

의학사서를 위한 시맨틱 웹의 소개

울산대학교 아산의학도서관

남 재 우(역)

목 적: 본 논문은 (1) 의학분야 사서들에게 시맨틱 웹(SW)의 주요개념과 원리를 소개하고 (2) SW기술을 이용하여 생의학정보를 운용하는 다수의 프로젝트를 개략적으로 기술한다.

방 법: 본 논문은 두개의 주요 부분으로 구성되어있다. “시맨틱 웹 테크놀로지”에서는 Extensible Markup Language (XML), Resource Description Framework (RDF), RDF Schema (RDFS), 온톨로지의 주요 표준과 개념에 대한 예제와 설명을 하고 정보검색에서의 그 효용성을 알아본다. 그리고 좀 더 발전된 SW언어와 그 특징에 대한 언급으로 단원을 종결한다. “생의학분야의 시맨틱 웹 어플리케이션과 연구 과제”에서는 Unified Medical Language System (UMLS), Generalised Architecture for Languages, Encyclopedias and Nomenclatures in Medicine (GALEN), HealthCyberMap, LinkBase, The National Cancer Institute (NCI)의 시소러스에 대해 간단히 살펴본다. 또한 이 논문은 이 SW로 인해 나타나는 다른 이점과 부작용으로 나타나는 것들을 알아보고 그와 관련된 프로젝트의 예를 든다.

토론과 결론: SW의 전망에 직면해있는 몇 가지 문제점이 나타났다. 특히 지식을 조직하고 정보를 구축하는 사서의 전문지식은 SW프로젝트에 공헌할 것이다.

서 론

웹은 정보의 개념을 변화시켰다. 웹의 크기와 양은 계속 증가하고 있고 무엇보다 구조적인 복잡성과 근본적인 구조에서 계속 변화하고 있다.

Extensible Markup Language (XML)과 Dublin Core (DC)와 같은 메타데이터 표준의 출현은 “시맨틱 웹”을 창조하기 위한 World Wide Web Consortium (W3C)의 주도로 정보가 컴퓨터에서 이해되고 읽힐 수 있는 웹기반 정보의 새로운 세계를 향하고 있다. Berners-Lee[1]의 논문에서 요약된 시맨틱 웹(SW)의 전망으로는 지

능형 검색 프로그램(software agents)이 웹기반 정보에 덧붙여진 메타데이터로부터 정교한 결론을 도출해낼 수 있다고 하였다. 이러한 전망과 기초기술은 아직 그 효과가 충분히 실현되지 않았지만 매우 큰 가능성을 보여주고 있다. 그렇기 때문에 현재 많은 연구가 진행 중이며 유럽과 미국에서 상당량의 예산이 배정되고 있다.

이 논문은 두 가지 목적을 갖고 있다. 첫째, 저자는 의학사서의 관점에서 SW 기술의 주요 개념에 대한 높은 수준의 포괄적 소개를 하고 의학정보과학에서 사용된 개념과 정확히 연계하려 한다. 둘째, 저자는 SW기술을 사용하고 있거나 사용이 계획되어 있는 의학정보의 운용을 포함하여, 몇 가지 어플리케이션과 진행 중인 프로젝트에 대해 간단히 알아본다. 저자는 정보과학자의 전문지식이 이 프로젝트에 공헌할 수 있는 방법이라고 역설한다. 의학사서들이 의학정보의 미래를 만들어가는 핵심역할을

이 논문은 Ioana Robu, Valentin Robu and Benoit Thirion이 공저하여 “An introduction to the Semantic Web for health sciences librarians”라는 제목으로 미국의학도서관협회지(J Med Libr Assoc) 94권 2호(2006년 4월) 198~205에 실린 내용이다.

계속할 것이고, 사서의 전문지식이 인공지능의 핵심 프로젝트에서 큰 자산이 될 것이라고 확신한다.

Semantic Web (SW) Technology

오늘날 대부분의 웹 콘텐츠는 사람들이 사용하기에 적절하다. 웹의 전형적인 사용은 정보를 탐색하고 소비하며, 다른 사람들을 찾아 관계를 맺는 것을 포함하고 있다. 이런 활동을 지원하는 소프트웨어는 그다지 잘 개발되지 않았는데, 사실상 주된 소프트웨어로는 검색엔진이 있다. 게다가 이런 소프트웨어들은 거의 발전이 없으며, 웹 콘텐츠는 기술적인 진보보다 더 빨리 성장하고 있다. 특히 정보검색은 제대로 지원되지 않는다. 당면한 주요 난관은 웹 콘텐츠의 의미를 컴퓨터가 인식할 수 없고[2], 컴퓨터의 지각(知覺)으로는 단어, 문장 그리고 그들의 관계를 해석할 수 없다는 것이다.

SW의 주요 목표는 현재의 웹에 논리를 더하는 것인데, 다시 말해서 컴퓨터가 정보에 접근할 수 있게끔 해주는 일련의 형식적 규칙을 통해서 데이터 의미, 대상의 특성, 그 사이에 존재하고 있는 복잡한 관계의 의미를 표현하는 것이다. 컴퓨터의 접근성은 정보가 나타나는 형태에서 벗어나, 데이터의 의미(시맨틱)를 바탕으로 쿼리(질의)를 만들 수 있게끔 정보를 표현하는 것으로 이해되어야 한다.

XML과 다른 표준 언어들은 이러한 목적을 달성하기 위해서 비롯되었으나, 이것은 단지 제한적인 방법일 뿐이다. 다음 3개의 XML tag를 살펴보자.

1. <article name="Duodenal ulcer">
 <author>Smith</author>
 </article>
2. <author name="Smith">
 <article>Duodenal ulcer</article>
 </author>
3. <bibRecord>

```
<author>Smith</author>
<article>Duodenal ulcer</article>
</bibRecord>
```

위의 모든 태그는 제목이 "Duodenal ulcer"이고 저자가 Smith라는 동일한 의미를 표현하고 있다. 그러나 컴퓨터에게 있어서 위의 세 표현은 서로 다르게 인식되는데, 이는 태그의 중첩된 순서가 다르기 때문이다. 위의 표현은 의미를 거의 가지고 있지 않다.

그렇지만 XML은 구조가 상대적으로 단순하고 정밀한 구문이나 형식이 강화된 많은 어플리케이션에는 충분하다. 이러한 표현의 수많은 예가 존재하고 있는데 가장 잘 알려진 것으로는 XML형태로 된 PubMed 레코드의 출력 일 것이다. 그러나 XML의 정확한 구조가 지켜지지 않는 한, 즉, National Library of Medicine (NLM)이 명세한 document type description (DTD)가 지켜지지 않으면 XML 코드는 유효하지 않다.

SW의 접근법은 XML표준의 상위층에 또 다른 층을 만들어서 코드화된 정보에 의미를 좀 더 부여하는 것이다. 이와 같은 특별한 층은 Resource Description Framework (RDF) / RDF Schema (RDFS)와 Web Ontology Language (OWL)과 같은 표준으로 표현된다.

SW는 대체가 아닌 "현재 웹의 확장"[1]으로 간주되어야 한다. 좀 더 발전된 "시맨틱" 지식 표현기술을 사용한 사이트들은 단순한 XML이나 HyperText Markup Language (HTML)을 사용해 구축한 사이트들과 계속 공존할 것이다.

Resource Description Framework (RDF)

RDF는 W3C <<http://www.w3c.org/RDF/>>의 표준으로 인정받은 기술(記述)언어로, 웹에서 표현되는 데이터의 논리를 포착하려는 데 목적이 있다. RDF는 XML을 기본표현으로 사용하지만 XML 구문에 개념적으로 의존하지 않는다는 점에 주의해야 한다. 좀 더 형식적으로 기술되

어진 RDF/RDFS는 XML과 추상화의 수준이 다르다.

RDF표준은 몇 개의 간단한 개념을 기초로 하고 있다. 첫 번째 개념은 “자원(resource)”이다. 자원은 도메인(논문, 책, 저자, electronic resource 등)에서 기술되는 기본적 대상이거나 “사물(thing)”이다. 모든 자원은 Universal Resource Identifier (URI)라는 고유한 식별자로 식별된다. 대부분의 어플리케이션에서 URI는 웹페이지의 Uniform Resource Locator (URL)나 웹페이지의 한 부분(anchor URL) 또는 웹서버에서 이용할 수 있는 문서의 링크이다. 그렇지만, URI는 웹 링크보다 좀 더 일반적인 개념이며 자원을 유일하게 식별한다. 예를 들면, 도서관에서 도서의 MARC 목록레코드는 URI로 생각할 수 있다.

두 번째 주요 개념은 “클래스(classes)인데 이것은 근본적으로 객체(object)¹⁾와 사물(thing)의 공통 특성에 대한 집합이다. 클래스와 클래스 또는 객체의 인스턴스(instance)²⁾를 구분하는 것은 매우 중요하다. 예를 들어, “article”이 클래스인 반면에 “Duodenal ulcer”는 “article”클래스의 객체 또는 인스턴스이다. 클래스는 더 나아가 하위클래스로 나누어질 수 있다. 예를 들면, “article”클래스는 “Printed Resource”의 하위클래스인데 “Printed Resource”는 좀 더 일반적인 클래스인 “Resource”의 하위클래스이다. (Resource는 printed와 electronic resources를 포함할 것이다.) 이는 정보학에서의 계층적 분류의 개념과 비슷하다.

세 번째 중요한 개념은 “특성(Property)”이다. 특성은 다른 클래스와 자원과의 관계를 기술한다. 예를 들면, “article” 클래스는 “has_author,” “appears_in_journal,” “has_publication_date” 등의 특성을 가질 수 있다. 유사성을 자의대로 취해

서는 안되겠지만, 이 관계는 어느 정도 패킷 분류에서의 패킷으로 볼 수 있다.

클래스와 속성을 정의했으면 “서술문(Statements)은 추론해 볼 수 있다. RDF의 서술문은 주어(subject)-술어(predicate)-목적어(object)의 기본 형태를 갖는다. 주어는 RDF의 자원이고, 술어는 RDF의 속성, 목적어는 또 다른 RDF의 특성이나 리터럴(이름, 숫자, 코드 등)이다. 위의 예를 사용해서, Duodenal ulcer has_author Smith,” 또는 “Duodenal ulcer has_publication_date 01.01.2001”과 같은 서술문을 만들 수 있다. 이 논문의 RDFS 용어에서는(아래 “RDF schema”를 보라) 주어는 특성의 영역이라 하고 목적어는 범위라 한다.

RDF Schema

RDF 스키마는 RDF와 함께 주어진 정보의 구조를 기술하기 위한 표준이고 여기에서 구조는 특정 도메인을 명확하게 나타낸 것이다. 이 표준은 단지 내용을 기술하기 위해 사용하는 DC같은 표준과는 본질적으로 다르다. DC에서 반드시 채워져 있어야 하는 필드는 dc:Creator, dc>Date 등과 같이 이미 채워져 있으며 이용자는 이 필드를 바꿀 수 없고 단순히 값을 할당할 수만 있다. 이와 마찬가지로 이용자는 목록레코드의 필드를 바꿀 수가 없다.

이와 대조적으로 RDFS는 RDF가 반드시 가져야 하는 구조를 규정하고 있는 표준이다. 즉, 클래스, 특성 그리고 그들의 관계는 RDF 세부 기술에 나타나도록 되어있다. 그리고 나서 RDF는 내용(사실상의 정보)을 상술하는데 사용한다. 위의 예로 돌아가면, RDFS에서 이용자는 다음과 같이 구조를 명시할 수 있다.

“article”클래스 has_author의 인스턴스는 “author”클래스의 인스턴스를 갖고 있거나 “article”클래스의 has_publication_date의 인스턴스는 “date”의 인스턴스를 갖고 있다. 그리고 나면 RDF에서는 이와 같은 구조를 “Duodenal ulcer has_author

1) 클래스에서 정의된 개념이나 존재.

2) 클래스에 의해 생성되는 객체의 한 부분으로 한 집합의 개별적 요소.

Smith” 또는 “Duodenal ulcer has_publication_date 01.01.2001.”와 같이 설명할 수 있다.

Ontologies

온톨로지의 개념은 철학에 뿌리를 두고 있다. 아리스토텔레스는 철학을 “the science of being qua being”으로 정의했는데, “qua”의 의미는 “~로서”이다[3]. 컴퓨터 과학에서 이 용어를 사용하면서 이 용어에 대한 많은 정의들이 생겨나게 되었다. Sowa[4]는 온톨로지가 특정 도메인에서 “존재하거나 존재할 수 있는 것들의 카테고리”이며, 그 도메인과 관련된 대상의 형식과 그 형식들 간의 관계를 세부적으로 기술한 목록을 만든다고 하였다. 이런 카테고리의 유형이 온톨로지이다.

다시 말해, 온톨로지는 주어진 도메인에서 철저하고 정밀한 개념 스키마(예: 개념들 간의 관계에 대한 지도)를 체계적으로 나타내려 한다. 가장 일반적으로 인용되는 정의는 Gruber의 “개념화에 대한 명시적 명세”[5]이다. Sabou는 “도메인의 공유된 개념화”[6]라는 문장으로 온톨로지에 대해 좀 더 명확하게 규정하였다. 여기서 “공유된”이라는 용어를 강조하였는데, 정보를 사용하는 모든 구성원은 서로 같은 표현을 사용해야 하기 때문이다. 예를 들어 서로 다른 사람이 카테고리화 하위카테고리로 된 어

면 시스템에서 아티클을 분류한다고 할 때, 카테고리라는 누구에게나 알려져 있고 일치되어 있어야 한다. 그림 1은 도서관 영역에서 가능한 온톨로지의 예를 보여준다.

그림 1에는 두 종류의 요소가 나타나는데 클래스는 직사각형으로 표현되었고 클래스의 특성은 타원형으로 표현되었다. 또한 RDF/RDFS에 의해 허용된 두 개의 주요 관계 형태에 따라 2가지 형태의 화살표가 있음에 주의해야 한다. 그림 1에서 끝부분이 두꺼운 화살표는 “subclass of” (또는 “is_a”) 관계이다.³⁾ 예를 들어 “book” 클래스의 어떤 인스턴스는 “Printed Resource”이고 “Resource”이기도 하다. 이것은 클래스 간의 계층구조와 일치한다.

관계의 두 번째 타입은 “영역(Domain)-속성(Property)-범위(Range)” 관계인데 그림 1에서 끝부분이 얇은 화살표를 따르고 있다. (예: resource는 author 또는 year의 인스턴스인 publication year을 갖는다). 상위 클래스에서 정의된 특성들은 그들의 모든 하위클래스에도 적용된다는 것을 명심하자. 예를 들면, 저자는 그림 1에서 모든 resource가 authors를 가질 수 있다고 명시했는데, 이것은 “Printed Resources”와 “Electronic Resources” 또한 author를 가지며, 따라서 book과 article 또한 그렇게 된다는 것을 의미한다. 그러나 특성 “hasJournal”은 오직 journal articles의 하위클래스에만 적용되고 적어도 이 온톨로지에서는 book 또는 electronic resources에 적용되지 않는다.

물론 이 예제는 매우 간단하다. 실제의 온톨로지들은(특히 생의학 영역에서) 수많은 클래스와 특성을 가지고 있을 것이다.

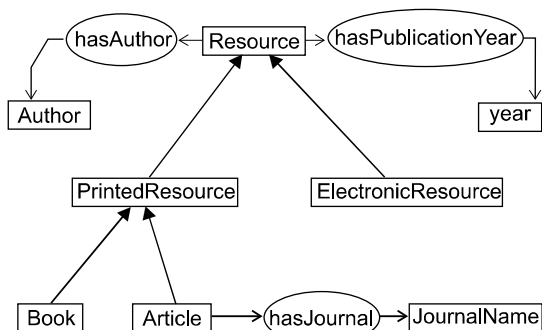


그림 1. 도서관 정보자원을 위한 온톨로지의 실례(實例).

3) is_a 관계는 “Printed Resource is a Resource”, “Electronic Resource is a Resource”와 같은 관계성을 나타내며 Resource 라는 클래스는 Printed Resource과 Electronic Resource의 공통 부모 클래스로서 Printed Resource와 Electronic Resource에게 공통적으로 있는 성질을 가지고 있다.

Extensible Markup Language (XML) Tag의 사용 예제

위의 세 섹션의 개념은 그림 1에 나타난 몇 개의 클래스와 특성에 상응하는 tag들을 통해서 예를 들 수 있다. 이 논문의 목적은 RDF나 RDFS 태그의 모든 구문을 제시하는 것이 아니지만, 위의 예를 좀 더 구체적으로 만드는 것은 중요하다.

먼저, RDFS는 구조를 표현하기 위해 사용된다는 것을 기억하자. 반면, RDF는 구조에 대한 실제 정보를 입력하기 위해 사용된다.

그림 1로부터의 온톨로지 구조를 잘 생각하라. RDFS 정의는 아래와 같은 태그를 포함한다. 가장 상위의 태그는 태그들이 사용된 XML의 도메인을 정의한다.

```
<rdf:RDF xmlns:rdf="http://www.w3c.org/.../rdf"
xmlns:rdfs="http://www.w3c.org/.../rdfs">
  <rdfs:Class rdf:about="Article">
    <rdfs:subClassOf rdf:resource="Printed-Resource">
  </rdfs:Class>
  <rdf:Property rdf:about="hasAuthor">
    <rdfs:domain rdf:resource="Resource">
    <rdfs:range rdf:resource="Author">
  </rdf:Property>
...</rdf:RDF>
```

일단 온톨로지 구조가 RDFS 태그에서 정의되어지면, 그것은 로컬서버의 "libStructure.rdfs" 파일에 저장된다. 이 파일에서 정의된 클래스와 속성들은, "lib"이라는 자체 정의된 도메인에 속하는 것으로 간주되어 다시 사용될 수 있다. RDF파일의 상부에는 이 정보의 참조가 포함될 것이다.

```
<rdf:RDF xmlns:rdf="http://www.w3c.org/.../rdf"
xmlns:lib="http://www.umfcluj.ro/.../libStructure.rdfs">
  <rdf:Description rdf:ID = "Duodenal ulcer">
```

```
<rdf:type>
  <rdfs:Class rdf:resource="Article">
</rdf:type>
<bib:hasJournal rdf:resource="Gastroenterology">
<bib:hasAuthor rdf:resource="Smith">
<bib:hasPublicationYear rdf:resource="2005">
</rdf:Description>
... </rdf:RDF>
```

정보검색에서의 효용

이와 같은 인코딩의 가장 중요한 이점은, 이들이 Resource Query Language (RQL)이라고 하는 전문 쿼리언어를 사용하여 좀 더 복잡한 정보검색을 가능하게 한다는 것이다. 그렇다면, 이 복잡한 정보검색이 단순한 XML로 가능할까? 적어도 간단한 방법으로는 아니다. 예를 들면 아래 두 개의 태그들과 같다.

```
1) <BibliographicResource>
  <Title>"Duodenal ulcer"</Title>
  <Author>Smith</Author>
  <PublicationYear>2001</PublicationYear>
</BibliographicResource>
2) <JournalArticle>
  <Title>"Treatment of duodenal ulcer"</Title>
  <Author>Smith</Author>
  <PublicationYear>2002</PublicationYear>
</JournalArticle>
```

첫 번째 태그 "BibliographicResource"는 어찌 보면 아직 발표되지 않은 임상보고서에 적용될 것이다. 쉬운 이해를 위해, 사용자가 이미 저자와 날짜가 XML 중첩의 두번째 수준에 있는 것을 알고 있다고 가정해보자. 이제, 쿼리는 저자가 Smith이고 2000년 이후에 출판된 모든 서지자원을 검색하기 위해 상위의 구조에서 만들어질 수 있다. 그리고 이것은 오직 Bibliographic Resource으로 명백히 표시된 첫번째 레코

드에게로 돌아갈 것이고 두번째 레코드에게는 아니다.

인간이 이해하기에 journal articles은 마찬가지로 정보의 서지자원이라는 것이 분명하다. 그러나 컴퓨터는 그 근본을 이루는 표현이(또는 온톨로지) 이런 정보를 포함하지 않는 한 이것을 알지 못한다. 그러므로 좀 더 의미론적으로 의미가 있는 표현들에게 이점이 있다. 게다가 좀 더 시맨틱한 표현을 사용할 때의 이점은 표현되는 정보의 크기와 복잡성이 증가한다는 것이다. 수많은 클래스와 특성이 있는 생의학영역에서의 온톨로지 사용은 종종 이와 같은 복잡성을 수반한다.

이 모든 구조, 규칙, 표준의 핵심 개념은 형식적인 논리를 강화하고, 컴퓨터의 정확한 해석에 의미를 붙이는 것이다.

RDF를 넘어서: 더 진보한 SW언어

RDF, 그리고 그와 관련된 RDFS는 “시맨틱”이라 부를 수 있는 가장 단순한 웹 언어이다. RDF는 이미 W3C에 의해 표준으로 인정받았고 가장 실제적인 환경에서 만나게 될 것이다.

RDF는 RDF로 표현될 수 있는 구조 형식에 대해 한계를 갖고 있다. SW이 포함된 좀 더 강력한 언어는 다음과 같다. Defense Advanced Research Projects Agency (DARPA) Agent Markup Language (DAML)는 미국의 DARRA의 지원으로 개발되었고, Ontology Inference Layer (OIL)은 유럽 암스테르담의 Artificial Intelligence Department of the Free University (VU)에서 개발되었다. RDFS보다 매우 강력한 이들 두 언어는 후에 통합되었고 Web Ontology Language (OWL)라고 하는 경량버전을 W3C 표준으로 제의하였다. 이들은 또한 Stanford Medical Infor-

matics에서 개발된 Protégé⁴⁾와 같은 온톨로지 언어에 “쓰기(write)”를 위한 프로그램이 존재하며, 생의학 온톨로지의 국가자원으로 NLM의 지원을 받는 지식베이스(knowledgebases)이다. RDF/RDFS에서는 명세될 수 없지만 OWL에서는 가능한 표현 유형의 예는 다음과 같다.

- 카디널리티(cardinality)⁵⁾ 제약: 예를 들어 사용자는 어떤 프린트 자원의 저자는 한 명 이상일 수 있으나 출판연도는 오직 하나라고 말할 수도 있다.

- 분류의 혼란스러움: 예를 들면 사용자는 어떤 자원의 인쇄나 전자 어느 쪽으로든 분류된다는 점을 강조할 수 있다.

- “유일성”이나 “부분적 범위”와 같은 속성의 다른 특성: “유일성”은 책은 단지 한 개의 International Standard Book Number (ISBN)을 가질 수 있다는 것이다.; “반대속성”의 예를 들면, 책에 적용된 “is published by”는 출판사에 적용된 “publisher of”의 속성과 반대이다.; “부분적 범위”는 속성의 도메인을 어떤 분류에 제한하는 것이다.

OWL이나 DAML+OIL을 설명하고 예를 드는 것은 이와 같은 소개성(introductory) 논문의 범위를 넘어서는 것이다. 관심 있는 독자는 Antoniou와 Van Harmelen의 저서인 Semantic Web Primer[2]를 참고해야 할 것이다.

생의학 분야에서의 SW 어플리케이션과 연구프로젝트

SW의 본질적인 개념은 적어도 지금, 그것이 명확하게 무엇이고 어떻게 작동하는지를 알 수 있게 해주는 웹사이트의 예를 드는 것을 매우 어렵게 한다[7]. SW의 실제 어플리케이션은 현실적 융합과는 거리가 멀고 계속 연구하는 수준이다. 그리고 실제에서 보다 개념에서 훨씬 더 가치가 있다고 판명될지도 모른다는 약간의 우려가 있다. 그러나 생의학과 보건학에서 존재하는 많은 어플리케이션과 연구프로젝트는

4) Java 기반의 open source 통합 온톨로지 구축 framework.
5) 객체들 간의 관계를 표현할 때 종속객체가 가질 수 있는 튜플(tuple :릴레이션을 구성하는 각각의 행 속성의 모임)의 개수.

SW기술을 기반으로 하고 있거나 중요한 범위에까지 SW기술과 결합되어 있다. 이 장에서는 독자들에게 어플리케이션에서 중요한 이슈가 무엇인지를 알려주기 위해 다수의 프로젝트에 대해 간단히 살펴보는 것이 목적이다. 생의학 어휘, 용어, 분류의 정보검색에서의 결정적인 역할을 고려해 볼 때[8], 이 들 대부분의 프로젝트들은 그들의 형식화와 관련이 있다. 즉 포함된 내용들의 클래스를 정의하고, 특히 그들 간의 관계를 정리하여 온톨로지로 발전시키는 것이다.

1. The Unified Medical Language System

NLM에 의해 개발된 Unified Medical Language System (UMLS)는 생의학 분야의 가장 많은 자원을 통합하고 있으며 Metathesaurus, Semantic Network, Specialist Lexicon 그리고 사전 프로그램들을 포함하고 있다. 특히 UMLS는 Semantic Network의 135개 시맨틱 형식 <http://www.nlm.nih.gov/research/umls/META3_current_semantic_types.html>으로 인해 온톨로지의 개념과 가장 가까운데, 이것은 백 개 이상의 어휘와 용어에서 추출된 백만 개 이상의 개념을 포괄하는 수준 높은 카테고리이다. 이 카테고리에는 두 개의 주 계층이 있으며, 한 개는 “Entity” 계층(“물리적인 사물”과 “개념적인 존재”를 포함)이고 다른 한개는 “Event” 계층(“활동”과 “현상이나 처리”를 포함)이다. 직접적으로 계층적인 is_a 관계 외에 의미론적인 형태 내에서 6,700개 이상의 계층적이고 연합적인 관계를 나타내기 위한 53종류의 정의된 관계가 사용된다. UMLS는 계속해서 개발 중이며 적용범위, 내용, 조직화 그리고 다른 온톨로지와의 호환성에 관한 연구가 진행 중이다[9-12].

온톨로지로서 UMLS는 다른 제작물이나 프로젝트를 개발하는 데 중요한 역할을 한다. 그 중에서 The Semantic Knowledge Representation (SKR) Project는 “현재 도서관에서 현재 이용할

수 있는 자원들을 구축함으로써 생의학 free text의 시맨틱 표현을 제공하는 프로그램(UMLS 지식 자원과 같은)[13].[†]을 개발하는 것”이 목적이다. 이 프로젝트의 일환으로 개발된 Medical Text Indexer[14]는 Medical Subject Headings (MeSH)에서 색인을 추천해주는 프로그램으로서 반자동 색인과 완전 자동색인에 사용될 수 있다(2002년부터 NLM에서 사용되고 있다).

2. Generalised Architecture for Languages, Encyclopaedias and Nomenclatures in Medicine

Generalised Architecture for Languages, Encyclopaedias and Nomenclatures in Medicine (GALEN): <<http://www.opengalen.org>>은 임상시스템에 용어자원을 제공하기위한 유럽연합의 프로젝트이다. GALEN은 기존의 시소러스나 분류법에 기초를 두고 있지 않기 때문에 처음부터 형식화가 가능했고 용어는 그 시스템에 포함되면 조정된다. GALEN의 웹 사이트에서는 “임상의들의 일상 업무를 지원하기 위한 유용하고 편리한 임상 어플리케이션의 구축을 좀 더 쉽게 하고자 한다.”라고 설명하고 있다. “임상용어집”인 GALEN Common Reference Model은 “분별력 있는 의학 개념들”과 자동으로 분류될 수 있는 카테고리들을 나타내는 데 목적이 있다[15]. GALEN은 또한 자연언어 처리 기술뿐만 아니라 기존의 코딩과 분류를 사용한다[16]. 지원하는 언어는 영어, 불어, 이탈리아어, 네덜란드어, 독어, 폴란드어, 스웨덴어이다.

3. The HealthCyberMap

HealthCyberMap <<http://healthcybermap.semanticWeb.org>>은 “검색과 탐색이 향상된 새로운 시맨틱 방법으로 사이버 공간에서 보건정보자원 분야의 엄선된 지도가 되는 것을 목적으로 하고 있다”[17, 18]. HealthCyberMap은 “주제” 요소에서 UMLS 용어를 포함하고 있는 DC 메타데이터를 사용한다.

4. LinkBase

LinkBase <<http://www.landcglobal.com/pages/linkbase.php>>는 “세계에서 가장 큰 공식 의학 온톨로지이고, 개념적으로 컴퓨터가 이해할 수 있는 의학의 표현”이 될 것이라고 단언한다. 상업적 용도로 Language and Computing 모델링 팀에 의해 개발되고 관리되었다가 후에 유럽연합으로부터 대규모 연구자금 지원을 받게 된다. LinkBase는 백만 개 이상의 언어독립적(language-independent)⁶⁾ 의학개념과 일반개념을 포함하고 표준용어와 자연어 표현이 가능하다. 의학개념은 언어독립적이고 약 3백만 개의 다양한 언어(영어, 불어, 스페인어 등)로 된 용어와 연결되어 있다. 개념들은 형식적인 관계를 표현하기 위해 약 450개의 다른 링크형식을 사용한 시맨틱 네트워크를 통해서 서로 연결되어 있다[19]. First DataBank, Healthgate 그리고 WebMD는 LinkBase의 온톨로지를 사용하고 있다.

5. 국립암학회(National Cancer Institute)의 시소러스와 온톨로지

이 온톨로지는 NCI가 만든 논리기반 용어의 공공 도메인에 대한 기술(記述)이고, 생물정보 caCORE의 배포를 위한 NCI센터의 구성요소로서 배포되었다[20]. 이것은 UMLS 메타시소러스에 기반을 두고 있고, 공용버전은 <http://ncimeta.nci.nih.gov>에서 이용할 수 있다. NCI의 시소러스를 OWL-Lite로 변환시키기 위한 연구가 진행 중이고 2005년 9월에, 이 ontology는 500,000 개 이상의 트리플⁷⁾을 포함했다(다운로드 버전 <http://www.mindswap.org/2003/CancerOntology>).

이 SW이 주장하는 또 다른 중요한 이점은 어떻게 해답이 도출되었는지에 대한 증명이다.

즉, 질의(querying) 어플리케이션이 사실을 신뢰할만한 어떤 근거를 제시할 수 있다는 사실이다. 적어도 도출된 사실들은 자원 덕분이라고 할 수 있고 시간에 걸쳐 완전함, 신뢰성 등에 대하여 자원을 평가하는 어플리케이션들이 개발되었다. 이것은 특히 의학분야에서 중요한데 의학분야에서의 정보는 신뢰의 가장 높은 수준이 보장되어야만 하기 때문이다[21].

6. MedCIRCLE

웹사이트에 따르면 MedCIRCLE은 다음과 같다.

MedCIRCLE <<http://www.medcircle.org>>은 유럽 보건 주제의 게이트웨이 또는 평가서비스의 협력으로, MedCERTAIN 프로젝트 내에서 안내되는 인터넷 보건정보를 평가하는 작업을 위해 설립, 확장되어 지속되고 있다. MedCIRCLE과 MedCERTAIN 두 프로젝트는 네트워크로 연결된 보건정보에서 신뢰받는 세계적 웹을 설립하고 이용자를 인터넷상의 신뢰할 수 있는 보건정보로 안내 할 수 있는 기술을 개발하여 활성화하기 위한 종합적 목적의 상호보완적 시맨틱 웹 프로젝트이다. 그리고 이용자가 웹상의 고품질 보건정보를 선별하여 분명하게 선택할 수 있도록 하기 위한 목적이 있다.

MedCIRCLE은 Health Information Disclosure, Description and Evaluation Language (HIDDEL) 어휘를 기초로, 우수하고 신뢰할 수 있는 용어가 바탕이 된 보건정보를 기재하고 있다[22, 23]. Catalog and Index of French-language Medical Resources(CISMef)는 이 프로젝트와 협력관계이다[24].

이미 프로토타입(prototype)⁸⁾을 넘어선 또 다른 프로젝트가 Joubert 등[25]에 의해 보고되었다. 그리고 그 논문에서는 보건영역에서 증명된 고품질 정보로의 접근을 제공하는 Web portal의 설계와 실행을 언급하고 있다. 프로젝트는 UMLS와 ARIANE 프로젝트를 기초로 하

6) 특정 언어에 구애받지 않음.

7) 주어-술어-목적어(subject-predicate-object)로 이루어진 구조로 RDF에서 사용되는 서술문(statements)의 표현

8) 정식버전이 나오기 전의 테스트버전.

고 있고 시맨틱 관계는 PubMed, Health on the Net (HON) databases, CISMef 또는 다른 공인된 자원에서 실행될 수 있는 자동 번역된 쿼리를 사용한다.

구조화되지 않은 데이터를 다룬다는 연구의 다른 분야는 온톨로지와 분류를 자동으로 구축하거나 통계적 클러스터링⁹⁾과 같은 자동기술을 통해 이미 존재하고 있는 온톨로지를 강화하는데 목적이 있다[16, 26]. 그렇지만, 저자는 일반적으로 텍스트 마이닝¹⁰⁾, 자연어 처리와 자동 정보 검색을 위한 다른 알고리즘(예를 들면 Google에서 사용된 것들)이 전통적으로 간주되어온 SW기술이 아니라고 지적하고 있다(그들이 정확히 형식논리에 의거하지 않기 때문이다). 그러나, 최근의 많은 연구는 Free text와 부분적으로 구조화된 자원으로부터 시맨틱 표현을 추출하는 데 초점을 맞추고 있다.

(예: The Semantic Knowledge Representation Project Website[13]).

또한, 의학교육자원[27]이나 의학 시청각 자원[28]에 관한 시맨틱 색인은 이미 보고되었고, 또 다른 SW는 영어외의 다른 어휘 자원을 구축하는 데 주력하고 있다[29, 30].

SW의 의구심

SW는 지식을 통제하고자 하는 첫번째 시도가 아니다[31]. 보편주의자들(Universalist)의 열망은 그 전에도 존재했었다[32]. (예를 들면, universal bibliographic control은 1970년대에 문헌정보협회에서 개발되었다). 비록 프로토콜과 표준이 형식적이고 상징적인 인공지능에 견고한 기반을 가지고 있지만, 지금까지는 인간의 마음이 너무 복잡해서 모델로 하기에는 매우 어

렵다고 입증되어왔다. 일반화, 모호성 그리고 퍼지의 관계는 어떤 분류 스키마에서도 본래부터 문제가 있었고 세상은 항상 이러한 속성의 용어로 기술될 것이다. Nelson[10]은 “시의 아름다움은 과학의 절망이다.”라는 인용으로 이것을 표현했다. 다른 한편, Berners-Lee는 “시의 아름다움”을 “사람이 중얼거리는 것”[21]이라고 보았다.

전통적인 SW는 일반적으로 컴퓨터의 정확한 해석을 넘어 정보의 의미를 포착하기 위한 형식적 논리를 사용함으로써 백 퍼센트의 정확도를 목표로 접근한다. 그러나 협정된 온톨로지에 따라서 이미 충분히 구조화된 정보를 제공하는 것은 매우 노동 집약형의 작업일 수 있다. 사실 대부분의 존재하는 정보는 비구조화되었거나(예: free text, HTML) 얼마간 구조화된(예: MeSH) 형태이다. NLM의 Semantic Knowledge Representation Project와 같은 몇몇의 프로젝트는 기존의 자원을 사용하면서, 생의학 free text에 근거한 의미론적인 표현을 자동적으로 만드는 프로그램을 구축함으로써 이 차이를 메우는 데 목적을 두고 있다.

SW 지지자들이 봉착한 또 다른 문제는 웹과 같은 열린 환경에서 정보를 표현하기 위한 온톨로지가 대체로 서로 일치하지 않는다는 사실이다(어떤 온톨로지의 “date” 필드는 다른 온톨로지에서는 “year”, “month”, “day”로 표현될지도 모른다). 이 문제는 온톨로지 통합이라고 하는 일련의 연구를 통하여 다루어진다.

그러나 SW 비전[1]의 최초 열망 중 일부는 아직 구체화되지 않았고, 여전히 많은 문제가 존재한다. 그럼에도 불구하고 SW기술은 생의학 영역과 같이 표현하고 검색하기가 복잡한 정보구조의 어플리케이션에서 매우 성공적으로 입증되었다.

사서들은 SW로 무엇을 해야하나?

통제어휘, 분류, 분류시스템, 정보검색, 양질

9) 인공지능 기법에서 통계적인 방법을 사용해 대량의 정보로부터 의미 있는 패턴을 자동적으로 분석하는 기술.

10) 비정형적 텍스트에서 여러 가지 기술을 적용하여 유용한 정보를 추출, 가공하는 것.

의 게이트웨이, 메타데이터 그리고 지식 관리의 사서직에 없어서는 안 될 요소를 구성한다. 심지어 현재 SW 연구가 직면하고 있는 많은 근본적인 문제들이 웹에서 발생하기 오래 전부터 사서들에게 수년간 연구되었다고 말할 수도 있다.

예를 들면, 패싯분류는(1930년대에 Ranganathan에 의해 최초로 제안되었다). 정보 자원의 구별과 검색을 위한 다양하고 신축성 있는 방법들을 허용하기 위해 클래스나 패싯과 같은 개념을 사용하였다. 이 접근은 RDF/RDFS에서 사용된 다양한 클래스와 속성과 비슷하다(비록 RDFS는 형식적인 논리에 근거한다는 명확한 이점이 있지만). 패싯접근에 많은 관심이 다시금 생겨나고 있는데, 그 이점이 웹 환경에서 검토되고 있다. 최근 보고된 미국의학협회[33]의 “간결한 의학 분류(concise medical taxonomy)”는 좀 더 정교하고 웹 친화적인 표현을 만들어 내기 위해서 패싯색인의 능력을 연구하였다. 도서관 커뮤니티에 알려진 바와 같이 패싯분류는 결코 실생활에서 떠나지 않을 것이다. 떠나지 않을 것이라는 주요 이유 중 하나는 패싯분류가 계층적 접근보다 인간의 논리에 덜 호소했다는 것이다. 오늘날 사용된 분류의 대부분은 계층구조이므로 개념적으로 더 단순하다. 이론의 여지는 있지만, 계층은 더 적은 표현성과 유연성을 제공한다. 그러나 표준이 더 이해하기 쉽고, 모든 부분에 한결같이 적용하기 쉽다는 주요 이점이 있다.

사서와 도서관 활동은 실제적인 수준에서 SW에 다양한 방법으로 공헌할 수 있다[7, 34, 35].

- DC와 같은 메타데이터 표준과 그들이 표현될 수 있는 다양한 방법(HTML, XML, RDF)을 잘 알고 있어야.
- 온톨로지의 개념과 XML이나 RDF를 사용한 정보표현에 대해 좀 더 학습하라.
- SW기술을 사용하는 프로젝트에 참여하라; 사서의 전문지식은 높게 평가될 것이다.

저자는 의학도서관과 SW커뮤니티가 서로 배울 점이 많다고 말하고 있다. SW접근은 정보 검색과 정보표현의 새로운 방법을 마련했다. 이런 접근은 형식적인 인공지능에 바탕을 두고 있는데, 인공지능에서 사용된 표현은 논리적으로 일치되고 그 의미 또한 컴퓨터가 접근할 수 있는 것을 보장하고 있다.

그럼에도 불구하고, 이 과정에서 의학사서의 중요성은 감소하지 않을 것이고 오히려 그 역할이 보다 더 중대하게 될 것이다. 다양한 온톨로지와 의학분류법을 개발하는 것은 사서들의 주요한 참여 없이는 어떤 유용한 실제 적용도 이끌어낼 수 없다. Soergel[36]은 현존하는 온톨로지 중에 가장 크고 유명한 WorldNet을 “온톨로지 구축에 시소러스의 구축 경험과 분류의 표준 방법을 적용함으로써 이점을 취한 훌륭한 시스템”이라고 말했다. 그는 또한 “중종 분류과학으로부터의 기존의 경험이 사용되지 않기 때문에, 크고 유용한 시스템은 필요 이상의 많은 노력으로 구축되고 있다.”라고 지적하였다. 저자들은 지금까지는 거의 평행선이었던 연구의 두 방향이 한데 모아지고 있다고 논하고 있다. 그리고 진행 중인 많은 연구 프로젝트가 이 저자들의 가정을 확인해 주고 있다.

이 논문은 기술적인 상세 설명을 떠나서 단지 가장 기본적인 개념과 요약을 다루고 있다. 이러한 기술들을 의학 도서관 환경에서 사용하기 전에 해결되어야 할 과제들이 있다. 그 과제들을 탐구하기 위해서는 그 이상의 훨씬 많은 연구가 필요하다.

Footnotes

*Based on an invited tutorial lecture delivered at the Ninth European Association for Health Information and Libraries Conference in Santander, Spain, September 2004. The original course is available in two parts at <http://www.chu-rouen.fr/>

documed/eahilsantander.html and <http://www.chu-rouen.fr/documed/semanticWeb.html>, respectively.

† See also the report from the National Library of Medicine at <http://lhncbc.nlm.nih.gov/lhc/docs/reports/2003/tr2003004.pdf>.

참 고 문 헌

- 1) Berners-Lee T, Hendler J, Lassila O. The Semantic Web. a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American* [serial online]. 2001 May 17;284(5). [rev. 15 Jul 2004; cited 25 Apr 2005]. <<http://www.sciam.com/article.cfm?articleID500048144-10D2-1C70-84A9809EC588EF21>>.
- 2) Antoniou G, Van Harmelen F. A Semantic Web primer. Cambridge, MA: MIT Press, Apr 2004. (Preliminary version available at: <<http://www.ics.forth.gr/isl/swprimer/>>. [rev. 3 Jul 2004; cited 25 Apr 2005].)
- 3) Ontology. [Web document]. WordIQ Dictionary and Encyclopedia. [rev. 13 Jul 2004; cited 25 Apr 2005]. <<http://www.wordiq.com/definition/Ontology>>.
- 4) Sowa JF. Knowledge representation: logical, philosophical, and computational foundations. Pontiac Grove, CA: Brooks Cole Publishing Co, 2000. Quoted in: Jacob EK. Ontologies and the Semantic Web. *Bull Am Soc Inform Sci Technol* 2003 Apr/May;29(4):19-22.
- 5) Gruber TR. A translation approach to portable ontology specifications. Technical report KSL 92-71. Sep 1992; revised Apr 1993. [Web document]. Knowledge Systems Laboratory. [cited 27 Nov 2005]. <<http://tomgruber.org/writing/ontologia-kaj-1993.pdf>>.
- 6) Web-based knowledge representation [course slides]. Free University of Amsterdam, 2002.
- 7) Palmer SB. Does it work? What Semantic Web applications are there? In: Palmer SB, ed. *The Semantic Web: an introduction*. [Web document]. 2001-09. [rev. 8 Feb 2005; cited 25 Apr 2005]. <<http://infomesh.net/2001/swintro/>>.
- 8) Lorence DP, Spink A. Semantics and the medical Web: a review of the barriers and breakthroughs in effective healthcare query. *Health Info Libr J* 2004 Jun;21(2):109-16.
- 9) Bodenreider O, Burgun A. Biomedical ontologies. [Web document]. In: Chen H, Fuller S, Hersh WR, Friedman C, eds. *Medical informatics: advances in knowledge management and data mining in biomedicine*. Springer-Verlag, 2005 (in press). (<<http://mor.nlm.nih.gov/pubs/pdf/2005-chaptermedont-ob.pdf>>. [rev. 18 Apr; cited 25 Apr 2005].)
- 10) Nelson JS. MeSH, UMLS, and the Semantic Web. [Web document]. 1/13/02. [rev. 13 Aug 2004; cited 25 Apr 2005]. <<http://www.nlm.nih.gov/mesh/presentations/taiwan2001/semanticweb/index.htm>>.
- 11) Kashyap V. The UMLS Semantic Network and the Semantic Web. [Web document]. In: *AMIA 2003 Symposium Proceedings*. 2003 Nov;351-5. [rev. 18 Aug 2004; cited 25 Apr 2005]. <<http://lhncbc.nlm.nih.gov/lhc/docs/published/2003/pub2003061.pdf>>.
- 12) The future of the UMLS Semantic Network: a workshop organized on April 7-8, 2005 at the US National Library of Medicine (NLM) in Bethesda, Maryland. [Web document]. [cited 27 Nov 2005]. <<http://mor.nlm.nih.gov/snw/>>.
- 13) National Library of Medicine. Semantic knowledge representation project. [Web document]. The Library. [rev. 1 Sep 2005; cited 14 Sep 2005]. <<http://skr.nlm.nih.gov>>.
- 14) Aronson AR, Mork JG, Gay CW, Humphrey SM, Rogers WJ. The NLM indexing initiative's medical text indexer. [Webdocument]. In: Fieschi M, Cuiera E, Y-CJ Li, eds. *MEDINFO 2004: proceedings of the 11th World Congress on Medical Informatics*. Amsterdam, The Netherlands: IOS Press, 2004. [rev. 1 Sep 2005; cited 14 Sep 2005]. <<http://cmbi.bjmu.edu.cn/news/report/2004/medinfo2004/pdf/papers/5122Aronson.pdf>>.
- 15) Rector AL, Rogers JE, Pole P. The GALEN high level ontology. [Web document]. In: *Fourteenth International Congress of the European Federation for Medical Informatics, MIE-96*; Copenhagen, Denmark; 1996. [rev. 24 Mar 2005; cited 25 Apr 2005]. <<http://www.opengalen.org/technology/technology.html>>.
- 16) Biczak DO, Amaral M, Roberts A, Rector AL. NLP techniques associated with the OpenGalen ontology for semi-automatic textual extraction of medical knowledge: abstracting and mapping equivalent linguistic and logical constructs. *AMIA Symposium 2000*. [Web document]. [5 pp.]. [rev. 24 Aug 2004; cited 25 Apr 2005]. <<http://www.amia.org/pubs/symposia/D200078.PDF>>.
- 17) Boulos MN. A first look at HealthCyberMap medical semantic subject search engine. *Technol Health Care* 2004; 12(1):33-41
- 18) Boulos MN, Roudsari AV, Carson ER. Towards a semantic medical Web: HealthCyberMap's tool for building an RDF metadata base of health information resources based on the Qualified Dublin Core Metadata Set. *Med Sci Monit* 2002 Jul;8(7):MT124-36.
- 19) Van Buggenhout C, Ceusters W. A novel view on information content of concepts in a large ontology and a view on the structure and the quality of the ontology. *Int J Med Inform* 2005 Mar;74(2-4):125-32.
- 20) Golbeck J, Fragoso G, Hartel F, Hendler J, Oberthaler J, Parsia B. The National Cancer Institute's thesaurus and ontology. *J Web Semantics* [serial online]. 2003 Dec; 1(1)

- [32 pp.]. [rev. 6 Feb 2005; cited 25 Apr 2005]. <<http://www.Websemanticsjournal.org/ps/pub/2004-6>>.
- 21) Berners-Lee T, Hendler J. Scientific publishing on the Semantic Web. *Nature* [serial online]. 2001 Apr;26; 410 (6832):1023-4. [rev. 8 Jul 2004; cited 25 Apr 2005]. <<http://www.nature.com/nature/debates/e-access/Articles/bernerslee.htm>>.
 - 22) Mayer MA, Darmoni SJ, Fiene M, Kohler C, Rothberghofer TR, Eysenbach G. MedCIRCLE: collaboration for Internet rating, certification, labelling and evaluation of health information on the World-Wide-Web. *Stud Health Technol Inform* 2003;95:667-72.
 - 23) Eysenbach G. An ontology of quality initiatives and a model for decentralized, collaborative quality management on the (semantic) World-Wide-Web. *J Med Internet Res* 2001 Oct-Dec;3(4):E34.
 - 24) Thirion B, Loosli G, Douyere M, Darmoni SJ. Metadata element set in a quality-controlled subject gateway: a step to a health Semantic Web. *Stud Health Technol Inform* 2003;95:707-12.
 - 25) Joubert M, Dufour JC, Aymard S, Falco L, Fieschi M. Designing and implementing health data and information providers. *Int J Med Inform* 2005 Mar;74(2-4):133-40.
 - 26) Soualmia LF, Darmoni SJ. Combining different standards and different approaches for health information retrieval in a quality-controlled gateway. *Int J Med Inform* 2005 Mar;74(2-4):141-50.
 - 27) Pouliquen B, le Duff F, Delamarre D, Cuggia M, Mouglin F, le Beux P. Managing educational resource in medicine: system design and integration. *Int J Med Inform* 2005 Mar;74(2-4):201-7.
 - 28) Cuggia M, Mouglin F, le Beux P. Indexing method of digital audiovisual medical resources with Semantic Web integration. *Int J Med Inform* 2005 Mar;74(2-4):169-77.
 - 29) Zweigenbaum P, Baud R, Burgun A, Namer F, Jarrousse E, Grabar N, Ruch P, le Duff F, Forget JF, Douyere M, Darmoni S. Umlf: a unified medical lexicon for French. *Int J Med Inform* 2005 Mar;74(2-4):119-24.
 - 30) Weske-Heck G, Zaiss A, Zabel M, Schulz S, Giere W, Schopen M, Klar R. The German specialist lexicon. *Proc AMIA Symp* 2002:884-8.
 - 31) Brooks TA. The Semantic Web, universalist ambition and some lessons from librarianship. *Inform Res* [serial online]. 2002;7(4). [rev. 24 Aug 2004; cited 2 May 2005]. <<http://InformationR.net/ir/7-4/paper136.html>>.
 - 32) Anderson D. Universal bibliographic control: a longterm policy, a plan for action. Pullach bei Mu'nchen, Germany: Verlag Dokumentation, 1974.
 - 33) Mcgregor B. Constructing a concise medical taxonomy. *J Med Libr Assoc* 2005 Jan;93(1):121-3.
 - 34) Greenberg J, Sutton S, Campbell DG. Metadata: a fundamental component of the Semantic Web. *Bull Am Soc Inform Sci Techno* 2003 Apr/May;29(4):16-8.
 - 35) Campbell DG, Fast KV. Academic libraries and the Semantic Web: what the future may hold for research-supporting library catalogues. *J Acad Libr* 2004;30(5):382-90.
 - 36) Soergel D. The rise of ontologies or the reinvention of classification. *J Am Soc Inf Sci* 1999 Oct;50(12):1119-20.