

시맨틱웹을 위한 온톨로지 구축방법에 관한 비교 연구

The Comparative Study on the Methodologies of Building Ontology toward Semantic Web

김은경* · 남영준**

Eun-Kyong Kim · Young-Joon Nam

차 례

- | | |
|------------------------------------|--------------------|
| 1. 서론 | 4. 온톨로지 구축방법 모델 제안 |
| 2. 이론적 배경 | 5. 결론 |
| 3. IEEE 표준에 기반한 온톨로지 구축
방법 비교분석 | · 참고문헌 |

초 록

본 연구는 IEEE Standard 1074-1997을 기준으로 사례별 온톨로지 구축방법의 절차를 분석하였다. 분석대상은 OTK, CommonKADS, ONIONS, Ontology Development 101의 네 가지 구축방법을 선정하였다. 분석된 결과를 기반으로, 각 온톨로지 구축방법의 장점을 수용하고 단점을 보완함으로써 여섯 단계로 이루어진 완성도 있는 온톨로지 구축방법 모델을 제안하였다. 그 단계는 다음과 같다. 1) 사전개발 단계로서 타당성 조사를 포함한다. 2) 초기 구축 단계로서 온톨로지 구축 목표를 성립한다. 3) 지식 재사용을 위해 기존 온톨로지를 통합하는 방안을 고려한다. 4) 주제 영역의 개념과 관계를 규정하면서 온톨로지를 구성한다. 5) 완성된 온톨로지를 평가하고 검증한다. 6) 사후개발 단계로서 유지관리 절차를 포함한다.

키 워 드

온톨로지, 시맨틱웹, 온톨로지 구축방법, 온톨로지 정보검색

- * 유네스코 아시아·태평양 국제이해교육원 사서
(Librarian, Asia-Pacific Centre of Education for International Understanding)
- ** 중앙대학교 문과대학 문헌정보학과 부교수
(Professor, Library & Information Science Dept., Chung-Ang Univ., namyj@cau.ac.kr)
- 논문접수일자: 2004년 6월 1일
- 게재확정일자: 2004년 6월 18일

ABSTRACT

This is the comparative study about analyzing the methodologies of building ontology with IEEE Standard 1074-1997. The methodologies are chosen to be analyzed. They are OTK, CommanKADS, ONIONS and Noy & McGuinness's Ontology Development 101. On the basis of analyzing, the ontology development process is proposed after the strong points of four methodologies are accepted but the weak points of them are supplemented. The sixth development steps are following: 1) Conducting the feasibility study about ontology building as the pre-development process 2) Setting up the purpose of the ontology development as the starting point of the building 3) Considering the integration of the existing ontologies for the knowledge reuse 4) Constructing the ontology by defining the concepts and relations 5) Evaluating and testing the ontology for the completeness 6) Containing the ontology maintenance for the sustainable use.

KEYWORDS

ontology, semantic web, building ontology, ontology's IR

1. 서론

웹의 등장은 전통적인 정보검색을 비롯하여 지식관리와 일반 상거래 등 사회 전 분야의 변혁을 초래하였다. 특히 웹 정보검색은 소장 자료를 대상으로 하는 제한된 검색에서 웹을 통해 접근할 수 있는 전자자원을 대상으로 하는 무한 검색을 가능하게 하였다. 이러한 검색대상 범위의 확대는 보다 정교한 검색을 필요로 하게 되었으며, 필연적으로 지능화된 정보검색 시스템 개발을 촉진하는 계기가 되었다.

온톨로지(Ontology)는 시맨틱웹을 구현할 수 있는 도구로써 지식개념을 의미적으로 연결 조직화하는 기술로 폭증하는 웹자원을 효과적

으로 관리할 수 있는 정보검색의 새로운 도구이다. 또한, 지식분야와 영역의 특성을 수용할 수 있도록 가변적으로 구축하는 도구라 할 수 있다. 즉 온톨로지는 현실 세계에서 어떤 특정 영역에 대한 개념을 가상공간에서 구조화할 수 있도록 정의한다. 이를 통해 가상공간에서 용어의 의미나 개념간의 관계를 인간이나 컴퓨터 모두 인식할 수 있도록 하는 새로운 기술이다. 따라서 기존의 웹은 온톨로지의 구축을 통해 시맨틱웹으로 발전해 나가게 된다. 그러므로 다양한 분야의 온톨로지 구축은 시맨틱웹의 완성도를 높이는 계기가 될 것이다.

이를 위해서는 개발 영역과 관련한 기술 지식을 정확히 파악하고 적합한 구축방법을 이용

하여 해당 분야의 개념을 구조화해야 한다. 이는 온톨로지 구축에 있어 구축대상의 주제영역과 이용대상 등에 따라 최적의 구축방법을 선정해야 함을 의미한다. 구축방법 선정은 온톨로지의 완성 결과에 많은 영향을 미치기 때문에 온톨로지 구축 시 우선적으로 어떤 방법을 선정할 것인지에 대해 정확히 파악해야 한다.

그러나 기존의 구축된 온톨로지는 개발 환경에 따라 구축방법이 다르기 때문에 모든 영역을 포함할 수 있는 표준화된 온톨로지 구축 모델이 없는 실정이다. 지금까지의 온톨로지 구축은 해당 영역에 가장 적절한 방법을 해당 주제의 특성에 따라 개별적으로 개발하기 때문에 이러한 방법은 실제적인 온톨로지 구축에 커다란 장애가 되고 있다.

본 연구는 이 점에 착안하여 기존의 주요 온톨로지 구축방법들을 분석하여 온톨로지 구축에 필요한 공통적인 요소와 절차를 파악하고, 주요 주제 분야에서 필요한 요소들을 추출하고자 한다. 이런 분석결과를 토대로 본 연구는 주요 온톨로지 개발 분야에 많은 요소를 수용할 수 있는 표준화된 구축방법 모델을 제시하는데 그 목적이 있다.

본 연구의 목적을 수행하기 위해 다음과 같은 연구방식으로 연구를 수행하였다.

첫째, 문헌조사 방법을 통해 온톨로지의 개념과 배경 및 구축방법의 내용을 고찰하였다.

둘째, 온톨로지는 하나의 소프트웨어라 할 수 있기 때문에 온톨로지 구축방법을 비교하기 위한 기준으로 IEEE Standard 1074 1997을

채택하였다. 이를 기준으로 하여 온톨로지 구축방법 중 OTK, CommonKADS, ONIONS, Ontology Development 101의 네 가지 구축방법 사례를 분석대상으로 삼아 그 절차와 내용을 파악하였다.

셋째, IEEE Standard 1074 1997의 개발 절차에 따라 사례별 온톨로지 구축방법을 분석하고 사례별 온톨로지 구축방법의 특징과 장단점을 파악하였다.

넷째, 사례별 온톨로지 구축방법의 비교 분석을 바탕으로 온톨로지 구축방법의 표준적인 모델을 제시하였다.

본 연구에서 비교분석을 위해 선정한 온톨로지 구축방법의 수는 제한적이다. 국내외에서 구축하고 있는 모든 분야의 온톨로지 구축방법을 포괄적으로 수용하고 있지 못하는 제한점을 갖고 있다. 제안한 온톨로지 구축방법 모델은 선정된 온톨로지 구축방법 사례를 비교분석한 결과만을 토대로 한다. 그러므로 본 연구의 온톨로지 구축방법 모델은 상대적인 대안이 될 것이다.

2. 이론적 배경

2.1 시맨틱웹과 온톨로지

시맨틱웹은 현재 우리가 사용하고 있는 웹의 확장된 개념이다. 가장 큰 차이점은 현재의 웹은 단순한 키워드나 디렉토리 방식으로 일반적으로 검색이 이루어지나 시맨틱웹은 의미 검

색방식으로 검색이 이루어지는 점이다. 이것은 인간뿐만 아니라 컴퓨터가 의미를 인식할 수 있는 기술수준이 가능하게 되었기 때문이다. 따라서 웹을 통해 완전한 시맨틱웹이 구축되는 시점의 정보검색은 단순 키워드 검색으로 인한 다량의 불필요한 정보까지 연계 되는 경험을 줄일 수 있어 검색결과의 만족도를 높일 수 있다. 이런 시맨틱웹으로의 발전을 이루기 위해서는 지식 분야에 대한 온톨로지가 반드시 구축되어야 한다. 시맨틱웹이 기존 웹의 확장이라고 하는 것은 이제까지 개발된 웹 기술을 바탕으로 의미검색이 가능하도록 온톨로지가 사전에 구축되기 때문이다. 즉, 시맨틱웹을 실제로 구현하는 기반기술이 온톨로지이다. 특히 전산학적인 측면에서는 온톨로지에 대해 다음과 같이 정리하고 있다(최호섭, 옥철영 2004).

- 일정한 체계에 의한 어휘사전이나 용어의 확보
- 특정영역 뿐만 아니라 보편영역의 기본 개념에 대한 정의와 그들 간의 관계에 대한 명세화
- 개념, 관계, 속성들의 유기적인 집합
- 전산적인 처리가 가능한 구조화와 구체화
- 공유와 재사용의 가능
- 논리적 추론
- 통합

이상과 같은 관점은 문헌정보학 분야의 주제명 표목을 비롯하여 시소러스와 같은 색인어휘사전의 구조화와 동일한 구축과정을 공유할 수 있음을 의미한다.

웹의 계층 구조는 자원 표기를 위한 URI를 비롯하여 다국어 지원코드를 기초로 데이터계층, 스키마 계층, 의미계층으로 구성되며, 이 가운데 의미계층에 해당하는 것이 온톨로지이다. 그 밖에 상위 계층으로 Rules, Logic framework, Proof, Trust가 있으나 이 분야의 연구는 아직 미비한 실정이다. 컴퓨터가 인식할 수 있는 의미를 부여하기 위한 노력은 XML이나 RDF(S) 기술을 통해 발전되어 왔다. XML의 경우 태그에 의미정보를 부가하여 의미검색이 가능하며, RDF는 구조화된 메타 데이터의 생성 및 교환, 재사용을 통해 의미검색을 시도한다. RDF 스키마는 RDF 기술을 위한 스키마 정의언어이다. 그러나 이런 XML과 RDF(S) 기술만으로는 구조적이고 충분한 의미를 전달하는 시맨틱웹을 실현하기 어렵다. XML은 시맨틱웹을 구성하는 기반구조를 제공하고 있으나, 개념정의를 부족하다. RDF(S)은 개념과 속성, 관계를 부분적으로 정의하고 있으나 완전한 개념관계를 표현하는데 부족하다. 그러므로 데이터나 스키마를 통해 의미를 부여하는 차원을 넘어서 온톨로지라는 상위 의미계층을 개발할 필요가 있다.

시맨틱웹의 구조가 기존의 웹과 또 다른 점은 의미계층에 해당하는 온톨로지 지식 기반 시스템과 온톨로지에 포함된 개념의 의미를 유추할 수 있는 추론 엔진 시스템이 필요하다는 것이다. 온톨로지 추론 엔진(Ontology Inference Engine)은 온톨로지에서 추론 규칙을 적용하여 개념을 유추하는 기능을 실행한

다. 왜냐하면 온톨로지는 추론 엔진을 통해서 질의를 처리하기 때문이다. 추론 엔진은 질의어의 관계를 파악하여 관련 용어를 검색 엔진에 전달하게 된다. 시맨틱웹은 현재 웹에서 온톨로지 의미체층과 그에 따르는 추론엔진 시스템이 부가된 확장 구조를 하고 있다.

온톨로지는 시맨틱웹 구현의 핵심 기술이다. 즉, 지식을 컴퓨터가 처리할 수 있는 상징체제로 만들고 인간이 그 체계를 이해할 수 있도록 구현한 커뮤니케이션 기술이다. 이런 온톨로지의 기원은 철학에서 그 어원이 발생하였다. 철학에서 말하는 온톨로지는 존재의 특성을 연구하는 이론을 의미한다. 철학에서 비롯된 온톨로지를 인공지능 분야에서는 용어사이 관계를 공식적으로 정의한 문서나 파일을 뜻하는 전문용어로 사용하기 시작하였다. 인공지능 분야에서는 온톨로지를 용어 확장을 정의하기 위한 관계 및 단어를 결합하기 위한 규칙으로 간주하고 있다. 또한 주제 용어를 포함하는 기본 용어와 관계를 정의하고 있다. 이 분야에서 온톨로지를 개발하기 시작한 것은 지식의 공유와 재사용을 목적으로 한 것이었다. 온톨로지는 1990년대 초에 지식공학, 자연언어 처리, 지식표현 분야에서 연구 과제로 부각되기 시작하였다. 현재에는 정보통합, 지식시스템, 정보검색, 전자상거래, 지식관리에서 온톨로지의 사용이 확산되고 있다. 이렇게 온톨로지가 널리 사용되는 이유는 인간과 어플리케이션 시스템간의 커뮤니케이션이 가능한 영역을 공유하여 이해할 수 있게 하기 때문이다(Ding et al.

2002).

온톨로지에 대한 여러 정의 가운데 '공유된 개념화에 대한 명시적이고 공식적인 상세'가 일반적으로 널리 알려져 있다. 여기서 공유되었다고 함은 온톨로지가 합의된 지식으로 공통된 이해를 바탕으로 하는 것을 의미한다. 개념화라는 것은 실제 세상에서 발생하는 현상에 대한 구체적 모델을 말하는 것이다. 명시적이란 것은 사용된 개념의 형태를 말하는데 그 사용에 대한 제한사항이 명시적으로 정의되는 것을 의미한다. 공식적이라고 함은 온톨로지가 기계 가독형, 즉 컴퓨터가 인식할 수 있어야 한다는 것이다(이재호 2003; 양정진 2003). 온톨로지는 지식 기반의 기초 분야가 될 수 있는 영역을 계층적으로 구조화된 용어의 집합으로 표현될 수 있다. 이 경우에 온톨로지는 지식 기반에서 나타나는 지식의 명시적 개념을 설명하는 수단이 될 수 있다. 간략하게는 클래스와 속성에 대한 정보의 집합이나 온톨로지 문서 내에 포함된 정보를 온톨로지로 표현한다(W3C 2004).

이런 온톨로지는 해당 지식 영역에서 일반적으로 공유되는 기본 용어를 제공한다. 즉, 주제 영역에 있어 통용되는 개념을 공유하게 하는 것이 온톨로지의 기본적인 역할이다. 더불어 온톨로지는 지식을 체계화시킨다. 이것은 인간이 설명하는 현상과 이론들을 개념, 용어로 잘 구성하는 것을 말한다. 여기서 온톨로지는 지식 체계화의 주요 바탕이 된다. 또한 온톨로지는 함축적인 지식을 공유할 수 있게 한다.

모든 인간 활동은 함축적인 가정을 내포하고 있다. 함축적인 가정 속에서 일반적인 기본 단어, 관계 및 제한사항에 대한 정의를 추출하여 예시를 만들고 이것을 온톨로지에 포함시킨다. 이것은 지식을 공식화하는 작업이다. 지식기반은 개발자가 갖고 있는 개념을 기반으로 만들어진다. 그러나 인간이 갖고 있는 개념은 함축적일 수밖에 없다. 인간이 가진 개념이 함축적이기 때문에 동일한 개념도 사람에 따라 다른 용어로 사용할 수 있다. 그러나 인간의 개념을 동일한 용어로 표현해야 공통적인 지식을 공유할 수 있다. 이렇게 인간 개념의 함축성은 지식의 공유를 어렵게 하는 요인이 된다. 그러나 온톨로지를 통해 이런 개념과 가정을 명시적으로 표현함으로써 지식의 공유와 재사용이 가능하다. 인간 개념의 함축성을 설명할 수 있다는 점은 온톨로지 구축의 중요한 이유 가운데 하나이다.

2.2 온톨로지 구축과정

온톨로지 구축은 앞의 분석에서 제시한 바와 같이 지능적인 웹정보검색을 위해 색인화에서 필수적인 과정이 되고 있다. 이러한 변화에 따라 온톨로지 구축방법은 이론과 이에 따른 실제 구현사례로 제한적인 분야에서 특정하게 이루어지고 있다. 포괄적인 온톨로지 구축방법은 목적 확인, 개념화, 기호화, 기존 온톨로지 통합, 평가, 문서화와 같은 기본 과정으로 이루어져 있다.

① 목적 확인은 온톨로지를 구축하는 목적을 분명히 함으로써 어느 분야나 시스템에서 온톨로지 사용에 대한 도입여부를 결정하는 과정이다. 온톨로지의 목적은 지식의 재사용인 경우가 많은데 이것은 언급한 것처럼 온톨로지 구축이유에 이미 내재되어 있다. 온톨로지의 목적은 소규모 단위에서 공유될 것인지, 다양한 어플리케이션에서 재사용될 것인지, 대규모 단위에서 재사용될 것인지에 따라 달라질 수 있다. 이 단계에서는 구체적으로 향후 이용대상이 누구이고 그들의 특성이 무엇인지 파악해야 한다. 적합질문 목록을 작성하고 개발환경이 어떠한지도 확인하는 것이 필요하다. 이 단계에서 작성한 적합질문은 추후 온톨로지를 평가하는 요소가 될 수 있다.

② 개념화는 어떤 주제 영역에서 핵심적인 개념이 무엇이며 개념간의 관계가 어떻게 설정될 수 있는지를 확인하는 과정이다. 이 과정에서는 각각의 개념과 관계를 문맥에서 정확하게 산출하고 관련 단어가 어디에 해당하는지 반드시 확인해야 한다. 즉 이 단계에서는 개념과 용어를 정의하고 확인하는 작업이 이루어진다. 개념화는 범주화, 정의 생성, 검토, 메타 온톨로지 작성으로 세분된다.

③ 범주화 작업은 주제 영역에 관련된 단어들을 최대한 나열하기 위해 브레인스토밍(자유 주제발언방식)을 통해 모든 의견을 취합한다. 주제 영역에서 취합된 단어들은 대략적으로 분류할 수 있는데 중복개념 등은 배제

하고 유사어 등은 통합한다. 이 단계에서 단어들 사이의 의미적 상호 참조도 확인하게 된다.

- ④ 정의 생성은 온톨로지 구축의 주요 작업이다. 이 단계의 정의 생성에 오류가 많으면 재작업을 해야 하기 때문에 매우 중요한 구축 단계가 된다. 해당 주제 영역에서 가장 기본이 되는 단어들을 정의한다.
- ⑤ 개념화 과정에서 정의를 생성한 후에는 정의 및 중요 결정사항 등을 다시 한 번 점검, 재검토한다. 그리고 기호화 단계의 기초가 되는 메타 온톨로지를 구축한다. 메타 온톨로지는 자연언어 정의를 사용하는 기초 온톨로지이다.
- ⑥ 기호화 단계에서는 자연언어로 작성된 메타 온톨로지를 가지고 온톨로지 표현언어로 기호를 생성하게 된다. 온톨로지 표현언어는 현재 OWL(Web Ontology Language)을 표준으로 권고하고 있다. 기존 온톨로지 통합은 개념화와 기호화 과정에서 기존에 생성된 온톨로지를 사용할 것인지를 결정하는 과정이다. 그러나 온톨로지 통합을 위한 지침과 도구를 제공하는 일은 포괄적인 온톨로지 구축 방법 개발에 있어 가장 어려운 과정이다.
- ⑦ 평가는 온톨로지의 기술적인 측면을 판별해 보고 관련 소프트웨어 환경을 살펴보는 과정이다. 또한 요구명세서와 적합질문 목록과 같은 기준으로 온톨로지를 평가할 수 있다.
- ⑧ 문서화는 온톨로지의 목적에 따라 다른 지침을 갖도록 하는 것이 바람직하다. 따라서

적합한 온톨로지 문서를 통해 효율적인 지식의 공유가 가능하기 때문에 문서화 작업이 온톨로지 구축방법 과정에 항상 포함된다.

3. IEEE 표준에 기반한 온톨로지 구축방법 비교분석

3.1 IEEE Standard 1074-1997

IEEE Standard 1074 1997은 기존의 IEEE/EIA 12207.0 1996과 IEEE/EIA J STD 016 1995를 통합 조절한 것이다. 주된 내용은 소프트웨어의 생성주기 개발에 따른 기준요소를 제시한 개발 표준이다. 즉, 특정한 소프트웨어를 개발할 때 개발의 절차를 설계하는 작업 표준기준을 제시하고 있다. 따라서 이 표준은 소프트웨어 개발의 절차를 설계하는 작업에 따른 국제적 기준을 갖고 있기 때문에 소프트웨어 개발 방법론에 대한 표준이라 할 수 있다(IEEE 1997). 따라서 사례별 온톨로지 구축 방법을 비교하기 위한 기준으로 IEEE Standard 1074 1997을 적용한다는 것은 온톨로지가 일종의 소프트웨어라는 전제를 가져야 한다. 소프트웨어란 사용자가 계산기를 유효하게 잘 다룰 수 있도록 제조자로부터 제공되는 여러 가지 프로그래밍 원수순단으로 정의된다(모기아키라 2000). 온톨로지는 어떤 영역의 개념과 관계를 조직화한 분류체계 시스템을 말한다. 소프트웨어라는 것도 컴퓨터 프로그램, 절차, 시스템 운영 관련 데이터 및 문서라

고 할 수 있다. 따라서 온톨로지도 부분적 또는 잠재적으로 소프트웨어라고 할 수 있다.

소프트웨어 개발 방법 표준인 IEEE Standard 1074 1997을 온톨로지 구축방법 비교 기준으로 적용하고자 하는 것은 온톨로지 구축과 소프트웨어 개발절차가 유사한 측면을 가지고 있기 때문이다. 소프트웨어 공학은 소프트웨어의 개발, 가동, 보수, 회수에 관한 체계적인 접근을 의미한다(전기전자공학사전편찬위원회 1996). 온톨로지 공학은 언급한 것처럼 온톨로지의 개념화, 디자인, 실행 및 배치에 이르는 모든 과정을 포함하는 활동을 말한다. 온톨로지는 잠재적 또는 부분적으로 소프트웨어라고 할 수 있다. 온톨로지 구축 과정에서의 확장성, 인식성, 유지관리의 편리성, 인터페이스 고려 등은 소프트웨어 요소의 상호 운용성이나 객체지향 디자인의 객체 클래스 구성에도 적용되는 공통적인 사항이다. 따라서 온톨로지 공학이 소프트웨어 공학과 유사성을 갖는다고 할 수 있다.

이 표준에서 소프트웨어 생성주기 절차는 65개의 활동으로 구성되며 이것은 17개의 활동 그룹으로 구분된다. 활동들은 전체적인 소프트웨어 개발 생성주기를 포함하고 있다. 이 표준에서 말하는 활동이란 실행될 작업 전체를 정의한 것을 말한다. 실행 작업에는 필요한 입출력 정보가 포함된다. 활동이란 입력 정보에서 출력정보로의 필요한 전환을 설명하는 것이다. 모든 입력 정보가 처리되고 모든 출력 정보가 산출되면 활동 실행은 완료된다. 그러므로 활

동은 입력 정보, 설명, 출력 정보의 세 가지 부분으로 나누어진다. 입력 정보는 전환될 필요가 있는 정보의 목록이다. 설명은 전환될 입력 정보의 대한 논의이다. 출력 정보는 입력정보가 전환되어 산출될 필요 정보의 목록이다.

3.2 온톨로지 구축방법 및 비교분석

온톨로지 구축방법은 이론적, 경험적으로 다양하게 제안되어 왔다. 온톨로지 구축방법은 SENSUS, KACTUS, OTK, CommonKADS, Tove, Methontology, Mikrokosmos, ONIONS, PHYSSYS 등과 같은 여러 종류가 있다. 본 연구에서는 여러 온톨로지 구축방법 중 일반적인 온톨로지 구축에 활용될 수 있을 것으로 판단되는 OTK, CommonKADS, ONIONS, Ontology Development 101의 네 가지 구축방법 사례를 선택하여 그 절차와 내용을 파악한다. 이 사례별 온톨로지 구축방법은 IEEE Standard 1074 1997로 비교분석 된다.

3.2.1 OTK(On-To-Knowledge)

1) 구축방법

OTKM(On To Knowledge Methodology for development and evolution of Ontology based Tools for Knowledge Management)로 명명된 이 온톨로지 구축방법은 기존의 다른 방법론과는 달리 어플리케이션 지향의 온톨로지 개발에 중점을 두고 있다. On To Knowledge는 온톨로지를

통해 지식을 관리할 수 있는 방법과 도구를 개발하는 전체적인 시스템이다(Davis 2003).

- ① 타당성 조사 단계는 개발될 온톨로지 시스템의 문제점과 가능성이 무엇인지, 잠재적인 솔루션은 어떤 것이 있을지를 확인하는 과정이다. 따라서 관련 작업의 특징을 구체적으로 확인하여 어떤 도구를 선택할지를 고려한다. 또한 지식 영역에서 개선점과 문제점을 구체화한다. 포함될 인력구성을 파악하는 것도 타당성 조사의 주요 내용이다.
- ② 시작 단계에서는 온톨로지 요구명세서(Ontology Requirements Specification Document:ORSD)의 작성이 이루어진다. 요구명세서는 어떤 개념을 포함하고 배제할 것인지, 온톨로지의 계층적인 구조를 어떻게 정할 것인지에 대한 지침을 제공하는 중요한 역할을 수행한다.
- ③ 정제(Refinement) 단계에서는 시작 단계에

서 작성한 온톨로지 요구명세서에 따라 어플리케이션 지향 온톨로지를 구축하는 단계로써, 지식을 공식화한다. 이를 위해 지식 도출 과정을 실행하여 온톨로지 초안을 수정, 확장해야 한다. 표현력을 고려하여 적합한 온톨로지 언어를 선정하고 온톨로지를 공식화한다. 이 단계는 평가단계와 연결되어 있다. 평가 단계에서 온톨로지를 분석하였는데 오류가 발생하면, 그 결과를 정제 단계에서 입력하여 재수정하는 과정을 반복적으로 실행해야 한다. 따라서 이 과정을 수행하는 기간에는 몇 차례의 반복적인 과정이 주기적으로 이루어진다.

- ④ 평가 단계에서는 요구명세서와 적합 질문에 온톨로지가 부합되는지 확인해야 한다. 시험 이용자들의 피드백을 모아서 온톨로지서 사용되는 부분과 사용되지 않는 부분을 추적해야 한다. 신중한 패턴 평가를 통해 자주 사

〈표 1〉 OTK 온톨로지 구축방법

단계	작업
타당성 조사	- 문제점 및 가능성 확인 - 지식관리 어플리케이션 초점 - 도구 및 포함 대상 확인
시작 단계	- 요구 명세서(ORSD) 작성 - 지식정보원 분석 - 초기 온톨로지 개발
정제 단계	- 초기 온톨로지 검증 및 정제 - 공식 온톨로지 생성
평가 단계	- 질의 분석
유지 및 개선 단계	- 지속적인 유지, 개선 과정

용되는 부분을 확장하고, 덜 사용되는 부분은 어플리케이션과의 상관성을 검토한다. 이 피드백 과정은 정제 단계로 넘어가 반복적인 수정 작업이 이루어진다.

- ⑤ 유지 및 개선 단계는 조직적인 과정으로 갱신, 삽입, 삭제에 대한 엄격한 규칙에 근거하여 통제된다. 개선사항이 발생하면 정제단계로 넘어가서 처리된다. 단, 어플리케이션에 대한 모든 효과를 철저히 시험한 후 변화 및 개선사항을 수집하고 새로운 버전에 대한 배치전환을 실행해야 한다.

2) 비교분석

OTK 온톨로지 구축방법은 타당성 조사를 비롯하여 시작, 정제, 평가, 유지 및 개선의 다섯 단계로 구성되어 있었다.

첫째, 타당성 조사 단계는 IEEE Standard 1074 1997의 타당성 조사 실행, 위험관리, 기능 분석 활동과 동일한 절차이다. OTK 온톨로지 구축방법에 있어서 타당성 조사는 무엇보다 시스템의 문제점과 가능성을 파악한다. 사용자와 이용 경우에 대한 분석, 작업분석, 지식항목 분석 등이 타당성 조사를 통해 이루어진다. 이런 조사 작업들은 IEEE Standard 1074 1997에서 의견이나 요구사항을 분석하고 소프트웨어를 개발하는데 제한사항과 이점이 무엇인지를 파악하는 타당성 조사 실행 활동과 동일한 절차이다. 시스템의 잠재적인 문제점을 파악하는 것은 위험관리 활동 절차를 수행하는 것이다. 작업분석의 경우는 소프트웨어 기능분석 활동과 동일한 절차이다.

둘째, 시작 단계는 IEEE Standard 1074 1997의 의견 요구 확인 활동, 의견 요구의 확정 및 정제, 소프트웨어 입수 방법 정의 활동과 동일한 절차이다. 이 단계에서는 온톨로지 요구 명세서를 통해 시스템에 대한 의견과 요구를 체계화한다. 이것은 실제 구축에 중요한 사전 작업이 된다. 요구명세서를 기술하는 절차는 의견 요구를 확인하고 확정하는 활동과 동일하다. IEEE Standard 1074 1997에서 소프트웨어 입수를 온톨로지 구축의 정보자원 입수로 대체한다고 가정하면, 시작 단계의 지식 정보를 입수 방식 설명 부분은 소프트웨어 입수 방법 정의 활동과 유사 절차로 판단할 수 있다.

셋째, 정제 단계는 IEEE Standard 1074 1997의 디자인 실행을 비롯하여 세부 디자인 실행, 실행 코드 생성 활동과 동일한 절차이다. 이 단계에서는 초기 온톨로지를 생성하기 위해 지식을 공식화한다. 이것은 디자인 실행 활동과 동일하다. 그리고 평가 단계에서 발생한 오류 내용을 반복적으로 수정, 정제하는 과정들이 단계에서 거치게 되므로 세부 디자인 실행 활동과도 동일한 절차로 판단할 수 있다. 또한 온톨로지 표현언어로 코드를 생성하는 작업은 소프트웨어 개발에서 프로그래밍 코드 생성 활동과 동일한 절차로 파악할 수 있다.

넷째, 평가 단계는 IEEE Standard 1074 1997의 검토실행을 비롯하여 추적 매트릭스 생성, 시험 실행, 평가 결과 보고 활동과 동일한 절차이다. IEEE Standard 1074 1997에서 검토실행과 추적 매트릭스 생성은 전체 생성주기

에 걸쳐 이루어지는 작업으로 진행 검토, 관리 검토, 진행 향상 검토, 사후 실행 검토 단계로 구성된다. 여기서 기술적인 결점을 보완하고 실제 결과와 계획을 비교하는 작업을 거치게 된다. 추적 매트릭스 생성은 기술적인 요구사항과 소프트웨어 요구사항이 무엇인지를 파악할 수 있도록 작성되는 것으로 검토작업 비교를 용이하게 한다. 이 작업은 온톨로지 요구명세서와 적합질문에 따라 구성된 온톨로지가 적합한지를 판단하는 평가 작업과 동일한 절차가 된다. 시험 실행은 시험 데이터를 전체 시스템에 사용해 보는 작업으로 반복적으로 행해질 수 있다. 이 점은 구축된 온톨로지를 시험 이용자들이 사용하여 피드백을 모으는 과정과 동일한 절차이다.

다섯째, 유지 및 개선 단계는 IEEE Standard 1074 1997의 소프트웨어 개선 사항 확인, 문제보고 방법 시행, 소프트웨어 생성주기 재적용 활동과 동일한 절차이다. 이 단계는 소프트웨어에서 개선되어야 할 사항들을 확인하고 문제점들을 보고서로 작성하고, 소프트웨어 생성주기를 반복하여 개선사항을 수정하는 활동으로 온톨로지 유지 및 개선 단계와 동일한 절차이다. 이 단계에서 온톨로지 구축의 개선사항이 발생하면 정제단계로 이관되어 삭제 및 갱신 작업이 이루어진다.

OTK 온톨로지 구축방법은 IEEE Standard 1074 1997의 전체 65개 활동에서 16개의 활동과 동일한 구축절차를 포함하고 있다. 즉, 프로젝트 관리 단계에서 1개 활동을 포함하여, 사

전 개발 단계에서 5개 활동, 개발 단계에서 3개 활동, 사후 개발 단계에서 3개 활동, 통합 단계에서 4개 활동과 동일하다. OTK 온톨로지 구축방법은 IEEE Standard 1074의 개발절차에 상응하는 과정을 포함한다.

3.2.2 CommonKADS

1) 구축방법

CommonKADS는 유럽의 ESPRIT IT 프로그램 개발사업의 한 부분이다. 이 사업은 대규모의 산업기반 지식 시스템을 만들기 위한 방법으로 약 15년 동안 장기간에 걸쳐 개발되었다. 이 구축방법은 온톨로지를 이용한 지식기반 시스템 개발에 광범위하게 활용되고 있다. 이 구축방법은 구조적인 지식공학을 지원하는 주도적인 방법론이며, 유럽의 대학 및 기업에서 이 방법론의 정확성을 검증해 왔다. 그러므로 유럽지역의 지식 분석과 지식 기반 시스템 개발의 사실상 표준이라고 할 수 있다. 산업 분야와 같은 지식 집약적인 작업, 처리에 대한 상세한 분석을 실행하게 하는 방법론을 제공한다 (Schreiber 1999).

CommonKADS에서는 지식 모델 구축 과정을 지식 확인, 지식 상세, 지식 정제의 세 단계로 구분하며 사전에 작업 및 조직 분석을 위해 타당성, 효과 조사를 수행한다. 작업 분석을 위해 행해지는 타당성 조사는 작업 및 조직 분석을 위해 필요한 절차로 설정되어 있다. 타당성 조사를 통해 문제점 및 가능성, 잠재적인 솔루션을 확인하고 배치한다. 가장 부각될 분야

〈표 2〉 CommonKADS 온톨로지 구축방법

단계	작업
지식의 확인 단계	-영역 파악(정보원, 용어 사전, 어휘 목록) -재사용을 위한 잠재적 모델 요소 목록(작업 및 영역 관련 요소)
지식의 상세 단계	-작업 템플릿 선택(초기 작업 분해) -초기 영역 개념 구성(주요 영역 정보형태) -지식 모델 상세 완성(부분적 지식기반을 가진 지식 모델)
지식의 정제 단계	-지식 모델 검증(문서 시뮬레이션, 추론 시스템 원형) -지식 기반 정제(지식 기반 완성)

와 목표로 하는 솔루션을 선정하기 위한 경제적, 기술적인 측면의 프로젝트 타당성을 결정한다. 목표로 하는 솔루션을 선정하기 위해 실행하는 효과 및 개선 조사는 작업의 상호관계에 대한 관찰을 수집하는 것이다. 지식 시스템 솔루션의 통합을 확인하기 위해 작업 변화를 결정하는 작업을 포함한다.

① 지식의 확인 단계에서는 실제 지식모델 상세를 위한 준비로서 지식 모델링에 사용될 만한 정보원을 확인한다. 정보원으로는 용어 사전이나 어휘목록 등이 있으며, 정보원의 작업 템플릿, 영역 스키마 같은 기존 모델링 요소를 조사한다. 보통 지식 항목 설명과 어플리케이션 작업의 특성화가 지식확인 시작점이 된다. 작업을 위한 정보원 탐색을 하는데, 용어사전이나 어휘목록을 작성한다. 보다 상세하게 작업의 성질을 연구하고 작업 형태를 바꾸거나 확인한다. 모든 잠재적 재사용 지식 모델 요소들을 구조화한다.

② 지식의 상세 단계에서는 지식모델에 대한

설명을 만든다. 이 단계의 목적은 지식 기반의 내용만 빼고 지식 모델 상세를 완료하는 것이다. 지식 확인 단계에서 확인된 재사용 모델링 요소의 목록을 이용해서 작업 템플릿을 선택하고 초기영역 스키마를 생성한다. 지식 모델 상세를 위해서는 두 가지 방식이 있다. 즉 추론지식을 가지고 시작하고 나서 관련 영역과 작업지식으로 이동하거나 영역과 작업 지식으로 시작해서 추론을 통해 이것을 연결시키는 방식이다. 이 단계에서 중요한 것은 영역 스키마의 존재이다.

③ 지식의 정제 단계에서는 지식 모델을 검증하기 위한 시도들이 이루어진다. 또한 완벽한 지식 인스턴스를 삽입한 지식기반을 완성하려 하는 단계이다. 초기 상세사항을 검증하는 중요한 방법은 지식 확인 동안 수집한 시나리오를 시뮬레이션으로 구성하는 것이다. 시뮬레이션은 문서나 소규모 프로토타입의 생성일 수도 있다. 시뮬레이션 결과를 통해 지식 모델이 요구된 문제 해결하는 방법

을 생성하는지 확인해야 한다. 긍정적인 검증 결과가 나오면 이것을 통해 지식 기반을 생성하는 것을 결정할 수 있다. 이 세 단계는 서로 연결되어 있고, 피드백의 과정을 거쳐서 이전 단계로 돌아갈 수 있다.

2) 비교분석

CommonKADS 온톨로지 구축방법은 타당성 조사와 더불어 지식 확인, 지식 상세, 지식 정제의 세 단계로 구성된다.

첫째, 타당성 조사 단계는 IEEE Standard 1074 1997의 타당성 조사 실행, 위험관리 및 기능분석 활동과 동일한 절차이다. 이 구축방법은 OTK 구축방법과 마찬가지로 타당성 조사 단계에서 어떤 영역의 문제점 및 가능성, 잠재적 솔루션을 확인하는 것으로 업무 분석을 위해서 행해지는 절차이다. 이것은 소프트웨어 개발 제한사항과 이점을 파악하는 타당성 조사 활동과 동일한 절차이다. 시스템의 문제점 파악과 업무분석은 IEEE Standard 1074 1997에서 기술적, 경제적 안전 및 보안을 반복적으로 분석하는 위험관리 활동 및 기능 분석 활동과 동일한 절차이다.

둘째, 지식 확인 단계는 IEEE Standard 1074 1997의 잠재적 접근법들 공식화, 소프트웨어 입수 활동과 동일한 절차이다. 이 단계에서는 특히 지식 모델링을 위해 필요한 정보원을 확인하는 작업으로 이루어진다. 지식 정보원 입수를 소프트웨어 입수로 대체한다고 가정했을 경우 소프트웨어 입수 활동과 동일한 절차가 된다. IEEE Standard 1074 1997은 소

프트웨어 개발 방법론으로 온톨로지 구축에서 필요한 지식 정보원을 입수하는 절차가 필요하지 않다. 그러므로 사전 개발 단계의 소프트웨어 입수를 정보원 입수로 대체한 절차로 간주하고자 한다. 지식 확인단계에서는 잠재적 재사용 지식모델 요소를 목록화한다. 이것은 잠재적 접근법 공식화 활동과 동일한 절차이다. 잠재적 접근법들은 시스템의 입수나 개발에 관련된 데이터나 요구사항을 기반으로 개발된다.

셋째, 지식 상세 단계는 시스템 아키텍처 개발을 비롯하여 구성 디자인 실행, 데이터베이스 디자인 활동과 동일한 절차이다.

지식 상세 단계에서는 지식 개념화 및 구조화가 이루어지며 이것들은 초기 지식 모델이 된다. 시스템에 대한 요구사항이나 기능 설명은 방법론, 표준 및 도구들에 의해 시스템 아키텍처로 전환된다. 이런 시스템 구조화는 개념 구조화 절차와 유사 단계로 간주할 수 있다. 구성 디자인 실행은 시스템 아키텍처를 상위 수준의 디자인 개념으로 전환하는 작업으로 역시 초기 지식 모델 생성을 이루는 과정과 유사성을 갖는다. 데이터베이스 디자인은 개념적 데이터베이스 디자인, 논리적 데이터베이스 디자인, 물리적 데이터베이스 디자인으로 구성되는데 개념을 상향식, 하향식, 중간식으로 계층 구조화하는 작업과 동일한 절차로 파악할 수 있다.

넷째, 지식 정제 단계는 검토실행, 시험절차 개발, 시험실행, 평가 결과보고 활동과 동일한 절차이다.

소프트웨어 개발의 전체 주기를 통해 검토

를 실행하는 것처럼 이 단계에서 시뮬레이션을 통해 지식 모델의 검증하는 절차를 거친다. 시뮬레이션은 문서나 프로토타입으로 형성될 수 있는데 이것은 시험절차를 개발하는 작업과 동일한 절차이다. 실행 코드, 모듈, 요소가 작동되어야 하며, 시험 데이터를 사용해서 전체 시스템과 통합 소프트웨어를 작동시키는 시험 실행 활동은 지식 모델을 검증하는 절차와 유사하다. 지식 모델에 따른 결과는 평가결과 보고 활동과 동일한 절차이다.

CommonKADS 온톨로지 구축방법은 IEEE Standard 1074 1997의 전체 65개 활동에서 15개의 활동과 동일한 구축절차를 포함하고 있다. 즉, 프로젝트 관리 단계에서 1개 활동, 사전 개발 단계에서 8개 활동, 개발 단계에서 2개 활동, 통합 단계에서 4개 활동과 동일하다. 사후 개발 단계에서는 동일한 절차로 판단되는 활동을 발견할 수 없었다. CommonKADS 온톨로지 구축방법에서는 운영 및 유지관리 절차가 명시되지 않았다.

3.2.3 ONIONS

1) 구축방법

ONIONS (Ontologic Integration On Naive Sources)는 영역 온톨로지를 통합하고 분석할 목적으로 개발되었다. 특히 ONIONS는 정보 입수에 있어 외래의 정보원을 통합하는 방법과 같은 지식 통합 문제를 해결하기 위해 시작되었다. ONIONS 접근법의 가장 구별되는 특징은 온톨로지 입수 방법이다.

ONIONS 구축방법은 온톨로지의 최종 표현 문제보다 모델링 과다방지(stopover: 어떻게 온톨로지를 과다하게 정제하는 것을 멈출 수 있는가) 및 지식 상관성(무엇이 개념적으로 상관성이 있는지 어떻게 언급할 수 있는가)과 같은 온톨로지 입수 문제에 관심을 둔다(Steve et al, 1997). ONIONS의 온톨로지 구축방법 6 단계는 다음과 같다(Jones 1998).

첫째, 어떤 분야의 확실한 정보원에서 언어 자료를 생성한다. 즉, 정보원은 각각 타당성을 입증 받은 것이어야 한다. 이 단계는 영역 지식도 구조화 되어 있지 않고, 명시적인 개념도 드러나지 않은 상태이다. 이 상태에서 검증 받은 언어자료를 수집하는 것이 이 단계의 목적이다.

둘째, 분류 구조의 적절성을 분석한다. 분류가 부족하면, 체계를 구축해야 한다. 이 단계의 목적은 용어 목록에서 주요 분류 구조를 발견하는 것이다.

셋째, 협의(local) 정보원을 분석한다. 이것은 단어의 개념 분석을 말한다. 상관 개념이 이용할 만한 것으로 판단되면 분류 기준에 중점을 두어야 한다. 분류 기준이란 해석 필드를 생성하기 위한 개념의 협의적 정의를 뜻한다. 정의가 부족하면 계층구조, 그룹화, 부울린 조합, 공리, 사전, 용어사전, 백과사전, 전문가 등으로부터 모든 암시를 찾아서 명백한 정의를 만든다.

넷째, 다중 협의(multi local) 정보원을 분석한다. 이 단계에서는 협의 정보원 분석 내용을 도식화하고 개념적 원칙(패러다임)을 유도

〈표 3〉 ONIONS 온톨로지 구축방법

단계	작업
1 단계	- 검증된 언어 자료의 수집
2 단계	- 분류구조 분석
3 단계	- 단어의 개념 분석
4 단계	- 정의된 개념 도식화
5 단계	- 온톨로지 라이브러리 구성
6 단계	- 온톨로지 라이브러리 분류 실행

하는 것이 목적이다. 도식화는 영역 온톨로지 구축을 통해 이루어 질 수 있다. 이런 영역 온톨로지는 공리를 필요로 하지만 템플릿이 된다. 표현 설명에서 사용된 요소들은 명백한 관계를 갖는다.

다섯째, 통합된 온톨로지 라이브러리를 만든다. 온톨로지 라이브러리는 모든 지역 정의와 다중 협의 구축에 사용된 통합적 정의인 개념적 원칙을 수록한다. 개념적 원칙을 설명하는 다른 온톨로지를 분명히 포함하는 영역 온톨로지를 구축한다. 그로 인해 충분한 협의의 정의를 만드는 것이 목적이다. 개념적 원칙을 분명히 하여 다른 종류의 협의의 정의와 연결하기 위해 협의의 정의를 충분하게 해주는 것이다. 이것은 협의에서 다중 협의 정도의 상태로 증가하는 수준까지 필요하다. 즉 모든 다른 종류의 정의를 포함하는 영역 온톨로지를 구축하는 정도를 의미한다. 사실 사전적 정의 보강은 협의 정보원 분석 단계에서 이루어졌다. 따라서 이 단계에서는 다중 협의로의 분류 상태로 처리하면 된다. 이 단계에서 일반 온톨로지

라이브러리를 구축하는 것은 기반을 잘 만든 상위 수준의 온톨로지를 구축하는 것과 동일한 효과를 얻을 수 있다.

여섯째, 온톨로지 라이브러리를 분류하고 응용한다. 자동 분류 시스템에서 영역 온톨로지를 실행하는 단계이다. 이 단계는 보급, 사용, 분류 및 모델 검증 단계에 속한다. 확장 및 부분 수용 또는 거부, 정제, 갱신, 다른 정보원과 통합하는 과정은 반복적으로 진행될 수 있다. 이전 단계에서 통합된 온톨로지는 적합한 상관성을 가지고 통합, 분석된 정보원으로 심층적인 역할을 할 수 있다. ONIONS를 방법론을 적용해 개발한 온톨로지로는 의학 분야의 ON9이 있다.

2) 비교분석

ONIONS 온톨로지 구축방법은 검증된 정보원의 체계화, 분류분석, 협의 정의 구성, 정의된 요소 조직, 재정리와 매핑, 분류기 실행의 여섯 단계의 절차로 구성되어 있다.

첫째, 정보원의 체계화 단계는 IEEE Standard 1074 1997의 입수 정보원의 평가

및 소프트웨어 입수 활동과 동일한 절차이다. 역시 소프트웨어 입수는 정보원의 입수로 간주하였다. 이 단계에서는 검증된 언어자료를 수집하는 것이 주된 작업이다. 타당성 있고 검증된 언어자료를 수집하는 일은 입수된 소프트웨어가 만족할 만한 매커니즘을 갖고 있는지 결정하는 입수 정보원 평가 활동과 동일한 성격의 절차로 판단할 수 있다.

둘째, 분류 분석 단계는 IEEE Standard 1074 1997의 시스템 아키텍처 개발 활동과 동일한 절차이다. 이 과정은 유사한 분류구조를 찾아내어 체계를 구축하는 개념화 작업은 시스템 구조화 작업과 동일성을 갖는다.

셋째, 협의의 정보원 분석 단계는 IEEE Standard 1074 1997의 기능 분석 활동과 동일한 절차이다. 단어의 개념 분석을 행하는 이 절차는 소프트웨어 개발에 있어 기능을 분석하는 활동과 동일한 성격을 지닌다.

넷째, 다중 협의의 정보원 분석 단계는 IEEE Standard 1074 1997의 구성 디자인 실행 및 데이터베이스 디자인 활동과 동일한 절차이다. 정보원을 분석한 내용을 도식화하고 개념적 원칙을 도출하여 초기 온톨로지를 생성하는 단계이다. 구성 디자인을 실행하고 데이터베이스를 디자인하는 실질적인 소프트웨어 개발과 동일한 절차이다.

다섯째, 통합된 온톨로지 라이브러리 생성 단계는 IEEE Standard 1074 1997의 세부 디자인 실행 활동과 동일한 절차이다. 통합된 온톨로지 라이브러리 생성이란 상위수준의 온톨

로지 구축과 같은 것으로 데이터베이스, 실행 코드, 소스코드를 통합하는 통합 실행과 동일한 절차이다.

여섯째, 온톨로지 라이브러리를 분류하고 응용하는 단계는 IEEE Standard 1074 1997의 시스템 운영, 소프트웨어 개선사항 확인, 문제보고 방법 시행, 소프트웨어 생성주기 재적용 활동, 시험 실행 활동과 동일한 절차이다. 이 단계는 보급, 사용, 분류 및 모델 검증 단계에 속하며 수정, 갱신이 반복적으로 진행되는 과정이다. 자동 분류 시스템에서 영역 온톨로지를 실행시키는 것은 소프트웨어 개발 절차에서 의도한 환경에 맞춰 설치된 소프트웨어 시스템을 사용하는 방식과 동일한 성격을 갖는다. 이 과정에서 피드백 데이터가 수집된다. 소프트웨어 개선사항에 대한 필요성을 확인하고 문제보고 방법 시행은 모델 검증 절차와 동일한 성격을 갖는다. 반복적인 수정 갱신 작업은 문제점을 수정하여 소프트웨어 생성주기에 재적용하는 것과 같다. 이 단계에서 소프트웨어 개발의 실행코드, 모듈, 요소의 작동을 시험하는 시험 실행 활동도 포함된다.

ONIONS 온톨로지 구축방법은 IEEE Standard 1074 1997의 전체 65개 활동에서 12개의 활동과 동일한 구축절차를 포함하고 있다. 즉, 프로젝트 관리 단계에서 포함되는 절차는 찾아볼 수 없었으며, 사전 개발 단계에서 4개 활동, 개발 단계에서 3개 활동, 사후 개발 단계에서 4개 활동, 통합 단계에서 1개 활동이 동일하다. ONIONS 온톨로지 구축방법에서는

운영 및 유지관리 절차가 명시되지 않았다.

3.2.4 Ontology Development 101

1) 구축방법

Ontology Development 101은 인공지능 분야 연구 기관인 스탠포드 대학 지식시스템 연구소의 핵심 연구과제 중 하나이다. 이 과제는 온톨로지 구축이 많은 어플리케이션에서 핵심요소가 되고 있는 시점에서 온톨로지를 구축하는 이유 및 방법론을 제공하고 있다. 이 연구를 수행한 Noy와 McGuinness는 와인 및 음식분야의 지식기반을 바탕으로 처음 온톨로지를 구축하고자 하는 개발자에게 도움을 주는 지침을 다음과 같이 제시하였다(Noy and McGuinness 2001).

첫째, 온톨로지의 영역과 범위를 결정한다. 온톨로지의 영역과 범위를 결정하는 일은 수록하려는 영역과 사용 이유, 온톨로지가 제공해야 할 답변이나 온톨로지의 이용자에 따라 정할 수 있다. 온톨로지 디자인 과정에서 영역과 범위는 계속 변화할 수 있으나, 기본적인 모델의 범위를 규정하는데 우선적으로 영역과 범위를 설정하는 것이 필요하다. 여기서 영역과 범위를 결정하는 방법 중 지식기반 시스템에서 답변해야 할 질문이 무엇인지 미리 예측하는 것이 효과적이다. 즉 적합한 질문을 설정하고, 이 질문 목록에 충분한 온톨로지 정보를 포함하고 있는지 확인할 수 있는 지침이 된다.

둘째, 기존의 온톨로지를 활용하는 방법을 고려한다. 특정 주제 영역의 온톨로지를 구축

할 경우에는 그 영역에 관련된 지식기반 시스템이나 온톨로지가 마련되어 있는지 조사해야 한다. 기존 온톨로지가 존재할 경우에 해당 온톨로지에 대한 소스를 활용한다.

셋째, 온톨로지서 중요 단어를 열거한다. 어떤 영역에서 이용자에게 설명하고자 하는 내용의 문장에서 중요하다고 생각되는 단어들을 모두 열거한다. 이 때 개념의 중복, 단어사이 관계, 개념이 속성인지 클래스인지를 고려하지 말고 우선 광범위하게 중요한 모든 단어를 열거하는 것이 필요하다.

넷째, 클래스와 클래스의 계층 구조를 정의한다. 해당 영역에서 열거한 단어 중 보다 독립적인 존재를 갖는 객체를 기술하는 단어를 선택하면, 이것이 클래스가 된다. 그리고 이 단어는 클래스 계층에서 앵커가 될 것이다. 클래스 A가 클래스 B의 상위 클래스라면, B의 모든 인스턴스도 A의 인스턴스이다. 이와 같이 클래스를 계층적 분류로 조직화할 수 있다. 클래스 계층 구조를 정의하는 데는 3가지 방법이 있다. 하향식 계층 정의는 영역에서 일반적인 정의와 그 다음에 생기는 개념의 특성을 정의하는 것으로 시작한다. 상향식 계층 정의는 특정 클래스, 최하단 클래스로부터 계층 구조를 정의해가는 방법이다. 그 다음 보다 일반적인 개념들을 그룹화한다. 복합식 계층 정의는 하향식, 상향식 계층 정의를 통합한 방식이다. 처음에 보다 부각된 개념을 정의하고, 다음 단계에서 그 개념들을 보다 적절하게 일반화, 특성화시킨다.

다섯째, 클래스의 속성인 슬롯을 정의한다.

〈표 4〉 Ontology Development 101 온톨로지 구축방법

단계	작업
1 단계	-온톨로지의 영역과 범위 결정
2 단계	-기존 온톨로지를 활용하는 방안
3 단계	-중요 단어의 열거
4 단계	-클래스와 클래스의 계층 정의
5 단계	-클래스의 속성(슬롯) 정의
6 단계	-속성(슬롯)의 패킷 정의
7 단계	-인스턴스 생성

클래스 자체로는 개념에 대한 충분한 정보를 제공할 수 없다. 그렇기 때문에 클래스를 정의한 후에는 개념의 내부 구조를 기술해야 한다. 중요단어를 열거한 목록에서 클래스로 선택하고 남은 단어들은 클래스의 속성이 될 수 있다. 속성을 어느 클래스에 붙여 설명할지도 결정해야 한다. 슬롯이 되는 객체의 속성에는 몇 가지 형태가 있다. 본래(intrinsic) 속성과 외래(extrinsic) 속성, 객체가 구조적일 경우의 부분들, 다른 개체와의 관계 등이 그것이다. 어떤 한 클래스의 하위 클래스는 그 클래스의 속성을 물려받는다. 슬롯은 그 속성을 가질 수 있는 가장 일반적인 클래스에 붙여져야 한다.

여섯째, 슬롯의 패킷을 정의한다. 슬롯은 가치 형태, 허용 가치, 가치 개수(cardinality) 및 슬롯이 취할 수 있는 가치의 다른 특징 등을 설명하는 다른 패킷을 가질 수 있다. 몇 가지 일반적인 패킷은 다음과 같다. 관계차수 패킷은 슬롯이 가질 수 있는 가치 값이 몇 개인가를 정의한다. 가치 형태 패킷은 슬롯에서 채울 수 있

는 가치 형태가 무엇인지를 설명한다. 일반적인 가치 형태는 String, Number, Boolean, Enumerated(Symbol), Instance 등이다. 인스턴스 가치 형태를 갖는 슬롯에 허용된 클래스들을 종종 슬롯의 범위라고 한다. 슬롯이 붙여진 클래스들을 슬롯의 영역이라고 한다.

일곱째, 인스턴스를 생성한다. 계층구조에서 클래스의 개체 인스턴스를 생성한다. 클래스의 개체 인스턴스를 정의하는 것은 다음 사항을 요구한다. 클래스의 선택, 선택한 클래스의 개체 인스턴스 생성, 슬롯 가치에 채우기가 그것이다. 인스턴스는 정의된 슬롯 가치를 갖는다.

2) 비교분석

Ontology Development 101 온톨로지 구축방법은 온톨로지 영역과 범위 결정, 기존 온톨로지 활용방안, 중요단어 열거, 클래스와 클래스 계층정의, 클래스의 속성 정의 속성의 패킷 정의, 인스턴스 생성의 7 단계의 절차로 구성되어 있다.

첫째, 온톨로지 영역과 범위 결정 단계는 IEEE Standard 1074 1997의 프로젝트 관리 계획 활동과 동일한 절차이다. 구체적으로 프로젝트 관리 계획 활동에서 소프트웨어 개발의 영역과 범위를 결정하는 내용을 포함한 것은 아니지만 프로젝트 조직구성을 상세히 하고 책임을 할당하는 작업을 하게 된다. 구성관리 및 평가, 문서화에 대한 표준, 방법론, 도구들이 설명되며 일정과 예산 분배까지 이 단계에서 진행하게 되므로 온톨로지 구축의 영역과 범위 결정 작업보다 더 포괄적인 활동이다. 프로젝트 관리 계획이라는 것이 소프트웨어 개발의 범위를 결정하는 것도 포함할 것으로 유추할 수 있다.

둘째, 기존 온톨로지 활용 방안 단계는 IEEE Standard 1074 1997의 소프트웨어 입수 방법 정의 활동과 동일한 절차이다. 소프트웨어 입수 방법 정의는 프로젝트에서 소프트웨어를 입수할 때 적용되는 절차이다. 온톨로지 구축방안에 이를 적용하여 보면 소프트웨어 입수를 온톨로지 입수로 간주할 수 있다. 어떤 소프트웨어를 입수하는 것이 가장 적합한지 판단하는 작업들은 기존 온톨로지를 어떻게 활용할 수 있는지 판단하는 이 절차와 동일한 성격을 갖는 활동이다.

셋째, 온톨로지에서 중요단어를 열거하는 단계는 IEEE Standard 1074 1997의 소프트웨어 입수 활동과 동일한 절차이다. 이 단계는 영역지식에서 필요한 모든 개념 입수하여 이것을 포괄적으로 나열하는 단계이다. 정보원 입수라는 측면에서 이것을 소프트웨어 개발의 입

수 활동으로 대체할 수 있다. 소프트웨어 입수는 통제된 방식으로 소프트웨어에 입수된 요소를 가져오는 것으로 전체 시스템 통합을 보장하게 된다.

넷째, 클래스와 클래스의 계층 구조를 정의하는 단계는 IEEE Standard 1074 1997의 시스템 아키텍처 개발 활동과 동일한 절차이다. 개념들을 상향식, 하향식, 조합식 접근법을 이용해 구조화하는 절차는 소프트웨어 개발에서 시스템의 구조를 개발하는 활동과 동일성을 갖는다. 소프트웨어 디자인 활동의 기반이 되는 시스템 아키텍처 개발은 온톨로지 개발의 기반이 되는 개념의 계층구조 정의와 동일할 절차이다.

다섯째, 클래스의 속성인 슬롯 정의 및 슬롯의 패킷을 정의하는 단계는 IEEE Standard 1074 1997의 구성디자인 실행 활동과 동일한 절차이다. 슬롯 및 패킷의 정의는 구체적으로 개념의 특성과 제한사항을 부여하는 절차로 온톨로지 구성을 완성해 가는 단계이다. 이것은 시스템 아키텍처를 상위수준 디자인 개념으로 전환하는 구성 디자인 실행 활동과 동일성을 갖는다.

여섯째, 인스턴스를 생성하는 단계는 IEEE Standard 1074 1997의 세부 디자인 실행 활동과 동일한 절차이다. 세부 디자인 실행 활동에서는 데이터 구조, 알고리즘, 소프트웨어 요소에 대한 통제 정보 등이 최종적으로 설명되어야 한다. 온톨로지를 구축하는데 있어 클래스의 인스턴스 생성 작업은 세부적인 디자인 실행 활동과 동일한 절차로 파악할 수 있다.

Ontology Development 101 온톨로지 구축방법은 IEEE Standard 1074 1997의 전체 65개 활동에서 6개의 활동과 동일한 구축절차를 포함하고 있다. 즉, 프로젝트 관리 단계에서 1개 활동, 사전 개발 단계에서 3개 활동, 개발 단계에서 3개 활동과 동일하다. 사후 개발 단계 및 통합 단계에 해당하는 절차는 찾아볼 수 없었다. Ontology Development 101 온톨로지 구축방법에서는 온톨로지의 유지 관리 및 평가 절차가 전혀 명시되지 않았다.

3.3 온톨로지 구축방법 비교분석 결과

IEEE Standard 1074 1997에 따라 OTK, CommonKADS, ONIONS, Ontology Development 101의 온톨로지 구축방법을 측정한 결과는 다음과 같다.

OTK 구축방법은 IEEE Standard 1074 1997의 65개 활동 중에서 16개의 활동과 동일한 구축절차를 포함하고 있다. 이 구축방법은 소프트웨어 개발 방법론의 프로젝트 관리, 사전개발, 개발, 사후개발 및 통제에 다섯 단계에 걸친 활동에 유사 절차를 포함하고 있었다. OTK 구축방법은 기존의 온톨로지 구축방법의 내용을 통합하여 작성된 것이므로 온톨로지 구축에 필요한 것으로 고려되는 절차를 전반적으로 모두 포함하고 있는 형태를 취하고 있다. 그로 인해 소프트웨어 개발 방안 전체 단계의 활동과 일치하거나 유사성을 갖는 OTK 온톨로지 구축 절차가 존재하였다.

OTK 온톨로지 구축방법은 타당성 조사를 통해 사전 개발 작업을 준비하고 시작과 정제 단계에서 본격적으로 온톨로지를 작성하는 작업을 진행한다. 그 후 평가 및 유지 개선 단계를 두고 있는데, 평가단계에서 온톨로지를 실험하고 부족한 패턴이 발견되면 정제 단계로 다시 돌아가서 반복 작업을 하게 된다. 평가 단계까지는 다른 온톨로지 구축방안의 반복적인 수정작업과 유사성을 갖는다. 그러나 사후 관리에 대한 유지 및 개선 단계를 구축방법 절차에 공식적으로 포함하고 있는 점이 다른 방법들과의 차이점이다. 온톨로지의 완성 후에도 지속적인 갱신이나 삭제 등이 발생한다는 가능성을 고려하여 엄격한 규칙을 통해 온톨로지를 갱신, 삭제하는 등의 체계적인 온톨로지 유지 관리의 필요성을 강조한 구축방법이다. 그러므로 이 구축방법은 타당성 조사 절차를 두어 온톨로지 구축의 사전 단계를, 유지 및 개선 절차를 두어 온톨로지 구축의 사후 관리 단계를 포함하고 있는 체계적인 구축방안을 갖고 있다.

CommonKADS 온톨로지 구축방법은 IEEE Standard 1074 1997의 65개 활동 중에서 15개의 활동과 동일한 구축절차를 포함하고 있다. 프로젝트 관리, 사전 개발, 개발 및 통합 단계에 걸쳐 이 온톨로지 구축방법 절차에 적용되는 활동들을 발견할 수 있었으나, 사후 개발 단계에서는 적용 활동이 존재하지 않았다. 이 점은 OTK 온톨로지 구축방법과 가장 큰 차이점으로 유지 관리 절차가 CommonKADS 온톨로지 구축방법에는 명시

되어 있지 않다. 개발 완료 후에 발생할 수 있는 온톨로지 유지관리에 대한 사후 절차가 포함될 때 온톨로지 구축의 완성도 및 지속성이 강화될 것이다.

CommonKADS 온톨로지 구축방법은 정확하게 온톨로지 구축을 위한 방법론이라기 보다 지식기반 시스템을 구축하기 위한 체계적인 방법론이다. 이 구축방법에서는 작업 모델에 따른 워크시트 문서 형식을 제공하고 있어 방법론을 체계적으로 적용할 수 있다. 절차와 형식의 틀을 체계적이고 충분하게 제공하고 있다는 것이 장점이다. 그러나 광범위한 지식기반 시스템 구축을 위한 방법을 제공하고 있어 부분적인 영역의 온톨로지 구축 등에는 모든 절차를 적용할 필요가 없다. 특정 온톨로지 구축에 적합한 절차를 선택하여 사용할 필요성이 있을 것으로 보인다.

ONIONS 온톨로지 구축방법은 IEEE Standard 1074 1997의 65개 활동 중에서 12개의 활동과 동일한 구축절차를 포함하고 있다. 프로젝트 관리 단계에 적용되는 절차는 없었으며, 사전개발, 개발, 사후 개발 및 통합단계의 활동들과 유사성을 갖는 절차를 포함하고 있다. 사전 개발 단계에서도 개념조사 활동에 속하는 절차가 없었다. 이것은 타당성 조사나 의견요구 확인과 같은 활동을 전혀 실행하지 않고 있다는 것을 나타낸다. ONIONS 온톨로지 구축방법은 실제 개발을 수행하기 이전에 필요한 절차들을 포함하고 있지 않은 단점이 있다. 그러나 최종 단계에서 온톨로지 모델을 검증하고 확장, 갱신, 정제를 반복적으로 실행

할 수 있도록 절차를 마련해 놓고 있다. ONIONS 온톨로지 구축방법의 최종 단계는 평가 및 유지관리가 통합된 복합절차이다.

ONIONS 온톨로지 구축방법은 동일한 주제 분야의 용어나 온톨로지를 통합할 목적으로 만들어진 방법론이기 때문에 그 절차에 있어서도 다른 양상을 보인다. 결국 통합된 온톨로지 라이브러리를 만들고 이를 분류 응용하는 절차라고 할 수 있다. OTK, CommonKADS의 경우 자체적인 온톨로지 생성을 위한 구조적인 절차를 포함하고 있다. 반면 ONIONS의 경우는 주제 영역의 기존 온톨로지나 용어를 활용하기 위한 통합방법을 제시하고 있다. IEEE Standard 1074 1997에 적용하여 12개의 활동에 그 절차가 적용될 수 있어 OTK 및 CommonKADS와 대응하는 형식에 차이가 없는 것으로 보인다. 그러나 ONIONS 온톨로지 구축방법은 내용상 기존 용어나 온톨로지의 통합에 중점을 두고 있어, 두 구축방법과 다른 성격을 지니고 있다.

Ontology Development 101 온톨로지 구축방법의 경우 IEEE Standard 1074 1997의 65개 활동 중에서 6개의 활동과 동일한 구축절차를 포함하고 있다. 사전개발 및 개발 단계에 대응하는 절차가 대부분이었다. 사후 개발과 통합 단계에 대응하는 절차를 발견할 수 없었다. 이 구축방법의 경우도 ONIONS 온톨로지 구축방법과 마찬가지로 온톨로지 개발을 위한 타당성 조사와 같은 준비 작업이 포함되어 있지 않다. 이 구축방법의 경우 간략하게 온톨로지를 구축하는 방법을 명시하고 있어서, 주로

사전개발 및 개발 단계 활동에 대응하는 절차만이 포함되어 있다. 사후 개발 및 통합 단계의 활동과 일치하는 절차가 전혀 포함되어 있지 않다. 구조적인 온톨로지 구축방법이라고 할 수 없다. 온톨로지 구축에 필요한 전체적인 구축 절차를 포함한 완성도를 갖춘 방법으로 판단할 수 없다.

Ontology Development 101 온톨로지 구축 방법은 개념 구조화의 기초적인 절차를 설명하고 있다. 온톨로지 구축에 필요한 내용을 상세하게 설명하고 있어 접근이 용이하다. 그러나 개념의 구조화에 국한된 기본적인 내용만을 제시하고 있어 온톨로지 구축에 필요한 전체 과정을 체계적으로 설정한 방법론으로 볼 수 없는 단점이

<표 5> IEEE 표준에 따른 온톨로지 구축방법 비교분석

Section	Activity Group	IEEE 표준기준	기존 온톨로지 구축방법				
		Activity	OTK	Common-KADS	ONIONS	Ontology Development 101	
프로젝트 관리	프로젝트 시작	소프트웨어 생성 주기 절차 생성					
		평가 실행					
		프로젝트 자원 할당					
		매트릭스 정의					
	프로젝트 계획	평가 계획					
		구성 관리 계획					
		시스템 이전 계획					
		설치 계획					
		문서화 계획					
		훈련 계획					
	프로젝트 모니터링 및 통제	위험 관리					
		프로젝트 관리					
		소프트웨어 생성주기 절차 향상 필요 확인					
		기록 보유					
	사전 개발	개념 조사	매트릭스 데이터 분석 및 수집				
			의견 요구 확인				
잠재력 접근법들 공식화							
타당성 조사실행							
의견 요구의 확정 및 정제							
시스템 할당		기능 분석					
		시스템 아키텍처 개발					
	시스템 요구 사항 분해						

	소프트웨어 입수	입수된 소프트웨어 요구 사항 확인			
		입수 정보원의 평가			
		소프트웨어 입수 방법의 정의			
		소프트웨어 입수			
개발	요구 사항	소프트웨어 요구 사항 개발 및 정의			
		인터페이스 요구 사항 정의			
		인터페이스 요구 사항 통합 및 우선 순위 결정			
	디자인	구성 디자인 실행			
		데이터베이스 디자인			
		디자인 인터페이스			
		세부 디자인 실행			
	실행	실행 코드 생성			
		운영 문서 생성			
통합 실행					
사후 개발	설치	소프트웨어 배포			
		소프트웨어 설치			
		운영 환경에서 소프트웨어 수용			
	운영 및 지원	시스템 운영			
		기술지원 및 상담 제공			
		지원 요구 로그 유지			
	유지	소프트웨어 개선 사항 확인			
		문제보고 방법 시행			
		소프트웨어 생성 주기 재적용			
	폐기	이용자에게 통보			
		병력 작동 실행			
		폐기 시스템			
통합	평가	검토 실행			
		추적 매트릭스 생성			
		결산 실행			
		시험 절차 개발			
		테스트 데이터 생성			
		시험 실행			
		평가 결과 보고			
	소프트웨어 구성 관리	구성 확인 개발			
		구성 관리 실행			
		상태 계산 실행			
	문서 개발	문서 실행			
		문서 작성 및 배포			
	훈련	훈련 자료 개발			
		훈련 프로그램 인가			
		훈련 프로그램 실행			

있다. <표 5>에서 사례별 온톨로지 구축방법은 IEEE 표준으로 분석한 결과를 살펴볼 수 있다.

4. 온톨로지 구축방법 모델 제안

온톨로지 구축방법은 개발 실례에 따라 구축표준안을 제시하지 않고 해당 분야나 기관의 특성만을 고려하고 경험만으로 구축된 경우가 대부분이다. 즉, 그 목적과 대상영역에 따라 다양한 측면에서 프로젝트가 진행되어 왔다. 본 연구의 온톨로지 구축방법 모델은 사례별 온톨로지 구축방법에서 공통적으로 채택하고 있는 절차를 기반으로 하였다. 더불어 IEEE Standard 1074 1997을 통한 비교분석 결과에 따라 온톨로지 구축에 필요한 요소로 판단되는 절차를 접목하여 제안되었다. 온톨로지 구축방법에 대한 사례를 고찰하고 각각의 장단점을 비교한 결과를 토대로 한 효율적인 온톨로지 구축방법은 아래와 같다.

첫째, 온톨로지 구축에 필요한 사전 준비작업으로 타당성 조사 절차를 포함해야 한다. 온톨로지 구축의 문제점 및 가능성을 파악하는 사전 개발 절차를 수행하여 실제 구축 시에 발생할 수 있는 오류를 최소화하고 올바른 구축 방향을 정립할 수 있다. 즉, 온톨로지 구축은 OTK 및 CommonKADS 온톨로지 구축방법에서와 같이 타당성 조사 절차를 포함하는 것이 바람직하다. 구조적인 온톨로지 구축방법일수록 실제적인 개발 이전에 타당성 조사와 같은 사전 작업 절차를 포함하고 있다는 것을

알 수 있다. 사전 개발 준비 절차를 통해 실제 개발의 진행을 신속하게 처리할 수 있으며 문제발생에 대한 효율적인 대처가 가능하다.

둘째, 온톨로지 구축 초기 단계에는 구축 목표와 주제 영역을 설정해야 한다. 다양한 정보원을 통해 주제 영역의 어휘를 수집하여 개념을 파악하는 작업이 필요하다. OTK에서 시작단계, CommonKADS에서 지식의 확인단계, ONIONS의 언어자료 생성, Ontology Development 101 온톨로지 구축방법에서 중요단어 열거 단계에서 공통적으로 주제 영역의 어휘를 파악하는 절차를 거치고 있다. 주제 영역에 대한 어휘는 검증된 개념을 수집하는 것이 바람직 할 것이나, 다양한 어휘 수집을 위해서는 브레인스토밍과 같은 방식을 취할 수 있다. 주제 영역의 어휘는 사전, 용어집에서 부터 주제 영역의 전문가들이 사용하는 언어개념에 이르기까지 다양한 정보원에서 수집하도록 한다. 그리고 온톨로지 구축을 통해 해결해야 할 질문이 무엇인지 설정한다. 즉 적합질문을 작성하여 구축된 온톨로지의 완성도를 평가할 수 있다. 그러나 적합질문 설정을 통해 개념과 관계를 추출할 수 있다. OTK의 시작단계, Ontology Development 101 온톨로지 구축방법에서 영역 및 범위 결정 단계에서 적합질문 작성을 채택하고 있다.

셋째, 기존 온톨로지를 통합하는 절차가 마련되어야 한다. 온톨로지 구축의 목적이 정보의 공유 및 재사용인 만큼 온톨로지 구축방법에 기존 온톨로지를 활용할 수 있는 방법이 마

련되어야 한다. 부분적으로는 기존의 온톨로지의 개념 설정과 관계구조 기반을 이용하는 방법과 전체적으로는 구축된 온톨로지 전체를 수용하는 방법 등으로 수준별 통합 절차를 설정해야 한다. 그러나 이런 기존 온톨로지 통합 절차가 구축방안에 포함되는 경우에도 이 절차를 수용해야 할 필요는 없다. 상황에 따라 해당 온톨로지 구축의 목적 및 환경에 맞춰 기존 온톨로지를 통합할 것인지 또는 새로운 온톨로지를 구축할 것인지를 결정해야 한다. 기존 온톨로지 통합 방안은 ONIONS 및 Ontology Development 101 구축방법에서 나타난다.

넷째, 주제 영역에서 수집된 어휘와 기존 온톨로지 통합 등의 과정을 거친 후에 개념과 관계를 규정하여 온톨로지를 구성한다. 온톨로지 구성은 사례별 온톨로지 구축방법에서 모두 공통적으로 나타나는 주요 절차이다. 여기서 온톨로지 구축 환경에 맞는 지원도구를 활용하는 방안에 대한 지침을 제공할 필요가 있다. 온톨로지 구축은 개발자의 지식과 경험에 따라 다른 결과를 나타낼 수 있을 것이다. 이런 개인적

인 편차를 최소화하고 개발의 객관성을 확보하기 위해 환경에 맞는 지원도구를 이용하는 방법을 마련해야 한다. 온톨로지 구축은 기술적으로 개발된 지원도구를 통해 이루어질 가능성이 많아지고 있으며, 이런 온톨로지 지원도구의 개발이 활성화되고 있다. 또한 온톨로지 구축 절차에 대한 공식적인 작업형식을 문서로 규정해야 한다. 사례별 온톨로지 구축방법에서 사용되고 있는 적합질문이나 온톨로지 요구사항서, 작업 모델별 워크시트와 같은 정형화된 형식을 통해 작업 절차를 공유하고 협업의 효율성을 높일 수 있기 때문이다. 그러므로 온톨로지 구축에 필요한 모든 절차에 필요한 형식을 공식화하여 작업 절차의 일관성 및 객관성을 확보할 필요가 있다.

다섯째, 개발 완료 후 시험 테스트, 시뮬레이션 등의 방법으로 구축 온톨로지를 검증하는 절차를 거쳐야 한다. 온톨로지 검증 절차로 시험 이용자들로부터 오류 사항에 대한 사용 피드백을 받을 수 있다. 또한 시뮬레이션을 구성하여 문제점을 확인할 수 있다. 이용 시나리오

〈표 6〉 온톨로지 구축 방법 모델 제안

단계	구축 방법 제안
1 단계	-온톨로지 구축 타당성 조사(온톨로지 구축의 문제점 및 가능성 파악)
2 단계	-온톨로지 구축 목표 성립(주제 영역 설정, 어휘 수집, 적합질문 작성)
3 단계	-기존 온톨로지 통합 방안(지식 재사용)
4 단계	-온톨로지 구성(개념 및 관계 규정, 지원 도구 확인, 정형화된 문서 형식 사용)
5 단계	-온톨로지 검증 및 평가(시험테스트, 시뮬레이션, 오류 사항 피드백)
6 단계	-온톨로지 유지 관리(수정, 갱신 사항 지속적 반영)

및 적합질문에 대한 결과 등으로 온톨로지를 평가할 수 있다. 이런 검증절차는 개발된 온톨로지의 결과를 평가하기 위해 반드시 포함되어야 할 절차이다.

여섯째, 사후 개발 단계로서 유지관리 절차를 공식적으로 포함시키는 것은 중요한 문제이다. 온톨로지 구축이 완료되었다고 하더라도 사후 필요한 수정, 갱신 사항이 발생하게 된다. 실제 온톨로지를 사용하면서 발생한 문제점을 다시 반영하여 온톨로지를 개선하는 작업을 지속적으로 유지함으로써 지속적인 온톨로지의 활용이 가능하기 때문이다. 따라서 구조화된 완성도가 높은 온톨로지 구축방법을 설계하기 위해서는 개발 이후 작업에 대한 유지관리 절차에 중점을 두고 이를 공식절차로 포함할 필요가 있다.

5. 결론

온톨로지는 시맨틱웹의 발전을 위해 반드시 구현되어야 할 기술이다. 필요기술 가운데 온톨로지 구축을 통해 지식 영역 기반을 형성하는 일은 웹의 의미검색에 필요한 핵심 기술이다. 본 연구에서는 이제까지 제안되어 온 대표적인 온톨로지 구축방법을 살펴보고 향후 온톨로지 구축에 적합한 구축방법 모델을 제시하였다.

성격이 다른 사례별 온톨로지 구축방법을 비교하기 위해 소프트웨어 개발방법론 표준을 적용하였다. 이를 통해 각 구축방법이 포함하

고 있는 절차가 구조적으로 완전한 것인지 파악하였다. 본 연구의 결과를 요약하면 다음과 같다.

- 1) 사례별 온톨로지 구축방법에서 사전, 사후 개발 단계까지 과정에서 모든 절차를 수용하고 있는 방법은 OTK 온톨로지 구축방법이다.
- 2) CommonKADS 온톨로지 구축방법의 경우 사후 관리 절차가 명시되지 않았고, 지식기반 시스템 구축에 중점을 둔 방안으로 모든 온톨로지 구축에 적용하기에는 적합하지 않다.
- 3) ONIONS 온톨로지 구축방법은 특정 주제 분야의 용어나 온톨로지 통합에 목적을 둔 방법론이다. 그 방법이 OTK 및 CommonKADS, Ontology Development 101 온톨로지 구축방법처럼 일반 온톨로지 구축과 다른 성격을 지니고 있다.
- 4) Ontology Development 101 온톨로지 구축방법은 경험사례를 바탕으로 기초적인 온톨로지 구성안을 제공한 것으로 구조적인 구축방법으로 보기 어렵다. 이 구축방법은 완성적인 온톨로지 구축방법을 제공한 것이 아니므로 다른 온톨로지 구축에 활용되는 방법으로 일반화할 수 없다.

이와 같은 결과를 바탕으로 통합적인 온톨로지 구축방법 모델을 다음과 같이 제안한다.

- 1) 사전개발 단계로서 타당성 조사를 포함한다. 온톨로지를 구축하기 이전에 타당성 조사를 실행함으로써 실제 구축 시 발생할 수 있는 문제점을 최소화한다.
- 2) 초기 구축 단계로서 온톨로지 구축 목표를

성립한다. 이 단계에서 주제영역을 정하고 어휘를 수집한다.

- 3) 지식 재사용을 위해 기존 온톨로지를 통합하는 방안을 고려한다.
- 4) 주제 영역의 개념과 관계를 규정하면서 온톨로지를 구성한다. 이 단계에서 지원받을 수 있는 도구를 확인하고 활용한다.
- 5) 완성된 온톨로지를 평가하고 검증한다.
- 6) 사후개발 단계로서 구축된 온톨로지를 지속적으로 사용하기 위한 유지관리 절차를 포함한다.

본 연구는 IEEE Standard 1074 1997을 기준으로 사례별 온톨로지 구축방법의 절차를 분석하였다. 이런 온톨로지 구축방법의 비교분석을 기반으로 각각의 온톨로지 구축방법의 장점을 수용하고 단점을 보완함으로써 완성도 있는 온톨로지 구축방법 모델을 제안하였다. 그러므로 향후 분야별 온톨로지 구축 시 본 연구에서 제안한 온톨로지 구축방법 모델안이 사용될 수 있을 것으로 기대할 수 있다.

참고문헌

(국 내)

김이란. 2001. 『온톨로지 의미정보를 이용한 RDF 스키마 생성에 관한 연구』, 석사학위논문, 연세대학교 대학원, 문헌정보학과.

모기 아키라 편저. 2000. 『한영일 전기전자 용어사전』, [서울]: 대광서림.

오삼균. 2002. 시맨틱웹 기술과 활용방안. 『정

보관리학회지』, 19(4): 297-320.

이현실. 2003. 『온톨로지 기반 한의학 처방 지식관리시스템 설계에 관한 연구』, 박사학위논문, 중앙대학교 대학원.

전기전자공학사전편찬위원회 편. 1996. 『전기전자공학 대사전』, [서울]: 집문사.

정도현. 2003. 『시소러스를 기반으로 한 온톨로지 시스템 구현에 관한 연구』, 석사학위논문, 연세대학교 대학원, 문헌정보학과.

이재호. 2003. 시맨틱 웹의 온톨로지 언어. 『정보과학회지』, 21(3): 18-27

양정진. 2003. 시맨틱 웹에서의 온톨로지 공학. 『정보과학회지』, 21(3): 28-35.

최호섭, 옥철영. 정보검색 시스템과 온톨로지. 『정보과학회지』, 22(4): 62-71.

(국 외)

Berners Lee, Tim. 2003. "WWW Past, present and future." [cited 2004.3.27]. <http://www.w3.org/2003/Talks/0922_rsoc_tbl/slide30_0.html>.

Davies, John, Fensel, Dieter, Harmelen Frank Van. 2003. Towards the semantic web: ontology driven knowledge management. [s. l.]: Wiley.

Dean Jones, Trevor Bench Capon, Pepijin Visser. 1998. Methodologies For Ontology Development. In Proc.

- IT&KNOWS Conference of the 15th IFIP World Computer Congress, Budapest*. [cited 2004. 4. 2]. <http://cweb.inria.fr/Resources/ONTOLOGIES/methodo_for_onto_dev.pdf>.
- Devedzic, Vladan. 2002. Understanding Ontological Engineering. *University of Belgrade, FON School of Business Administration, Department of Information Systems*. [cited 2004.4.4]. <<http://fon.fon.bg.ac.yu/~devedzic/CACM2002.pdf>>.
- Ding, Ying, Fensel, Dieter, Klein, Michel, Omelayenko, Borys. 2002. The Semantic Web: Yet Another Hip?. *Data and Knowledge Engineering*, 41. [cited 2004.3.30]. <www.cs.vu.nl/~mcaklein/papers/DKE41.pdf>.
- Engineering and Managing Knowledge: commonKADS. [cited 2004.4.1]. <http://www.commonkads.uva.nl/frameset_commonkads.html>.
- Gamper, Johann, Nejd, Wolfgang, Wolpers, Martin. Combining Ontologies and Terminologies in Information Systems. [cited 2004. 4. 4]. <<http://www.kbs.uni-hannover.de/Arbeiten/Publikationen/1999/tke99/>>.
- Gruninger, Michael, Fox, Mark S. 1995. Methodology for the Design and Evaluation of Ontologies. *Workshop on Basic Ontological Issues in Knowledge Sharing*. pp.1-10.
- Lopez, Fernandez M. 1999. Overview Of Methodologies For Building Ontologies. *In Proceedings of the IJCAI 99 Workshop on Ontologies and Problem Solving Methods: Lessons Learned and Future Trends*, CEUR Publications. [cited 2004. 4. 5]. <<http://www.ontology.org/main/presentations/madrid/analysis.pdf>>.
- Noy, Natalya F., McGuinness, Deborah L. 2001. Ontology Development 101: A Guide to Creating Your First Ontology. *Stanford Knowledge Systems Laboratory Technical Report KSL 01 05 and Stanford Medical Informatics Technical Report SMI 2001 0880*. [cited 2004. 3. 13]. <http://www.ksl.stanford.edu/people/dlm/papers/ontology_tutorial_noy_mcguinness.pdf>.
- Schreiber, Guus et al. Knowledge engineering and management the

- CommonKADS
MIT ed, 1999, *Methodology*, MA: The MIT Press, pp.167-186.
- Software Engineering Standard Committee of the IEEE Computer Society sponsored, IEEE Standard Board approved, IEEE Standard for Developing Software Life Cycle Processes, The Institute of Electrical and Electronics Engineers, Inc, 1997, [cited 2004, 4, 18]. <<http://kom.aau.dk/~oa/Teaching/2004/E4/E46/docs/IEEE1074.pdf>>.
- Steve, G., Gangemi, A., Pisanelli, D.M, 1997, Integrating Medical Terminologies with ONIONS Methodology, *Information Modelling and Knowledge Bases VIII, Amsterdam, IOS Press*, [cited 2004, 4, 3]. <<http://www.loa.cnr.it/Papers/onions97.pdf>>.
- Ushold, Mike, Gruniger, Michael, 1996, "ONTOLOGIES: Principles, Method and Applications." *Knowledge Engineering Review*, 11(2): 14-34.
- Ushold, Mike, 1996, Building Ontologies: Towards a Unified Methodology. *Proceedings of Expert Systems, the 16th Annual Conference of the British Computer Society Specialist Group on Expert Systems*, pp.1-6.
- W3C, 2004, OWL Web Ontology Language Overview, [cited 2004, 3, 27]. <<http://www.w3.org/TR/2004/REC-owl-features-20040210/>>.
- W3C, 2004, OWL Web Ontology Language Guide, [cited 2004, 3, 27]. <<http://www.w3.org/TR/2004/REC-owl-guide-20040210/>>.
- Zyl, J., Corbett, D, 2000, A Framework for Comparing Methods for Using or Reusing Multiple Ontologies in an Application, *International Conference on conceptual Structures (ICCS)*, [cited 2004, 4, 13]. <<http://www.cis.unisa.edu.au/~cisjvz/vanzylICCS.pdf>>.