 분석에 있어서, 김태수(2000)는 반의어를 연관 관계로 설정하였는데, 그 이유는 마음속에 저장되어 있는 용어들이 유기적인 의미망으로 조직되어 있기 때문에 검색효율의 개선을 위해 연상의 강도가 높은 용어를 시소러스에 도입했다고 밝혔다. 그러나 남영준(2004)은 반의어를 대립어(UFC: used for coordinate words)의 하위 개념으로 파악하고 시소러스를 활용하는 시스템이나 환경을 고려하여 구축할 수 있다고 기술하면서, 대등/동의 관계에 포함시켜서 UFC/A(used for antonym)라는 새로운 대등/동의 관계 기호를 부여하였다. 따라서 시소러스마다 반의어를 상이한 개념간 관계 유형에 배치하고 있음을 알 수 있고, 반의어의 의미 분석에 대한 명확한 기준을 제시하지 못함으로 해서 의미 관계의 모호성은 여전히 해결해야 할 문제로 남아 있다.

3번 정의는 동의 관계(USE/UF)가 계층 관계를 표현할 수도 있다는 것을 나타내고 있다. 예는 다음과 같다:

biological competition(생물학적 경쟁)
 UF interspecific competition(중간의 경쟁)
 UF intraspecific competition(중내의

경쟁)

'interspecific competition'과 'intraspecific competition'간의 적절한 구별은 검색에서는 필요하지 않은 것으로 보인다. 따라서 더 일반적인 범주로 표현하는 것이 바람직하다. 그리고 위의 예를 김태수(2000)와 남영준(2004)은 모두 상향처리(posting)로 인식하고, 이와 같이 상향처리를 하는 이유를 어휘의 수를 줄이거나 시소러스의 규모나 수준에 따라 대등/동의 관계로 처리한다고 기술하였고, 개념적이거나 논리적인 방법과는 무관한 것으로 생각하였다.

3. 2. 2 계층 관계

계층 관계를 설정하기 위해 사용된 BT/NT의 의미 관계가 명확하지 않은 사례가 있는데, 그 이유는 상이한 관계 유형을 합성해서 관계를 설정했기 때문이며 그 예들은 다음과 같다.

1) <includesSpecific> 관계

'적혈구(erythrocytes)는 혈액 세포(blood cell)의 특정한 한 종류이다'라는 명제는 다음과 같은 관계로 표현된다.

blood cell
 NT erythrocytes(적혈구)
 NT leukocytes(백혈구)

이 관계는 김태수(2000)의 시소러스의 의미 관계에 따르면 계층 관계 중에서 부류포섭관계와 동일한 의미 관계다.

2) <hasComponent> 관계

'혈액(blood)은 구성요소인 혈액 세포들 (blood cells)을 내포한다'라는 명제는 다음과 같은 관계로 표현된다.

blood
NT blood cells

이 관계는 김태수(2000)의 시소러스의 의미 관계에 따르면 계층 관계 중에서 부분-전체포섭관계와 동일한 의미 관계다.

3) 다음은 불분명하게 정의된 관계 모형화 원칙을 적용한 상이한 시소러스들 간의 불일치를 분명하게 보여주는 사례이다.

AGROVOC과 CABI(Commonwealth Agricultural Bureau Incorporate) 시소러스에서는 다음과 같이 관계를 정의하고 있다.

water
NT ice
NT water vapor
...
NT fresh water
NT drinking water

그러나 ASFA(Aquatic Sciences and Fisheries Abstracts) 시소러스에서는

water
RT ice
RT water vapor

...

NT fresh water 로 정의 되어 있다.

'water vapor'와 'ice'는 물의 상태를 표시하는 개념이고, 'fresh water'와 'drinking water'는 물의 종류를 지시한다. 따라서 AGROVOC과 CABI에서는 계층 관계가 의미 속성이 상이한 개념 관계를 일괄적으로 묶어서 정의하고 있다. 그리고 ASFA에서는 물의 상태와 관련된 관계 유형을 관련어(RT)로 처리하고 있지만, 관계 유형들을 통합하는 사례는 일관성을 상실하게 된다(Soergel et al. 2004). 따라서 상이한 의미 속성을 지니는 계층 관계들을 통합해서 정의하는 것은 용어가 지닌 개념의 본질적 속성을 제대로 표현하지 못할 뿐만 아니라 용어간의 의미 관계를 불분명하게 만들고 결과적으로 의미 관계의 일관성을 상실하게 된다.

3. 2. 3 연관 관계

BT/NT 관계처럼, 연관 관계인 RT도 보다 더 특정한 관계로 개선되어야 하며, 관련된 예들을 제시하면 다음과 같다.

1) <hasMember> 관계(Anglophone Africa <hasMember>Botswana)

Anglophone Africa(영어를 사용하는 아프리카 국가)

RT Botswana
RT Gambia
RT Ghana
RT Kenya

RT Lesotho

...

‘영어를 사용하는 아프리카 국가’라는 개념과 실제로 영어를 사용하는 개별 국가 간의 관계는 기존의 시소러스에서 표현하기 어려운 관계 유형이라고 할 수 있다. 그러나 김태수(2000)는 계층관계 중에서 사례관계의 정의 즉 “사례관계는 사물의 범주와 사례와의 관계로서 고유명으로 표현되는 해당 범주의 개개 사례가 모여 상위의 개념을 형성하게 된다”고 하였으나, 개별 사례가 상위 범주를 형성한다는 논리가 AGROVOC에서는 그대로 적용될 수 있을지 명백하지 않다. 오히려 개별 사례가 상위의 범주를 형성하는 것으로 보기 어렵고, 이미 ‘영어를 사용하는 아프리카 국가’라는 개념을 상징하고 개별 국가들을 파악하는 하향적 논리가 더 타당하다고 볼 수 있다. 따라서 AGROVOC과 김태수의 관계 유형은 서로 일치하지 않으며 시소러스간의 비일관성이 존재한다는 증거라고 할 수 있다.

2) <causes> 관계(bleaching<causes>discoloration)

bleaching(표백)

RT discoloration(변색)

표백 작용의 결과로 변색이라는 현상이 나타나게 되는데, 이 두 용어간의 관계를 단순히 연관 관계만으로 설정하는 것은 의미 관계를 명백하게 표현했다고 할 수 없다.

위에 제시한 예에서 AGROVOC에서의 개념간 관계가 모호한 측면이 있음을 확인하였고, 앞으로 정보 검색에서 지능형 정보 처리가 요구된다는 점에서 볼 때, 상이한 정보 시스템간의 상호운용성을 확보하기 위해서는 기존의 지식조직시스템을 통합할 필요가 있다(Soergel et al, 2004). 따라서 시스템의 구조가 보다 정밀해 질 필요가 있으며, AGROVOC의 UF, NT, BT, RT 관계를 개선하려는 연구가 이루어져야 할 것이다.

4. 온톨로지의 개념간 관계

4. 1 온톨로지의 개념 모형

시소러스의 단점을 보완할 수 있는 온톨로지의 개념간 관계를 설명하기 전에 온톨로지의 개념간 관계의 바탕을 이루는 개념 모형에 대해서 서술하겠다. 이 모형은 시소러스의 의미 관계 구조를 분석해서 도출한 것으로 UMLS의 개념 구조에 많은 영향을 받았다. 그리고 시소러스보다 정밀한 의미 구조를 생성하는데 유용한 모형이 될 수 있을 것이다.

이 모형의 가장 큰 특징은 개념 수준, 용어 수준, 문자열 수준을 명백하게 구분했다는 점인데, Büster(2003)가 용어학 이론에서 제안한 모형과 매우 유사하다. Büster가 의미전달에 초점을 맞추어서 개념과 의미를 구분한 점에서는 차이를 보였으나, 문자열에서 용어 단계로 다시 개념 단계로 진행되는 과정은 일치한다.

그러나 기존의 시소러스는 개념에 관한 정보

와 용어에 관한 정보를 구분하지 않고 혼합해서 표현함으로써 용어간의 분명한 의미 구조 구축에 한계를 드러내고 있으며, <그림 1>의 모형에서는 개념과 용어를 명확하게 분리시키고 있음을 알 수 있다(Soergel et al, 2004).

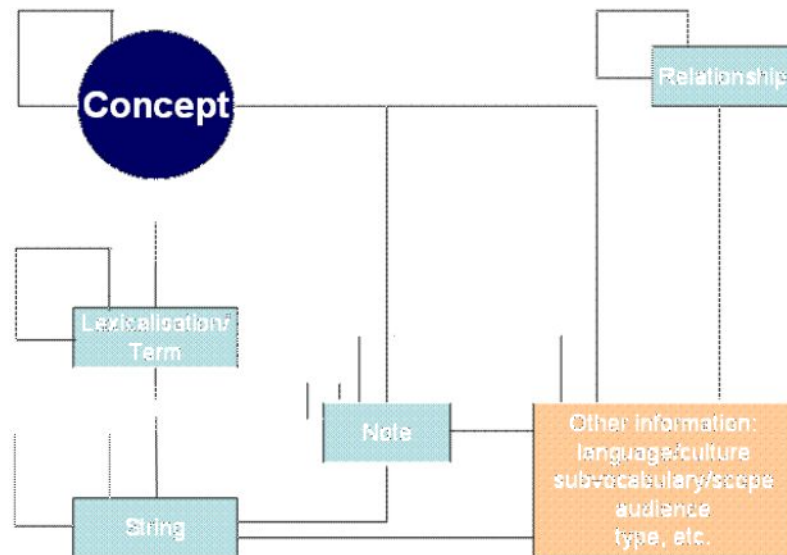
이 모형은 개념 구조의 틀을 제시하는 데 의미가 있다. 먼저 개념, 용어, 문자열에 대해 설명하면 다음과 같다. 개념은 용어나 어휘에 의해 표현되거나 지시되며, 이들은 단일어이거나 구 또는 복합어일 수 있으며, 용어는 단·복수, 격변화, 철자변형, 약어, 두문자어와 같은 다양한 어형 변화를 가질 수 있다. 따라서 하나의 개념은 다양한 어휘 표현을 가질 수 있으며 용어는 다양한 기호를 가질 수 있다.

각각의 개념, 용어, 기호는 URI를 갖는데, 개념의 경우에 있어서 UMLS는 개념 유일 식별자(Unique Concept Identifier: UCI)를

사용하는 반면에 토픽맵(Topic Map)에서는 공적 주제 식별자(Unique Subject Identifiers)를 사용한다. UCI를 사용하는 것은 개념에 대해서 분명한 참조를 가능하게 한다. 개념에는 DDC번호와 같은 기호들을 부여할 수 있고, 개념들을 계층적으로 제시함에 있어서 논리적이고 의미 있는 순서를 유지하기 위해서 기호법들을 이용할 수 있다.

그리고 개념들은 시소러스나 온톨로지 정보 모형에서 핵심 요소이다. 따라서 개념간의 관계가 중심이고, 개념들은 계층에 따라 배열되고 네트워크 내의 다른 개념들과 부가적인 관계를 맺는다. 특정한 경우에는 역관계를 가질 수도 있는데 예를 들어 <componentOf> 는 <has-Component>와 역관계이다.

더불어 중요한 점은 온톨로지에서의 개념간 관계 유형을 결정하기 위한 원천(source)은 시



<그림 3> 시소러스와 온톨로지의 통합을 위한 개념 모형

소러스 내에 존재하는 개념간 관계를 상세하게 분석하는 것이므로, 기존 시소러스의 개념간 관계 구조에 대한 지식이 중요하다. 각 개념은 과정, 기능, 재료, 살아있는 유기체처럼 하나의 엔티티 유형이나 패킷이어야 하는데, UMLS 의미망에서 제시한 의미 유형이 좋은 예가 될 수 있다.

개념이 용어로 표현됨으로써 개념 수준과 용어 수준이 연결되는데, 동형의이거나 동의어의 경우, 특정 용어가 복수의 개념을 갖는 경우에는 혼란을 피하기 위해서 개념들은 독립적으로 존재하며, 각 용어의 의미 속성이 분명한 용어들만 상응하는 개념과 연결된다.

이러한 이유로 개념간 관계를 구조화하는 시소러스나 온톨로지나 같은 지식조직시스템에서는 개념과 용어를 분리시켜서 모형화하는 것이 타당하다고 할 수 있다. 따라서 용어가 특정 개념을 표현하는데 사용되는 것처럼 시소러스에서는 디스크립터를 선택할 수 있으며, 해당 시소러스의 이용자가 선호하는 개념에 상응하는 디스크립터의 선택을 융통성 있게 결정할 수도 있다.

또한 <hasSynonym>, <hasAntonym>, <has-Cognate>(동일한 어원을 갖지만 언어가 다른 용어), <hasTranslation>과 같은 여러 관계를 통해서 용어들은 연결될 수 있다. 이미 남영준(2004)과 김태수(2000)가 유사동의어와 반의어를 시소러스의 의미 관계에 포함시킴으로써 그러한 바탕 위에서 이러한 온톨로지의 의미 관계를 설정할 수 있는 것이다.

그리고 용어는 문자열로 표현되며, 문자열은 <hasCaseVariant>, <hasSpellingVariant>, <hasAbbreviationOrAcronym>, <pluralOf/>

<singularOf>와 같은 관계를 통해서 연결될 수 있다. 시소러스는 그 용어를 표현하기 위해 사용된 문자열 중에서 선호하는 이형(variant)을 선택하면 된다. 특히 반의어에서 하나의 문자열은 여러 용어들에 동시에 포함될 수 있고, 그러한 경우에 분명한 의미 속성을 제시할 필요가 있다.

이처럼 개념, 용어, 문자열을 독립된 엔티티/패킷 유형으로 정의하는 이유는 두 가지이다. 첫째는 엔티티/패킷 유형 각각은 상이한 유형의 정보를 가지고 있기 때문이고, 둘째는 이러한 구별이 혼란을 피할 수 있다는 점이다. 특히 AGROVOC과 같은 표준 시소러스에서는 색인과 검색에서 사용되는 각각의 개념에 상응하는 우선어를 선택하고, 용어에 상응하는 우선 문자열이 선택되는데, 이 문자열이 디스크립터이다. 따라서 비디스크립터는 디스크립터와 동의 관계에 의해서 연결될 뿐이며 비디스크립터들 간에는 어떠한 관계나 연결도 설정되지 않아서 의미 구조상의 한계를 가질 수밖에 없다.

이러한 한계는 의미 관계를 구조화하는 기존의 시소러스가 갖는 약점 중의 하나라고 할 수 있으며, 기존의 시소러스의 이러한 약점에 대한 철저한 분석을 통해 보다 개선된 온톨로지의 구축이 가능해 질 수 있으며, 앞에서 제시한 시소러스와 온톨로지의 통합을 위한 개념 모형이 그 토대가 될 수 있을 것이다.

4.2 온톨로지의 개념간 관계 유형

여기에서 제시한 개념간 관계 유형은 AGROVOC 시소러스에 대한 언어학적, 개념적 분석에 기초한 결과이다. 이미 알려진 유형들도 있

으나, 이 관계들은 식량 및 농업 영역에서 필요한 관계라는 것을 인정해야하며, 일부는 새로운 의미를 추가했다.

그리고 관계 유형을 완벽하게 제시하려는 의도는 아니고, 보다 특정한 관계를 통해 전달될 수 있는 추가적인 정보와 개념 구조의 명확성을 설명하기 위한 예를 제시하려는 것뿐이다. 현재 사용되고 있는 시소러스의 관계들인 BT, NT, RT, UF에 대해서 AGROVOC을 개선함으로써 대체할 관계들로 통합하기 위해 비교 연구를 포함한 더 많은 연구가 필요할 것이다.

4. 2. 1 종속 관계

4. 2. 1. 1 X<includesSpecific>Y/Y<isa>X <isa> 관계는 'X<inheritsTo>Y/Y<inheritsFrom> X'의 관계도 포함하고, 대부분의 NT 관계를 변환한 것으로, 예는 아래와 같다.

- alcoholic beverages <includesSpecific> beers
- bovine spongiform encephalopathy(牛海綿狀腦病)<includesSpecific>spongiform encephalopathy(海綿狀腦病)
- cells<includesSpecific>blood cells
- blood cells<includesSpecific>leukocytes
- leukocytes(백혈구)<includesSpecific>lymphocytes(림프구)
- lymphocytes<includesSpecific>T-lymphocytes(UF 관계의 예)

위의 cells(세포)부터 T-lymphocytes(T-림프구)까지의 속종 관계를 계층 수준별로 나

타내면 다음과 같으며, 이것은 시소러스의 계층 수준보다 온톨로지의 층위가 더 세분되어 있음을 보여주는 예라고 할 수 있다.

- cells
- blood cells
- leukocytes
- lymphocytes
- T-lymphocytes

4. 2. 2 부분-전체 관계

다양한 관계가 여기에 속하는데, 대부분은 NT 관계이고, 계층적 상속의 방향이 때로는 부분에서 전체로, 때로는 전체에서 부분으로 나타난다. 여기에 제시된 모든 예는 전체에서 시작되는 관계들이다. 그리고 AGROVOC에서는 BT/NT 또는 RT 관계로 표현되어 있으며, 실제 예는 아래와 같다.

4. 2. 2. 1 X<containsSubstance>Y/Y<substanceContainedIn>X와 X<hasIngredient>Y/Y<ingredientOf>X

Y는 자연적으로(<containsSubstance>) 또는 인공적으로(<hasIngredient>) 만들어진 X의 물질이나 재료이다. 일단 X와 합쳐지면 Y는 자신의 정체성을 잃어버린다.

- blood<containsSubstance>blood fibrinogen(섬유소원)
- blood<containsSubstance>blood haemoglobin
- blood<containsSubstance>blood sugar

blood<containsSubstance>blood cells
cocoa beverages<hasIngredient>cocoa powder(RT 관계의 예)

4. 2. 2. 2 X<yieldsPortion>Y/Y<portionOf>X

X<yieldsPortion>Y 는 덩어리 X와 덩어리 X에서 취한 조각 Y와의 관계를 기술하는데, 예를 들면 아래와 같다.

roots<yieldsPortion>cuttings(조각)(RT 관계의 예)

4. 2. 2. 3 X<spatiallyIncludes>Y/Y<spatiallyIncludedIn>X

이 관계는 공간적인 범위를 갖는 대상을 위해서 사용된다. X는 Y에서 식별할 수는 있지만 분리할 수 없는 부분이다. 이들은 몸의 일부와 지리적 위치 등을 포함한다.

Asia<spatiallyIncludes>East Asia(NT 관계의 예)

특정한 공간과 관련된 관계는 전이성이라는 특징을 갖는 것이 일반적이는데, 'X<spatiallyIncludes>Y 이고 Y<spatiallyIncludes>Z 일 때 X<spatiallyIncludes>Z 이다' 라는 명제에 의해 개념간 관계의 전이성의 예를 제시하면 아래와 같다.

Asia<spatiallyIncludes>East Asia
East Asia<spatiallyIncludes>Korea Republic

Asia<spatiallyIncludes>Korea Republic

4. 2. 2. 4 X<hasComponent>Y/Y<componentOf>X

X가 Y의 일부분으로서 전체에 포함되더라도 X의 대상으로서의 정체성을 유지하는 관계를 말하며, X는 열거할 수 있거나 명명할 수 있어야 한다.

plough<hasComponent>ploughshare(RT 관계의 예)

woody plant<hasComponent>plant anatomy(식물의 부분들)(RT 관계의 예)

nucleus<hasComponent>chromosome(NT 관계의 예)

4. 2. 2. 5 X<includesSubprocess>Y/Y<subprocessOf>X

ATP cycle(腺苷三磷酸)<includesSubprocess>phosphorylation(磷酸化作用)(RT 관계의 예)

4. 2. 2. 6 X<hasMember>Y/Y<memberOf>X

집단, 그룹 또는 조직체 안의 구성원의 관계를 말한다.

Francophone Africa<hasMember>Benin(RT 관계의 예)

biotope(소생환경/군락생활지역)<hasMember>plant

pesticide crops<hasMember>Artemisia absinthium(서양쑥)(RT 관계의 예)

4. 2. 3 추가된 관계

여기에서는 Schmitz-Esser(1999)가 특정 주제 영역에 국한되지 않으면서, 동시에 다언어 시소러스의 설계 및 통합에 관한 연구에서 제안한 개념간 관계 유형의 일부를 함께 포함하여 제시하였다.

4. 2. 3. 1 X<causes>Y/Y<causedBy>X

X라는 원인에 의해 Y라는 결과를 낳게 되는 관계를 말하는데, 예는 아래와 같다.

overgrazing(과도한 방목)<causes>desertification(RT 관계의 예)

Serpulina hyodysenteriae(돼지이질小蛇菌)<causes>swine dysentery(돼지이질)
preharvest sprouting<causes>crop losses

4. 2. 3. 2 X<instrumentFor>Y/Y<performedByInstrument>X

이 관계는 X는 결과인 개념 Y를 성취하기 위한 수단이 된다. 예는 다음과 같다.

plough<instrumentFor>ploughing(RT 관계의 예)

상정한 도구는 생명체, 기계 또는 시스템이 사용할 수 있다. 이 관계의 의미는 도구를 사용함으로써 이루어지는 결과에 초점을 맞춘다.

4. 2. 3. 3 X<processFor>Y/Y<usesProcess>X

X가 Y안에 포함된 과정인 경우를 지시한다.

soil injection<processFor>fertilization A GROVOC에서의 RT

gonadectomy(생식선제거술)<processFor>sterilization[of organisms](BT 관계의 예)

4. 2. 3. 4 X<beneficialFor>Y/Y<beneficialFrom>X

fertilization<beneficialFor>crop yield(A GROVOC에는 없는 예)

4. 2. 3. 5 X<treatmentFor>Y/Y<treatedWith>X

pentosan(戊聚糖) polysulphate(PPS: 변종 인간광우병 치료를 위한 실험적 약물)
<treatmentFor>bovine spongiform encephalopathy(AGROVOC에는 없는 예)

4. 2. 3. 6 X<harmfulFor>Y/Y<harmedBy>X

AGROVOC에서는 간접적으로만 발견됨. 예를 들어:

preharvest sprouting<causes>crop losses로부터

preharvest sprouting<harmfulFor>crop yield 라는 결론을 내릴 수 있다.

4. 2. 3. 7 X<growsIn>Y / Y<growthEnvironmentFor>X

Halophytes(鹽生植物)<growsIn>saline soils(RT 관계의 예)

4. 2. 3. 8 X<hasProperty>Y/Y<propertyOf>X

fertilization<hasProperty>application rate(RT 관계의 예)

blood circulation<hasProperty>blood pressure(NT 관계의 예)

4. 2. 3. 9 X<similarTo>Y/Y<similarTo>X

bovine spongiform encephalopathy<similarTo>Creutzfeld-Jacob syndrome (RT 관계의 예)

4. 2. 3. 10 X<oppositeTo>Y/Y<oppositeTo>X

crop losses<oppositeTo>crop yield(AGROVOC에는 없는 예)

온톨로지의 개념 관계 유형은 상대적으로 시소러스의 개념간 관계보다 의미 구조의 표현에 있어서 보다 분명했고, 시소러스에서는 설정하지 않은 새로운 개념간 관계를 설정함으로써 개념간 관계가 보다 다양하게 나타났으며, 개념간 계층 수준에서도 층위가 세분된 사례들을 보임으로써 기존의 시소러스가 갖는 의미 관계의 한계를 부분적으로 극복한 것으로 평가할 수 있다. 그러나 개념간 관계 유형의 수가 많은 것이 오히려 단점이 될 수도 있다.

5. 온톨로지의 개념간 관계 유형의 적용

여기에서는 법원도서관에서 간행한 법률분

야 관련어집(Legal Thesaurus)의 용어간 관계를 온톨로지의 개념간 관계 유형을 사용하여 법률 용어간의 의미 관계를 표현하였다. 대상이 된 법률분야 관련어집의 용어들 중에서 사례가 있는 온톨로지의 개념간 관계들만 제시하였고, 여기에 제시된 예들은 속종관계를 제외하면 대부분 연관관계를 변환한 것이다. 그리고 관계에 대한 개념적 설명은 3장에 제시하였으므로 생략하였다.

5. 1 종속 관계

5. 1. 1 X <includesSpecific> Y/Y <isa> X
국가적공권<includesSpecific>형벌권(NT 관계의 예)

공무방해에관한죄<includesSpecific>공무집행방해죄(NT 관계의 예)

배당<includesSpecific>이익배당
위증제<includesSpecific>모해위증죄(NT 관계의 예)

5. 2 부분-전체 관계

5. 2. 1 X<hasMember>Y/Y<memberOf>X
대법원<hasComponent>대법원장
법원<hasComponent>법원공무원
헌법재판소<hasComponent>헌법재판소재판관

5. 3 새로운 관계

법률분야 관련어집의 RT관계로 설정된 관계 유형을 온톨로지의 새로운 관계유형으로 표

현하면 다음과 같다.

5. 3. 1 X<causes>Y/Y<causedBy>X

근무태만<causes>감급
 산업재해<causes>근로손실
 이혼<causes>가정붕괴
 주식소각<causes>자본감소

5. 3. 2 X<instrumentFor>Y/Y<performedByInstrument>X

국적취득<instrumentFor>귀화
 저항권<instrumentFor>시민불복중운동
 책임보험<instrumentFor>손해배상책임

5. 3. 3 X<processFor>Y/Y<usesProcess>X

구속절차<processFor>구속
 불신임절차<processFor>불신임
 심리절차<processFor>심리

5. 3. 4 X<beneficialFor>Y/Y<beneficialFrom>X

공정거래법<beneficialFor>독점규제
 리콜<beneficialFor>소비자보호
 치매관리사업<beneficialFor>노인복지
 컴퓨터프로그램보호법<beneficialFor>소프트웨어

5. 3. 5 X<harmfulFor>Y/Y<harmedBy>X

불법복제<harmfulFor>컴퓨터프로그램보호법
 불법인출<harmfulFor>홈뱅킹
 중간착취<harmfulFor>근로조건

환경범죄<harmfulFor>환경

5. 3. 6 X<hasProperty>Y/Y<propertyOf>X

공무원<hasProperty>겸직금지
 사회보장제도<hasProperty>각출
 이자<hasProperty>이율
 배당<hasProperty>배당액
 회사<hasProperty>회사명

5. 3. 7 X<similarTo>Y/Y<similarTo>X

이 관계는 유사동의어인데 남영준(2004)과 김태수(2000) 모두 대등관계로 파악하였다.

귀책원인<similarTo>귀책사유
 만성적실업<similarTo>구조적실업
 보수<similarTo>급여
 불법<similarTo>위법
 사회질서<similarTo>공공질서

5. 3. 8 X<oppositeTo>Y/Y<oppositeTo>X

남영준(2004)이 제시한 대립어, 반의어, 대비어 중 대립어 사례가 대부분을 차지하는 것으로 나타났다.

무한책임<oppositeTo>유한책임
 부동산<oppositeTo>동산
 불문법<oppositeTo>성문법
 피고<oppositeTo>원고

온톨로지의 관계 유형을 시소러스에 적용한 결과는 다음과 같다. 온톨로지의 관계 유형 중에서 속종관계인 <includesSpecific>은 시소러스의 NT 관계와 같은 관계 유형이고, <hasProperty>, <oppositeTo>, <similarTo> 등은

연관관계를 변환한 것이었으며, 이 세 관계 유형들은 시소러스의 개념간 의미 관계를 확장하거나 보완하는데 활용할 수 있을 것으로 보이며, 온톨로지에서 새로운 개념간 관계 유형으로도 이용될 수 있을 것이다. 그리고 나머지 관계 유형들은 사례 수가 적은 단점은 있으나 온톨로지에서의 활용가능성을 확인할 수 있었다.

더불어 개념간 관계 유형의 분석에서 복합어를 구성하는 용어간에 일정한 패턴을 가진다는 사실을 발견하였다. 예를 들면, 대부분의 경우 '배당'의 NT관계는 <includesSpecific> 관계와 RT관계는 <hasProperty> 관계와 일치하고, '배당' 이 다른 용어 뒤에 구성요소로 추가되면 대부분 <includesSpecific> 관계가 되고, '배당' 뒤에 다른 용어가 추가되면 <hasProperty> 관계와 일치하였다. 따라서 단일어와 복합어간의 관계 유형을 일정한 형식으로 정형화하거나 시소러스에서 난제 중의 하나인 복합어를 처리하는 방법론에 관한 후속 연구로 연결시킬 필요가 있다.

6. 결 론

시소러스의 개념간 관계를 구조화하는 수단인 동의 관계, 계층 관계, 연관 관계 각각이 의

미 관계를 구조화함에 있어서 완성도가 떨어지는 단점을 보완하는 대안으로 온톨로지를 고려할 수 있다. 그러나 시소러스의 개념간 관계 유형에 대한 철저한 분석을 통해서만 온톨로지의 개념간 관계 유형이 도출될 수 있고, 더 나아가서 시소러스에서는 실현하지 못했던 추론 규칙에 의한 새로운 개념간 관계 유형의 생성이나 정보를 찾아내는 연구도 가능할 것이다. 따라서 온톨로지에서는 시소러스가 가지고 있던 개념 모형과는 차별화되는 개념 모형을 고안해야 하며, 이 논문에서 제시한 개념, 용어, 문자열을 구분해서 개념을 모형화하는 방법을 보다 확장시키는 연구가 뒤따라야 할 것이다.

이와 관련하여 현재 우리 문헌정보학 영역에서는 시소러스 자체에 대한 심층적인 연구나 시소러스 구축에 관한 연구보다, 오히려 온톨로지에 대한 관심에만 집중되는 경향이 있는 것이 사실이다. 그러나 온톨로지가 표준화된 웹 자원 조직 수단으로서의 기능을 제대로 수행하려면, 시소러스에 관한 연구도 함께 병행되어야 한다. 시소러스가 의미 관계를 표현함에 있어서 한계를 가지고 있는 것은 사실이나 전통적인 문헌정보학의 주요한 연구 영역인 시소러스의 연구 성과들이 바탕이 될 때에만, 웹에서 자원을 조직하는 기능을 제대로 수행할 수 있는 온톨로지를 개발할 수 있을 것이기 때문이다.

참 고 문 헌

김태수. 2000. 『분류의 이해』. 서울: 문헌정보처리연구회.
김태수. 2001. 용어 정의를 도입한 시소러스 개발

연구. 『정보관리학회지』, 18(2): 231- 254.
남영준. 2004. 대등관계 설정의 확장 지침에 관한 연구. 『정보관리학회지』, 21(2): 1-21.

- 법원도서관. 1998. 『법률분야 관련어집』. 서울: 법원도서관.
- 손대형, 김태수. 1998. 패킷분류체계를 이용한 시소러스 작성에 관한 연구. 『제5회 한국정보관리학회 학술대회』. 1998년 8월.
- 정도현. 2003. 『시소러스를 기반으로 한 온톨로지 시스템 구현에 관한 연구』. 석사학위논문. 연세대학교대학원. 문헌정보학과.
- 조현양, 남영준. 2004. 시소러스와 온톨로지의 상호 호환성에 관한 연구. 『정보관리학회지』. 21(4): 27-47.
- 한상길. 1999. 시소러스의 기본 용어관계에 관한 연구. 『한국도서관정보학회지』. 30(4): 107-136.
- 한상길. 2000. 시소러스의 부가관계에 관한 연구. 『정보관리학회지』. 17(2): 119-138.
- AGROVOC. 2004. [online].
<<http://www.fao.org/agrovoc/>>.
- ERIC. 2004. [online].
<<http://www.ericfacility.net/extra/pub/thesbrowse.cfm>>.
- Hodge, Gail M., Marcia Lei Zeng and Dagobert Soergel. 2003. *Building a meaningful web: from traditional knowledge organization systems to new semantic tools*. Proceedings of the third ACM/IEEE-CS Joint Conference on Digital Libraries. [online]. in Houston, TX on May 31 2003. [cited 2004. 7. 20].
<<http://portal.acm.org/citation.cfm?id=827140.827243&dl=GUIDE&dl=ACM&type=series&idx=827140&part=Proceedings&WantType=Proceedings&title=International%20Conference%20on%20Digital%20Libraries>>.
- Schmitz-Esser, W. 1999. "Thesaurus and beyond: an advanced formula for linguistic engineering and information retrieval." *Knowledge Organization*, 26(1): 10-22.
- Soergel, Dagobert, Boris Lauser, Anita Liang, Frehiwot Fisseha, Johannes Keizer and Stephen Katz. 2004. "Reengineering thesaurus for new applications: the AGROVOC example," *Journal of Digital Information* [online]. 4(4): [cited 2004.7.15].
<<http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Soergel>>.
- UMLS. 2004. [online].
<<http://www.nlm.nih.gov/research/umls>>.
- Wüster, Eugen. 2003. "The wording of the world presented graphically and terminologically." selected and translated by Juan C. Sager. *Terminology*, 9(2): 269-297.