

# 메타데이터의 생성 및 관리를 위한 RDF 문서 검증기와 N-Triple 생성기

조성훈\* · 송병열\*\* · 조현규\*\*\* · 최의인\*\*\*\*

## 요약

전자상거래가 활성화됨에 따라 웹 상에서 공유되는 기업 정보와 카탈로그(catalog) 같은 비즈니스 정보의 양이 급속도로 증가하였다. 이를 효율적으로 처리하고자 RDF 메타데이터 프레임워크를 이용한 메타데이터의 관리 필요성이 대두되었다. RDF 메타데이터의 관리는 RDF/RDFS의 지식에 대해 제약받지 않고 웹 자원을 기술할 수 있어야 하고 메타데이터의 유효성을 보장하기 위한 환경을 지원하여야 한다. 그러나 이를 적절히 지원할 수 있는 연구가 아직까지 미흡한 상황이다. 이를 위해 본 논문에서는 RDF/RDFS 문서를 다양한 인터페이스를 통해 생성 및 관리가 가능하며, RDF를 이용하여 기술된 메타데이터를 N-Triple로 제공함과 동시에 메타데이터의 유효성 검증이 가능하도록 하였다.

## RDF Document validator and N-Triple Generator for Creation and Maintenance of Metadata

Sung-Hoon Cho\* · Byoung-Youl Song\*\* · Hyun-Gyu Cho\*\*\* · Eui-In Choi\*\*\*\*

## ABSTRACT

The quantity of business information like to be shared enterprise information and business catalog quickly increased because of activating e-commerce. Necessary of metadata management which is using RDF was required. RDF metadata management does not have to be restricted by RDF/RDFS knowledges, has to describe web resources and has to support metadata validation. But research which supports this is not enough. In this paper we enable to create and manage RDF/RDFS document using various interfaces, change from metadata to N-Triple and verify metadata validation.

**키워드 :** RDF/RDFS, 메타데이터(Metadata), N-Triple

### 1. 서론

거대한 분산정보 시스템처럼 다양하고 무수히 많은 정보를 가지고 있는 현재 웹은 정보교환과 비즈니스를 위한 중심적 위치를 차지하고 있다. 이 수많은 정보들을 효율적으로 관리, 검색하기 위해 정보 자원에 대한 기술 정보(descriptive information)를 제공하는 메타데이터가 등장하게 되었다.

메타데이터는 "데이터에 대한 구조화된 데이터"라고 정의할 수 있는데 이 메타데이터는 웹 상에서 URL을 통해 불러낼 수 있는 자원(resource)에 대한 위치나 내용을 기술하여 이를 관리하고 검색하는데 효율적으로 사용할 수 있다. 이러한 메타데이터의 중요성의 인식과 더불어 다양한 형식

의 메타데이터가 연구되었지만 특정 단체를 중심으로 연구되고 사용되어온 메타데이터는 어휘나 구조에 있어서 서로 다른 표현법(notation)을 사용하고, 상의한 의미 구조를 가지고 있어서 특정 단체 중심으로 사용하는 메타데이터는 서로 다른 응용프로그램이나 시스템에서 상호운용성(Interoperability)을 보장할 수 없는 문제점이 존재한다.

이를 위해 W3C가 상호운용성 보장을 목적으로 RDF를 제안하였다. RDF(Resource Description Framework)는 Warwick Framework에 많은 영향을 받으며 개념적 수준에서 메타데이터의 구조(framework)를 제시하였고, RDF는 이를 실제적으로 구현한 것이다.

RDF는 특정 분야에서 개개의 자원을 기술하는 것이 아니라 적용 분야에 따라 메타데이터 요소를 한정하기 위한 것으로 이를 이용하면 특정 메타데이터의 요소를 다른 메타데이터에서도 사용할 수 있다는 장점이 있다. RDF는 어휘의 표준화를 통해 여러 분야에서 동일한 의미로 사용하거나 확장할 수 있고 분산된 속성을 통합할 수 있는 구조로서, 메타데이터를 별도로 등록할 필요가 없다[1, 2, 9, 10].

근래에 들어 웹 상의 e-비즈니스에서 카탈로그와 같은 비

\* 본 연구는 한국전자통신연구원의 지원으로 수행되었음.

† 준회원 : 전남대학교 대학원 컴퓨터공학과

\*\* 정회원 : 한국전자통신연구원 지능형로봇연구단/지능형서비스플랫폼연구팀

\*\*\* 정회원 : 한국전자통신연구원 책임연구원, 지능형로봇연구단/지능형서비스플랫폼연구팀장

\*\*\*\* 정회원 : 전남대학교 컴퓨터공학과 부교수

논문접수 : 2004년 5월 18일, 심사완료 : 2004년 8월 17일

즈니스 정보에 대한 메타데이터를 RDF로 기술한 전자상거래 시스템을 구축하려는 움직임을 보이고 있다. 그러나 카탈로그에 대한 메타데이터를 저작할 수 있는 연구가 미흡한 실정이다. IsaViz, RDFedt, RDFAuthor, Protégé-2000과 같은 저작도구가 제시되었지만, RDF 요구사항(requirement)을 완벽히 만족시키지 못하고 있다. 즉, 다양한 인터페이스로 지원 가능한 RDF의 표기법을 모두 만족 시키는 저작도구가 없는 것이 현실이다.

따라서 본 논문에서는 RDF의 표기법에 적합한 다양한 저작 인터페이스를 제공하고, 생성과 동시에 웹 상에서 직접 활용 가능하도록 유효성 검증까지 제공하는 상위 수준의 RDF 저작도구의 설계하였다.

이를 위해서 하나의 RDF 문서를 텍스트, 트리, 그래프 기반으로 생성, 편집, 분석 가능하며, 각 저작 모듈 간 상호 전환하여 메타데이터를 처리할 수 있도록 구성하였다. 그리고 단순한 편집기 기능을 확장하여 생성된 문서의 유효성(validation)을 편집기 자체에서 검증할 수 있는 유효성 검증기를 제안하였고 N-Triple 생성기를 만들어서 RDF의 또 다른 표현 형식으로 표현 가능토록 함으로써 다른 저작도구와 RDF 문서 외에 상호 교환할 수 있도록 제안하였다.

본 논문의 구성은 2장에서 RDF와 관련 연구 내용을 살펴보고, 3장에서는 RDF 문서의 파싱과 유효성 검증을 위한 모듈을 제안하고, 4장에서는 설계한 RDF 문서 생성기와 N-Triple 생성기의 모듈을 실행하여 각 각의 기능과 설계 내용을 알아보고 5장에서 어플리케이션간의 기능 비교를 통해 장점을 알아보고 마지막으로 6장에서 결론과 향후 과제를 제시한다.

## 2. 관련 연구

메타데이터는 구조적인 기술, 효율적인 관리, 상이한 어플리케이션에서의 상호운용성을 보장하기 위해 RDF와 같은 메타데이터 표준을 필요로 한다. RDF는 W3C가 상호운용성 보장을 목적으로 제안한 것으로서 상이한 메타데이터 체계를 인정하고 이를 하나의 통합된 구조 속에 모두 수용할 수 있기 때문에 메타데이터 간의 상호운용성을 보장한다. 또한 RDF는 어휘의 표준화를 통해 여러 분야에서 동일한 의미로 사용하거나 확장할 수 있으며, 분산된 속성을 통합할 수 있는 구조로서, 메타데이터를 별도로 등록할 필요가 없으며 XML에 기반한다는 장점을 가지고 있다[3, 8].

이러한 RDF의 특징과 장점을 각 분야에 적용하기 위해서 RDF 메타데이터 문서를 생성하고, RDF로 기술된 웹 자원을 가시적으로 표현하여 효율적인 자원 관리를 하기 위한 연구가 진행되었다.

### 2.1 IsaViz

IsaViz는 Emmanuel Pietriga에 의해 그래프의 형태로 표현되는 RDF 모델을 브라우징하고 저작하기 위한 비주얼한

환경을 제공할 목적으로 개발되었다. IsaViz는 Xerox사의 비주얼 변환 머신(Visual Transformation Machine)을 기반으로 하고 있으며, 휴렛팩커드 연구소의 Jena, Apache Software의 Xerces Java 2를 포함하고 있으며 AT&T 연구소에서 개발한 Graphviz 라이브러리를 사용하였다. 또한 IsaViz는 저작 인터페이스로 그래프를 기반한 환경을 지원하는데 지원하는 기능으로는 확대, 축소 기능과 그래프 내에서의 노드 탐색 기능을 제공하며, 타원, 박스, 화살표를 이용하여 그래프를 생성하고 편집한다. 그리고 RDF/XML, N-Triple import 기능과 저작된 RDF 문서를 RDF/XML, N-Triple, SVG, PNG로 export하는 기능을 지원한다. 그러나 그래프 저작 인터페이스를 통한 저작 시 각 노드에서의 자원 기술과 편집이 쉽지 않기 때문에 RDF 문서의 저작하기 어렵다는 문제를 가지고 있다[11].

### 2.2 RDFedt

RDFedt는 복잡하고 구조화된 RDF 문서를 빠르고 쉬운 방법으로 만들 수 있는 방법을 제공하기 위해 개발되었다. 저작 환경으로 트리 기반 인터페이스를 지원하여 트리를 이용한 RDF 문서의 저작이 가능하다. RDFedt는 Element-tree를 이용하여 복합 데이터 구조에 대한 개요를 사용자에게 제공하며 다양한 Element Set을 지원한다. 그러나 RDFedt는 단순히 트리 인터페이스를 통한 저작 환경만을 지원하기 때문에 RDF의 자원간의 기술 구조나 자원간 관계를 가시적으로 표현이 불가능하다는 단점을 가지고 있다. 또한 RDF 문서에 대한 파싱만을 수행하므로 RDF 기술 구조만을 검사하고 유효성을 보장할 수 없다는 문제점을 가지고 있다.

### 2.3 RDFAuthor

RDFAuthor는 RDF 인스턴스 데이터 생성의 어려움을 좀 더 쉽게 해결하기 위해 디자인되었는데, 그래픽 인터페이스와 텍스트 인터페이스를 저작 인터페이스로 제공하여 RDF 문서를 저작할 수 있도록 한다. 구현을 위해서 Apache AXIS library, HPL Jena toolkit을 이용하였고, Apple사의 Cocoa 프레임워크상에서 개발되었다. 따라서 MacOS에서만 사용할 수 있는 소프트웨어이다. 제공하는 기능으로는 PDF, EPS, SVG로 RDF 문서를 export하는 기능과 RDF 질의 기능, RDF import 기능을 지원한다.

RDFAuthor는 그래프와 텍스트 저작 환경으로 문서의 저작 효율을 높였지만, Mac OS에서만 동작되기 때문에 크게 활용되고 있지 못하다. 또한 복잡한 자원간의 관계 표현에 제한적이다[12].

### 2.4 Protégé-2000

Protégé-2000은 국립의료도서관(National Library of Medicine), NSF(National Science Foundation), DARPA(Defense Advanced Research Projects Agency)의 후원을 받아 스탠포드 의과대학의 의료정보학과(Stanford University

School of Medicine, Stanford Medical Informatics)에서 개발되었다. Protege-2000은 온톨로지와 지식 기반을 수정하고 생성하기 위한 확장이 용이하며, 플랫폼 독립적인 환경을 지원하며, 저작을 위해 직관적이고 사용하기 쉬운 GUI를 제공한다. 필요할 때마다 데이터베이스 back-end를 로드하고, 기억장치에 캐싱하여 사용할 수 있다는 장점을 가지고 있다.

그러나 RDF를 이용한 자원의 기술과 표현에 중점을 두고 있다기 보다는 온톨로지와 지식 기반 저작에 목적을 두고 있다. 또한 RDF를 이용해 웹 자원을 기술하기에 다소 복잡한 입력 작업과 제약사항이 뒤따라서 사용이 어렵다는 문제를 가지고 있다[13].

### 3. 메타데이터 문서 검증기

메타데이터는 XML에 기반한 RDF를 이용하여 생성된다. 이는 곧 메타데이터를 XML에서 well-formedness와 유효성을 검사하는 것과 마찬가지로 검사할 수 있다는 것이다. 이를 통해 메타데이터의 유효성이 보장된다. 따라서 본 논문에서는 RDF에 기반하여 생성된 메타데이터 문서의 유효성과 well-formedness를 보장하기 위해 RDF 파서와 유효성 검증기를 설계하였다.

#### 3.1 개발 환경과 API

본 논문에서는 RDF 기반 메타데이터 문서 처리를 위해 개발 환경으로 Java 기반의 J2SE v1.4를 이용하였고, 파서(parser) 구현을 위한 API로 Xerces 2 Java Parser 2.0.2 Release와 JAXP(Java API for XML Processing) v1.2를 사용하였다[4]. RDF 모델을 처리하기 위해 Jena v1.4를 사용하였다[5].

#### 3.2 모듈의 구성 요소와 세부 구현

##### 3.2.1 RDF 파서(parser)

RDF 저작도구를 통해 RDF 문서를 작성하고 이 RDF 문서를 직접 웹 상에서 사용하기 위해서는 반드시 RDF 문서가 올바르게 작성되었는지를 검사해야 한다. 또한 RDF의 모든 표현 형식을 제공하기 위해서는 RDF 전용 파서(parser)를 통해 파싱(parsing)의 과정을 거쳐야 한다. 파싱은 RDF 문서의 구문 검사를 의미한다. 본 논문에서 파싱 과정은 세 가지 인터페이스인 텍스트, 트리, 그래프에서 저작되거나 저작중인 RDF 및 XML 문서에 대해 파싱을 하도록 하였다.

RDF 파서는 RDF 문서의 well-formedness 만을 검사하도록 구현하였고, DTD 및 Schema에 따른 유효성 검사는 RDF 문서 유효성 검증기(Validator)를 통해 수행하도록 하였다. 파싱 루틴은 RDF 문서가 문법에 맞게 구성되어 있는지를 검사하는 과정으로 파싱 루틴은 텍스트와 트리 인터페이스를 통해 편집되거나 편집 중인 RDF 문서에 대해 파

싱을 지원한다. 본 논문에서는 SAX를 기반으로 여러 문서에 대해 빠르게 처리할 수 있도록 구성하였다.

##### 3.2.2 RDF 문서 유효성 검증기(Validator)

RDF는 XML로부터 기초하며 자원을 기술하는데 있어서 의미(semantic)를 정의할 수 있도록 하고 있다. RDF에서는 동일 의미를 가진 여러 표현들 사이의 불명확성을 고려하기 위해 XML의 네임스페이스(namespace) 메커니즘을 적용하고 있다. 네임스페이스는 어휘집(vocabulary)을 이용하여 관리함으로써 특정한 속성 형식을 사용하는 여러 규정에서 다르게 나타나는 의미를 식별할 수 있도록 하고 있다. 또한 네임스페이스를 이용하여 기존에 메타데이터를 표현 하였던 Dublin Core를 처리할 수 있도록 하고 있다. RDF에서는 XML 네임스페이스로 "RDF"와 같이 "DC"라는 네임스페이스를 이용하여 Dublin Core에서 정의된 어휘들을 선언하고 있다.

따라서 이와 같이 다양한 정보를 포함하고 확장 가능한 RDF 구문의 유효성을 검증하기 위한 "RDF 유효성 검증기"를 설계하였다. 유효성 검증기는 RDF 문서를 생성하기 바로 전 단계에서 유효성을 검증하도록 구현한다.

RDF 파서가 RDF 문서의 well-formedness를 검사하는 것과 달리 RDF 유효성 검증기는 RDF 파서의 파싱 루틴을 거쳐 DTD와 Schema에 대한 유효성을 검사한다. (그림 1)은 유효성 검증 처리 과정에 대한 구성도이다. (그림 1)과 같이 유효성 검증 루틴은 prepararsing과 postparsing의 두 단계를 거치게 된다. prepararsing 단계에서는 well-formedness 만을 검사한 뒤 DTD 및 Schema의 위치 정보를 가져온다. post-parsing 단계에서는 유효성 검증을 위한 설정을 하고 실질적인 유효성 검증을 수행하게 된다.

(그림 1) 유효성 검증 루틴의 구성 및 흐름도

## 4. RDF 문서 생성을 위한 편집기와 N-Triple 생성기

### 4.1 RDF 문서 생성기와 N-Triple 생성기의 구성

메타정보 모듈은 RDF 파서, 텍스트 기반 메타데이터 편

집기, 트리기반 메타데이터 편집기, 그래프 기반 메타데이터 편집기, RDF 문서 유효성 검증기, N-Triple 생성기, RDF 문서 생성기로 구성하였다. RDF 메타정보와 N-Triple을 생성하고 웹 자원을 그래프로 표현하는 모듈의 구성도는 (그림 2)와 같다.

(그림 2) 메타정보 모듈의 구성도

#### 4.2 RDF 문서 생성을 위한 텍스트, 트리, 그래프 기반 메타데이터 편집기

##### 4.2.1 텍스트 기반(Text-based) 메타데이터 편집기

대부분의 편집기들은 한 가지 방식의 인터페이스만을 채용하였기 때문에 다른 유형의 편집 방식과 연동할 수 없었다. 이러한 문제로 인해 전체적인 RDF 자원의 분석이 비효율적으로 처리되었다. 이를 위해 본 논문에서는 텍스트 기반 편집기에서 그래프 기반 방식과 트리 기반 방식으로 쉽게 전환하여 메타데이터를 처리하도록 하였다.

(그림 3)은 기본이 되는 텍스트 기반 메타데이터 편집기의 구성도이다.

(그림 3) 텍스트 기반 메타데이터 편집기의 구성

##### 4.2.2 트리 기반(Tree-based) 메타데이터 편집기

많은 양의 정보는 그룹화하여 하위 주제로 분리하는 것이 효과적이다. XML에서는 이를 위해 계층 구조안에 정보를 그룹화한 뒤 트리 구조로 표현할 수 있다. 트리는 관련되어 있는 객체간의 부모/자식(parent/child) 관계나 형제/형제(sibling/sibling) 관계를 표현할 수 있다. 따라서 RDF에서 제안한 데이터 모형과 동일한 의미를 가지면서 RDF 메타데이터를 편집할 수 있는 또 다른 유형의 편집 방법을 적용할 수 있다. 따라서 본 논문에서는 이러한 트리 구조에 기반한 편집기를 설계 및 구현하였다.

트리 기반 메타데이터 편집기에서는 DOM(Document Object Model)을 사용하여 파싱하였다. 즉, DOM 파서를 이용

하여 RDF/XML 문서를 읽어 들이고 DOM Document를 생성한 뒤 트리 형태의 인터페이스로 변환하여 사용자에게 문서상의 내포관계를 구조적으로 보여줄 수 있도록 하였다. 추가적으로 RDF와 RDFS, 그리고 RDF에 삽입되어 사용되는 Dublin Core, RSS의 Element Set을 트리 기반 편집기에서 노드를 추가할 수 있도록 하였다.

##### 4.2.3 그래프 기반(Graph-based) 메타데이터 편집기

RDF는 가장 일반적인 XML 표현법 외에도 노드 앤 아크 다이어그램으로 표현할 수 있다. 노드 앤 아크 다이어그램에서 타원형으로 표현되는 노드가 자원을 표현하는 것이고, 아크는 이름이 부여된 특성을 표현한 것이다[6]. 문자열 리터럴을 표현한 노드는 직사각형으로 표현된다. 노드 앤 아크 다이어그램은 N-Triple의 Subject, Predicate, Object(Literal)를 구성요소로 가진다. 이러한 노드 앤 아크 다이어그램은 자원간의 관계를 분석하는데 용이하게 사용할 수 있다. 따라서 본 연구에서는 RDF 문서를 분석할 수 있는 노드 앤 아크 다이어그램을 그릴 수 있는 그래프 기반의 편집기 모듈을 설계하였다.

그래프 기반의 메타데이터 편집 모듈을 제공하기 위해서는 먼저 파싱의 과정을 거쳐야 한다. RDF 문서를 노드 앤 아크 다이어그램으로 변환하기 위해서는 RDF 문서의 파싱 요청에 의해 ARP(Another RDF Parser) 파서가 RDF 문서 전체를 파싱하게 된다. ARP 파서의 파싱으로 생성된 닷 파일(dot file)은 N-Triple 형태의 Subject, Predicate, Object(Literal)를 가지며 노드에 대한 포지션(position) 값은 없는 상태이다. 포지션 값은 Graphviz를 통해 노드 앤 아크 다이어그램의 각 노드의 포지션 값을 할당 받게 된다. 각각의 노드에 대한 포지션 값이 할당되면 Graph Painter가 RDF 노드 앤 아크 다이어그램을 생성한다. 그래프 기반 메타데이터 편집기의 구성 및 처리 흐름도는 (그림 4)와 같다. 또한 RDF 문서 외에 다른 형식으로 데이터를 상호 호환 가능하도록 W3C의 규격에 맞는 SVG, PNG, N-Triple의 Exporter를 추가하였다. (그림 5)는 RDF 문서를 파싱하여 노드 앤 아크 다이어그램을 그려주는 그래프 기반 메타데이터 편집기 인터페이스이다.

(그림 4) 그래프 기반 메타데이터 편집기의 구성 및 처리 흐름도

기가 어렵다. 이러한 점을 감안하여 RDF 문법에 대한 세부 지식이 없어도 RDF 문서를 생성할 수 있도록 하였다.

사용자에 의해 직접 작성된 N-Triple 파일이나 N-Triple 생성기를 통해 생성된 N-Triple 파일을 Import 하면, N-TripleReader에 의해 판독되어 메모리 상에서 RDF 문서를 만들기 위해 N-Triple 모델을 생성한다. RDF serializer는 생성된 N-Triple 모델에서 Subject, Predicate, Object(Literal)를 조합하여 RDF 문서로 생성해주도록 하였다.

### 5. 기능 비교

본 논문에서 구현한 RDF 문서 검증기와 N-Triple 생성기에서 지원하는 기능과 타어플리케이션들과의 비교를 통해 장점을 살펴보도록 한다.

<표 1>에 제시된 바와 같이 기존에 개발된 어플리케이션들의 기능을 대부분 포함한다. 특징적인 기능으로 그래프 편집기 상에 구현된 RDF 문서에 기술된 웹 자원의 추적(trace) 기능이 있다. 이 기능은 대개의 RDF 문서가 많은 웹 자원을 기술하기 때문에 특정 자원의 위치와 자원간의 관계를 파악하기 어려워 자원의 효율적인 관리가 힘들다는 문제를 해결하기 위해 개발되었다. 이 기능은 다른 RDF 메타데이터 관련 어플리케이션에서 지원이 미미한 상황이다. 또한 PNG와 SVG export 기능을 구현하여 그래프화된 웹 자원을 이미지로 표현할 수 있도록 하였다. 이 기능의 경우 IsaViz나 RDF Author에서도 지원하고 있다. 그러나 IsaViz는 많은 자원을 기술한 그래프에 대해 완벽히 SVG를 표현하지 못하는 문제를 가지고 있다.

<표 1> 어플리케이션간의 기능 비교

기능 어플리케이션	그래프 인터 페이스	트리 인터 페이스	자원 추적	N- Triple 생성	RDF 문서 유효성 검증	PNG· SVG export	RDF import & export
R·V·N·G (본 논문)	○	○	○	○	○	○	○
IsaViz	○	x	x	○	○	○	○
RDF Author	○	x	x	○	x	○	○
RDFedt	x	○	x	x	x	x	x
Protégé-2000	○	○	○	○	○	○	○

편집기에 관련된 기능으로서 다른 편집 모드로 저작 중이더라도 RDF 문서의 일관성을 지원하여 편집 모드에 따른 기술된 웹 자원의 불일치 문제를 해결하며, 3.2절에서 기술한 바와 같이 문서 자체의 유효성을 검사를 처리함으로써 유효한 웹 자원 기술 문서로 보장하고 있다. 편집 모드에 관련하여 IsaViz, RDF Author, RDFedt는 하나의 편집 모드만을 지원하여 저작 효율이 비교적 낮은 상황이다.

N-Triple 생성 기능은 IsaViz, RDF Author와 본 논문의 N-Triple 생성기 등에서 원활히 지원한다. 하지만 RDFedt는 단순한 편집만을 지원하기 때문에 n-triple을 생성하지 못한다는 문제를 가지고 있다.

(그림 5) 그래프 기반 메타데이터 편집기의 실행화면

### 4.3 N-Triple 처리를 위한 N-Triple 생성기와 N-Triple Reader

#### 4.3.1 N-Triple 생성기

RDF는 다른 RDF 구문에 대한 구문(Statements about Statements)을 작성하는데 사용될 수 있다. 이러한 고차원 구문(higher-order statements)은 RDF에서 표현하기 위해 모델화해야 한다. 웹 상의 자원들을 표현하기 위해서는 반드시 구체화된 구문(reified statement)이 필요하다. 그리고 트리플(triple)을 사용하여 고차원의 구문이 어떻게 구체화되는가를 설명할 수 있다.

실질적으로 웹 상의 복잡한 메타데이터를 생성하고 편집하기 위해서는 구체화 과정이 필요하고 이 과정을 위해서 RDF는 트리플을 사용하고 있다. 따라서 본 연구에서는 RDF 문서를 3-Triple로 변환하는 과정은 반드시 필요하고, 웹 상에서 사용 가능한 메타데이터를 분석하고 생성·편집하기 위해서는 이런 기능이 필요하다.

ARP 파서를 통해 파싱된 RDF 문서를 N-Triple 생성기를 통해 Subject, Predicate, Object(Literal)에 따라 N-Triple 테이블에 저장하게 된다. N-Triple 생성기를 통해 처리된 결과는 (그림 6)과 같다.

(그림 6) N-Triple 생성기 실행 화면

#### 4.3.2 N-Triple Reader

일반 사용자가 RDF를 사용하기 위해서는 RDF 문법을 숙지하여야 한다는 문제점이 있다. 전자상거래 분야와 같은 비 전문분야에서는 RDF를 사용하여 메타데이터를 표현하

별도로 Protégé-2000은 많은 서드파티들에서 개발한 플러그인(plug-in)을 통해 <표 1>의 기능을 모두 지원하고 있으며 현재 가장 진보된 온톨로지 저작툴로서 인식되고 있다.

### 6. 결론 및 향후 과제

자원 기술(Description)을 위한 구조로서 RDF를 이용할 경우 통합된 형식을 바탕으로 웹 자원이 기술되기 때문에 일관성 있게 관리될 수 있다. 그러나 RDF 문서를 처리하기 위한 모듈 환경은 아직까지 텍스트 인터페이스(Text Interface)나 콘솔(Console)에 의존적이기 때문에 효율이 매우 낮다. 또한 기존에는 RDF 문서가 유효한지에 대한 검증이 이루어지지 않았기 때문에 유효하지 않은 RDF 문서를 생성하고 배포할 수 있다는 문제점을 가지고 있었다. 이와 같은 효율과 유효성 문제를 해결하기 위하여 본 논문에서는 텍스트, 트리, 그래프를 통한 처리 환경과 RDF 문서의 유효성 검증을 위한 파서(parser)와 유효성 검증기(validator), RDF 문서 생성기를 설계 구현하였다.

본 논문에서 제안한 텍스트, 트리, 그래프, N-triple을 이용한 인터페이스는 처리 환경에 독립적으로 유효성 검증 및 RDF 문서 생성을 위한 통합된 모듈들로 구성하였으며, 모듈은 내부적으로 RDF를 처리하기 위해 커스터마이징(customizing)되어 있는데, XML에 기반한 모든 문서에 대해서도 파싱과 유효성 검증이 가능하다. 또한 기존에 개발되었던 RDF 처리 모듈에서 지원할 수 없었던 다양한 처리 환경과 유효성 검증 기능을 제공해줌으로써 RDF의 모든 표현 형식을 유연하게 지원하고, 유효한 RDF 문서를 편리하게 작성하여 웹 상에서 직접 활용 할 수 있도록 하였다.

현재 그래프 편집기 상에 기술된 메타데이터가 위치한 노드를 정확히 검색하지 못하여 방대한 웹 자원을 기술할 때에는 그 효율이 낮다. 따라서 향후 과제로서 그래프 편집기 상에서의 정확한 노드 검색을 통한 웹 자원 기술에 관한 연구를 진행할 예정이다.

### 참 고 문 헌

[1] W3C, Resource Description Framework(RDF) Model & Syntax Specification, <http://www.w3.org/RDF/TR/1999/REC-rdf-syntax-19990222/>, 1999.

[2] RDF Vocabulary Description Language 1.0 : RDF Schema, <http://www.w3.org/TR/2002/WD-rdf-schema-20020430/>, 2002.

[3] Jeen Broekstra, Michel Klein, and Stefan Decker, Enabling knowledge representation on the Web by extending RDF Schema, In Proceedings of the 10th World Wide Web conference, Hong Kong, China, pp.467-478, May, 2001.

[4] Apache XML Project, <http://xml.apache.org/xerces2-j/>, August, 2002.

[5] Hewlett-Packard, The jena semantic web toolkit, <http://www.hpl.hp.com/semweb/arp.html>, July, 2001.

[6] Olivier Corby, Rose Dieng and Cedric Hebert, A Conceptual

Graph Model for W3C Resource Description Framework, ICCS2000, Darmstadt, Germany, August, 2000.

[8] Decker et al., The Semantic Web : The Roles of XML and RDF, IEEE Internet Computing, 2000.

[9] RDF Issue Tracking, RDFCore Working Group, 2001.

[10] Dave Beckett, RDF/XML Syntax Specification(Revised), 2001.

[11] W3C, Resource Description Framework(RDF), <http://www.w3.org/2001/11/IsaViz/>, 2001.

[12] Damian Steer, RDFAuthor, <http://rdfweb.org/people/damian/RDFAuthor/>, 2003.

[13] Protégé-2000, <http://protege.semanticweb.org/>.

### 조 성 훈

e-mail : shcho@dblab.hannam.ac.kr  
 2001년 한남대학교 컴퓨터공학과(공학사)  
 2003년 한남대학교 컴퓨터공학과(공학석사)  
 2004년~현재 한남대학교 컴퓨터공학과  
 박사과정  
 관심분야 : RDF, 시맨틱 웹, 온톨로지,  
 ebXML, 웹 서비스

### 송 병 열

e-mail : sby@etri.re.kr  
 1995년 전북대학교 전자공학과(공학사)  
 1997년 전북대학교 전자공학과(공학석사)  
 1997년~현재 한국전자통신연구원 지능형로봇연구단/  
 지능형서비스플랫폼 연구원  
 관심분야 : 웹서비스, 시맨틱웹, 지능형로봇

### 조 현 규

e-mail : hkcho@etri.re.kr  
 1988년 한국외국어대학교(학사)  
 1990년 고려대학교 대학원(경영학석사)  
 1997년 한남대학교 대학원(경영학박사)  
 1990년~현재 한국전자통신연구원  
 책임연구원, 지능형로봇연구단/  
 지능형서비스플랫폼연구팀장  
 관심분야 : 시맨틱 웹, 온톨로지, 웹 서비스, 지능형 에이전트, 인공지능, 비즈니스 프로세스

### 최 의 인

e-mail : eichoi@dblab.hannam.ac.kr  
 1982년 승전대학교 계산통계학과(학사)  
 1984년 홍익대학교 전자계산학과(이학석사)  
 1995년 홍익대학교 전자계산학과(이학박사)  
 1985년~1988년 공군 교육사 전산실장  
 1992년~1996년 명지전문대학 전자계산과  
 조교수  
 1996년~현재 한남대학교 컴퓨터공학과 부교수  
 관심분야 : 실시간 데이터베이스, 주기억 데이터베이스, 클라이언트/서버 데이터베이스