

NewsML을 고려한 기술동향정보 메타데이터에 관한 연구

A Study on Metadata of Technology Trends Information Based on NewsML

이 성 숙(Sung-Sook Lee)*

송 인 석(In-Seok Song)**

목 차

- | | |
|--------------------------|------------------------|
| 1. 서 론 | 4. 1 국내 기술동향정보와 NewsML |
| 1. 1 연구의 배경과 목적 | 의 요소 비교 |
| 1. 2 연구의 방법과 범위 | 4. 2 기술동향정보 요소 선정을 위한 |
| 2. 국내 기술동향정보의 서비스 현황과 | 분석 |
| 적용 메타데이터 | 5. 기술동향정보 메타데이터 |
| 3. IPTC NewsML의 요소와 적용사례 | 5. 1 설명메타데이터 범주 |
| 3. 1 NewsML의 개념과 특징 | 5. 2 내용메타데이터 범주 |
| 3. 2 NewsML의 적용과 확장사례 | 5. 3 관리메타데이터 범주 |
| 3. 3 NewsML의 요소 | 5. 4 권한메타데이터 범주 |
| 4. 국내 기술동향정보와 NewsML의 | 6. 결 론 |
| 요소 비교·분석 | |

초 록

기술동향정보는 연구개발 활동과 과학기술정책의 결정과정에서 매우 중요한 요소이다. 그러나 기술동향정보 메타데이터에 관한 연구는 미비한 실정이다. 이 연구는 뉴스분야에서 표준적 위치에 있는 NewsML을 참고하되, 국내 기술동향정보 서비스에서 사용된 요소를 고려하여 기술동향정보 메타데이터 요소와 스킴을 제안하였다. 이 연구 결과는 앞으로 국내의 실정에 맞는 기술동향정보 메타데이터 표준을 개발하기 위한 기초자료로 이용될 수 있을 것이다.

ABSTRACT

The Information of Technology Trend is one of the critical factors for research activities. In addition, this information tremendously affects the decision-making process of Science and Technology Policy. However, metadata of this information has not gain much attention so far. This research suggests metadata element and encoding scheme that is based on NewsML and using metadata of domestic Information of Technology Trend service system. The result of this research provides the basic resources for developing standard metadata of Information of Technology Trend which is suitable for domestic information technology situation.

키워드: 기술동향정보, 메타데이터

Metadata, Technology Trends Information, NewsML

* 충남대학교 사회과학대학 문헌정보학과 강사(inflee@hanmail.net)

** 한국과학기술정보연구원 표준화기술지원실 선임연구원 (sis@kisti.re.kr)

논문접수일자 2005년 8월 15일

게재확정일자 2005년 9월 15일

1. 서론

1.1 연구의 배경과 목적

오늘날 기술동향정보는 연구개발 활동과 과학기술정책의 결정과정에서 매우 중요한 요소이다. 그러므로 국내외 첨단 기술동향정보의 조직적 수집과 활용능력을 강화하여 기술보호주의에 능동적으로 대처할 수 있도록 하며, 관련 기술개발정보 및 정책동향을 신속히 수집함과 동시에 수집된 정보를 용이하게 활용할 수 있는 데이터베이스를 구축해야 한다. 또한 기술동향정보를 적극적으로 활용하기 위해서는 개별적으로 중복 제작되고 있는 기술동향관련 정보를 통합적으로 공유하고 재사용할 수 있는 체제가 마련되어야 한다.

현재 기관에서 서비스되고 있는 기술동향정보는 체계적인 기준이 마련되어 있지 않고, 기관마다 다양한 메타데이터가 존재한다. 그로 인해 상호운용성 확보가 어렵고, 데이터의 관리와 이용자의 검색에 있어서 비효율적이고, 비용낭비의 결과를 초래하고 있다. 따라서 기술동향정보를 구축하는 기관이 메타데이터를 체계적으로 구축할 수 있도록 하고 통합검색을 지원하기 위한 지침 마련이 절실한 실정이다.

기술동향정보는 대개 인터넷 뉴스사이트, 신문, 잡지, 학술지 등에 게재되는 과학기술과 산업기술에 관한 새로운 소식으로, 기술동향정보가 가지는 속보성 때문에 주로 뉴스의 형식을 띠고 있다. 또한 실질적으로 기술동향정보 서비스 기관도 뉴스제공기관에서 제공하는 정보를 기반으로 하여 서비스하고 있다. 따라서 보다 효율적으로 기술동향정보를 서비스하기 위해서

는 뉴스산업분야의 NewsML(News Markup Language)을 살펴보고, 기술동향정보와의 연계를 고려해 볼 필요가 있다. NewsML은 IPTC(International Press and Telecommunications Council, 국제언론통신평의회)에 의해서 전 세계적으로 통용되는 기술표준으로, 현재는 뉴스의 작성, 편집, 관리, 인쇄, 교환, 아카이브 구축의 전 영역에서 표준의 위치에 있다.

그러나 NewsML을 기술동향정보에 활용하기 위해서는 국내의 기술동향정보와의 호환성을 보장할 수 있는 기술요소의 수용이 필요하다. 이는 NewsML이 뉴스의 작성, 편집, 관리, 전송에 초점을 두고 있다면, 이 연구의 대상인 기술동향정보는 신문사와 같이 사건 취재를 통한 뉴스의 생산보다는 이미 완성된 뉴스의 재가공에 초점을 두기 때문이다.

이 연구는 기술동향정보를 위한 메타데이터 요소와 스킴을 제안하기 위한 것이다. 이 연구에서는 기술동향정보 메타데이터를 선정하기 위해 국내 기술동향정보 서비스 사례에서 사용된 요소와 뉴스분야에서 표준적 위치에 있는 NewsML을 고려하였다.

이 연구 결과는 기술동향정보의 검색과 관련 DB 통합을 위한 체제 마련에 활용될 수 있다. 또한 향후 국내의 실정에 맞는 기술동향정보 메타데이터 표준을 개발하기 위한 기초자료로 활용될 수 있을 것이다.

1.2 연구의 방법과 범위

이 연구의 방법과 범위는 다음과 같다.

첫째, 문헌조사 및 기술요소 선별 부분에서는 1) 기술동향정보 메타데이터의 특성을 파악

하기 위해 국내 기술동향정보 서비스의 현황과 적용 메타데이터를 분석한다. 2) NewsML의 개념, 요소와 각국에 미친 영향과 적용 사례를 조사한다.

둘째, 요소 개발 및 설계 부분에서는 1) 국내 기술동향정보에 사용된 요소와 NewsML 요소의 분석 결과를 토대로 요소 세트를 개발한다. 2) 기술동향정보의 기술에 필요한 각종 스킴을 제시한다. 이는 메타데이터 생산을 용이하게 하기 위함은 물론 정보를 일관되게 기술하게 함으로써 효율적인 검색을 제공하기 위한 것이다.

이 연구에서 제안된 기술동향정보의 요소와 스킴의 범위는 기술(description) 메타데이터에 초점을 둔다. 또한 이 연구는 메타데이터 요소에 초점을 두며, 구조는 검토하지 않았다.

2. 국내 기술동향정보의 서비스 현황과 적용 메타데이터

기술동향정보는 학술지, 인터넷 전문사이트, 신문, 잡지 등에 게재되는 최신 과학기술 뉴스 정보이며, 과학기술계의 새소식은 물론 주요 정책동향과 연구동향을 제공한다. 정보서비스 기관에서 제공하는 기술동향정보 자체는 뉴스에 포함되는 하위 범주라고 볼 수 있다. 그러나 기술동향정보는 신문사에서 사건을 직접 취재하여 생산하는 뉴스와는 달리 이미 발표된 뉴스를 선별, 가공하여 제공하거나 링크를 제공한다는 특성이 있다.

과학기술정보의 특성은 지리적 경계를 초월하는 공통성과 첨단 정보에 대한 필요성과 속보성, 신뢰성, 검증가능성이라고 할 수 있다. 기술

동향정보는 이러한 특성을 모두 만족시켜야하는 서비스이다. 그래서 과학기술분야는 다른 분야보다 기술동향에 관한 뉴스의 중요성이 더 높다고 할 수 있다. 기술동향정보는 이용자들의 이용도가 높은 정보 중의 하나이다. 예를 들어 2001년 조사에 의하면 KISTI의 경우, 기술동향정보 서비스인 '해외과학기술동향정보'는 KISTI 내 정보 서비스 중 전체 조회 횟수의 75%를 차지하는 대표적 인기 정보 서비스이다(최성배, 한선화 2001).

이 연구에서는 국내 기술동향정보 서비스에 적용되는 메타데이터를 살펴보기 위하여 국내 기관에서 제공되는 서비스 중 5개를 선정하였다. 대상 선정 기준은 자체적으로 해당 주제분야의 기술동향을 제공하는 기관으로서, 기관정보보다는 기술동향 제공의 성격이 강한 서비스, 분석의 성격보다는 동향의 성격이 강한 서비스를 대상으로 하였다. 조사 대상 기관의 서비스 특징과 사용된 메타데이터는 <표 1>과 같다.

국내 기술동향정보의 적용 메타데이터 분석 결과, 이 연구의 메타데이터 선정시 고려해야 할 점은 다음과 같다.

첫째, 국내 기술동향정보는 기본적인 요소만을 제공하여 상세성이 부족하고 동일한 의미의 메타데이터에 대해 다양한 요소명을 사용하고 있는 것으로 나타났다. 따라서 제안될 메타데이터는 기술동향정보를 충분히 표현하고 효과적으로 검색·관리할 수 있어야 하고, 상호운용성 확보를 위해 표준화된 메타데이터 형식에 기반해야 한다.

둘째, 국내 기술동향정보에 적용된 메타데이터는 제목, 본문, 본문키워드, 정보출처, 원문언어, 국가코드, 주제분야(주제코드), 원문출판일,

〈표 1〉 국내 기술동향정보 서비스 현황과 적용 메타데이터

번호	기관명	서비스명	주제분야	정보원	서비스 특징	적용 메타데이터
1	한국과학기술정보연구원 (KISTI) http://www.kisti.re.kr	-과학기술동향 -기술산업동향	-전반적인 과학기술 분야 -산업분야	-국내·외 과학기술분야의 신문, 인터넷전문사이트, 잡지, 학술지 등	-뉴스 전달 -해외자료 번역 -자체 분류코드 부여	제목, 본문, 본문키워드, 정보출처, 원문언어, 국가코드, 주제분야, 원문출판일, 등록일, 작성자, 직책, 소속기관, 이메일, 최종학위, 전문분야, 자료유형, 원본파일, 그림파일
2	전자정보센터 http://www.eic.re.kr	-일일동향	-전자분야	-국내 신문, 저널 등	-뉴스 전달 -자체 분류코드 부여	제목, 본문내용, 기사출처, 기사날짜, 기자명, 저작권자
3	정보통신연구진흥원 ITFIND http://www.itfind.or.kr	-오늘의 IT News -IT839 전략 -해외IT 정보	-IT 관련 분야	-국내 주요 신문, 저널(영자 포함) 등	-뉴스 전달 -일부기사에 대해선 자체 분류코드 부여	제목, 본문내용, 기사명, 출처, 기사날짜, 주제명, 작성자, 등록일, 첨부파일
4	산업정보망 http://www.iin.co.kr/	-산업뉴스	-전반적인 산업분야(국내, 중국, 일본, 러시아 등)	-연합뉴스	-뉴스 전달 -국가별 분류	제목, 뉴스내용, 기사출처, 기사날짜, 저작권자
5	한국무역협회 http://www.kita.net/	-품목별시장정보	-전반적인 산업분야	-국외 저널, 잡지 등	-뉴스 전달 -해외자료 번역 -자체 분류코드 부여	제목, 작성자, 입력일, 첨부파일, 내용출처

자료유형, 원문파일, 그림파일, 저작권자, 등록일, 기자명(작성자), 작성자직책, 작성자소속기관, 작성자이메일, 작성자최종학위, 작성자전문분야 등으로 조사되었다. 기술동향정보 메타데이터 선정시 이러한 요소를 포함시키는 것이 타당하다.

셋째, 제공되는 기술동향정보의 정보원은 신문의 기술동향 기사가 대부분을 차지하지만, 인터넷전문사이트, 잡지, 학술지 등에 수록된 최신동향 뉴스도 포함된다. 개발되는 기술동향정보의 자료유형은 온라인과 오프라인의 신문, 잡지, 학술지 등을 포괄해야 한다.

넷째, 국내 기술동향정보 서비스는 단순 링크만을 제공하기도 하고, KISTI의 '해외과학기술동향정보'와 같이 이미 작성된 기사를 번역하고 간단한 분석을 첨부한 기사를 제공하기도 한다. 제안될 메타데이터는 단순히 뉴스를 전달하는 서비스와 뉴스를 가공(번역, 분석 등)하는

서비스의 유형을 포괄하도록 해야 한다.

다섯째, 국내 기술동향정보의 저작권 관련 정보는 기사 하나 하나에 대한 저작권자와 날짜를 위한 메타데이터는 유지하지 않으며, 특히 저작권날짜에 대한 메타데이터가 부족한 것으로 나타났다.

한편, 한국온라인신문협회에서 제정한(2005년 3월 1일) 디지털뉴스이용규칙에서는 디지털 저작물의 이용에 관하여 꺾임이 아니라 링크(단순링크 및 직접링크)를 권고하고 있다. 직접링크(deep link)에 대한 법적 판단은 아직 진행 중이며, 일시적으로 비영리·일반 개인 네티즌에 한해 한정적 범위에서 허용하고 있다(한국온라인신문협회 2005). 이에 따라 한국항공우주연구원 등 몇몇 기관은 기술동향정보 서비스를 중단하였다. 제안될 기술동향정보는 저작권과 관련된 메타데이터를 충분히 고려해야 한다.

3. IPTC NewsML 의 요소와 적용사례

3. 1 NewsML 의 개념과 특징

NewsML은 뉴스 유통의 비효율성과 기술적 낙후를 개선하기 위해 1998년 로이터 통신이 처음 설계해 1999년 IPTC에 표준화를 제안하면서 시작되었다. IPTC는 원래 통신사가 콘텐츠를 서비스하기 위한 표준을 만들기 위하여 조직된 단체로서, 로이터, AP, AFP 등 국제적인 통신사를 비롯해 뉴욕타임스, BBC, 일본신문협회, 세계신문협회와 같은 언론단체, 시스템 벤더 등 총 55개사가 회원사로 가입되어 있다. IPTC는 2000년 10월 NewsML의 버전 1.0을 발표하였고, 2003년 10월에 발표된 NewsML 버전 1.2는 IPTC가 발표한 가장 최신의 국제 표준 뉴스 포맷이다.

NewsML은 XML을 기반으로 하여 디지털 뉴스에 관한 정보를 효과적으로 교환·저장·검색할 수 있도록 만든 기술표준이며, 간략하고 확장성이 있으며 유연한 구조체계이다. NewsML은 원래 뉴스 교환을 위한 표준 포맷으로 설계되었으나 점차 아카이브 구축, 뉴스의 작성, 편집, 관리, 인쇄의 전 영역에서 표준이 되어가고 있다.

NewsML은 외부유통시 전송처에 비례해서 증가하는 관리·가공·시스템 운용비용을 줄일 수 있고, 콘텐츠 관리 체계를 표준화하여 업무 프로세스도 표준화함에 따라 생산성을 높일 수 있다. 또 표준규격에 의한 메타데이터(콘텐츠와 이를 설명하는 모든 정보)를 활용해 상품 개발과 콘텐츠 자동처리가 용이해지고, 웹·모바일 등 멀티미디어에 콘텐츠를 제공하는 것

도 쉬워진다(미디어 오늘 2005).

기사 작성과 관련하여 NewsML의 장점은 다음과 같다. 첫째, 뉴스와 관련된 텍스트, 이미지, 동영상 등을 대등하게 처리할 수 있다. 둘째, 뉴스를 작성하는데 편리하고, 뉴스 작성에서 흔히 발생하는 수정기사의 처리에 편리하다. 셋째, NewsML은 여러 개의 뉴스를 통합해서 하나의 NewsML 파일을 생성하기에 용이하다. 넷째, 개별 뉴스간의 관계를 설정할 수 있다. 다섯째, 다양하게 표현된 동일한 내용을 효과적으로 처리할 수 있다. 여섯째, 특정 부분을 포함시킬 때 내용을 직접 쓰지 않고 참조에 의해서 표현할 수 있다. 일곱째, 표준 스키마와 비표준 스키마를 지원하기 위한 메타데이터를 추가할 수 있다(IPTC NewsNL Home 2005).

3. 2 NewsML 의 적용과 확장 사례

3. 2. 1 NewsML의 적용 사례

전세계적으로 통신사와 신문사를 비롯한 여러 기관이 NewsML을 채택한 시스템을 운영하고 있다. NewsML을 사용하고 있는 대표적인 기관으로는 NewsML을 설계하였고 IPTC에 표준안을 제안하였던 로이터통신을 비롯하여, 여러 국가의 언어로 다양한 멀티미디어를 통해 뉴스를 제공하는 AFP, NewsML에 기반한 고속의 뉴스 배포 시스템을 개발한 Business Wire 이외에도 UPI, Dow Jones' WSJ.Com을 들 수 있다(Compton 2005; Le Meur 2005; Lorenzen 2005).

일본은 뉴스분야에 NewsML을 성공적으로 도입하여 적극적으로 활용하고 있다. NewsML이 뉴스기사 표준 형식으로 채택된 것은 이미지

와 텍스트가 함께 처리될 수 있고, 콘텐츠 관리에 적합하도록 메타데이터를 추가할 수 있으며, 플랫폼과는 독립적으로 사용가능하기 때문이다(Fujiwara and Tanaka 2003).

또한 일본신문협회는 일본 내에서 NewsML의 사용을 장려하기 위하여 각 신문사의 대표들과 솔루션 제공업체 등으로 기술위원회를 구성하였다. 이 기술위원회는 각 신문사 기사를 기초로 NewsML의 각 요소들을 비교분석하여 표준안 수용여부를 검증하였고, 일본의 여러 언론사간의 뉴스기사 교환이 효율적으로 이루어질 수 있도록 NewsML 가이드라인을 만들었다. 또한 교토통신사에 NewsML 도입을 제안하여 2001년 뉴스전달포맷으로 NewsML을 공식 채택하도록 하였다. 이후 2001년 NewsML의 일본 표준인 nskNewsML 버전 1이 정식 발표되었으며, 마이니치 신문사가 nskNewsML 버전 1에 기초한 NewsML 소스를 공개하였다. 이러한 계속된 노력으로 일본에서는 NewsML을 채택한 뉴스제공사가 지속적으로 증가하는 추세이며, 일본 내 유일한 표준으로 자리잡고 있다(한국언론재단 2004).

NewsML을 적극적으로 수용한 일본의 시스템 사례로는 2002년 NewsML에 기반한 새로운 시스템을 개발하기 시작해 2004년 2월부터 뉴스를 NewsML 방식으로 공급하는 일본의 대표적인 통신사인 교토통신(KYOTO NEWS), NewsML에 기반한 MAITY시스템과 PALACE 시스템을 운영하고 있는 마이니치신문(MAINICHI), NewsML에 기반한 Photo DB를 개발한 주니치신문(CHUNICHI), NewsML 사진관리 시스템을 개발한 류쿠신보를 들 수 있다(Fujiwara and Tanaka 2003; 한국언론

재단 2004; 황정근 2004).

중국도 2002년 1월 중국 NewsML 협회를 구성해 대만, 홍콩과 함께 NewsML을 도입하기 위한 논의를 하고 있다. 중국 NewsML 협회는 NewsML 컨소시엄을 구성하고, 정보집적시스템으로 NewsML에 기반한 CIDAX를 개발하고 있다. 또한 중국의학정보를 위해 CIDAX를 활용한 중국의학동향정보시스템을 구축하고 있다(Wong 2004).

국내에 NewsML이 알려진 것은 2002년 로이터 통신이 국내 파트너사인 뉴시스 통신사에 NewsML기반의 뉴스 시스템 구축을 권고한 후, NewsML기반의 시스템이 구축된 이후부터이다. 이외는 별도로 연합뉴스는 입체화된 뉴스 정보 전달을 위해 NewsML을 응용한 krNewsML을 개발하였다. 그러나 이런 시도는 NewsML의 전면 도입을 고려한 것은 아니었다. 이에 2004년 10월 한국언론재단에서는 NewsML 전문 기업인 와이즈미디어와 함께 뉴스 포맷 표준화 논의의 진전을 위해 NewsML을 번역하여 소개함과 동시에 NewsML-KR을 개발하였다. NewsML-KR은 NewsML 버전 1.2의 사양을 그대로 준수해 한글화하고 확장한 것으로써 모든 뉴스 제공사가 즉시 사용할 수 있도록 국내 표준 수준으로 설계된 것이다(한국언론재단 2004).

NewsML-KR을 기반으로 한 사례로는 한국언론재단의 카인즈(KINDS)를 들 수 있다. 2004년 말 한국언론재단에서는 NewsML 전문 기업인 와이즈미디어와 함께 개발한 NewsML-KR을 토대로, NewsML 기반의 전송포맷 표준화 방안 토론회를 통해 각계의 의견을 수렴하였다. 2005년 1월에는 한국언론재단의 카인즈 서비스(www.kinds.or.kr)를 NewsML에 기

만한 181개 매체의 통합검색이 가능한 사이트로 개편하였다.

NewsML 표준 포맷의 확산을 위하여 한국 언론재단은 NewsML에 관심있는 온라인 뉴스 업계 관계자들의 모임인 'NewsML 포럼'을 만들고 사이트(<http://www.newsml.or.kr>)를 운영하고 있다. NewsML을 적용한 시스템 개편은 각 매체사들이 표준화된 형식으로 시스템을 개편하고자 할 때 지원 가능한 노하우를 축적하는 계기가 되었고 향후 멀티미디어 뉴스 제공 및 다양한 맞춤형 서비스 수용에 능동적으로 대처할 수 있는 기반을 마련한 것으로 평가되고 있다(오수정 2004: 2005).

3. 2. 2 NewsML을 특정 분야에 적용·확장한 사례

NewsML을 스포츠, 날씨 등 특정 분야에 적용하여 NewsML을 확장한 사례로는 SportsML, ProgramGuideML, EventsML, WeatherML 등이 있다.

첫째, SportsML은 스포츠 기사의 다양한 특성을 고려한 표준에 대한 요구에 의해 개발되었다. IPTC는 2001년 3월 SportsML 프로젝트를 시작하였고, 그 결과 경기의 점수, 경기 일정표, 선수들의 순위표, 여러 경기결과와 통계와 같은 스포츠 자료를 교환하기 위한 SportsML을 완성하였다. SportsML은 가능한 이해하기 쉽고, 이용되기 쉽도록 설계되었으며, 뉴스제공기관, 스포츠데이터제공업자, 이용자 간에 스포츠데이터를 공유할 수 있도록 한다.

둘째, ProgramGuideML은 NewsML에 기반한, 라디오/TV 프로그램 정보를 교환하기 위한 표준이다. 즉 방송국이 방송하는 하루분의

방송내용에 대해 하나의 방송프로그램 단위로 NewsML화하여 표현하는 것이다. Program-GuideML은 뉴스제공업자와 방송국이 프로그램 편성표, 이미지, 실황방송, 방송용 뉴스기사, 표준적인 프로그램 정보 등의 다양한 정보를 효율적으로 교환할 수 있도록 한다.

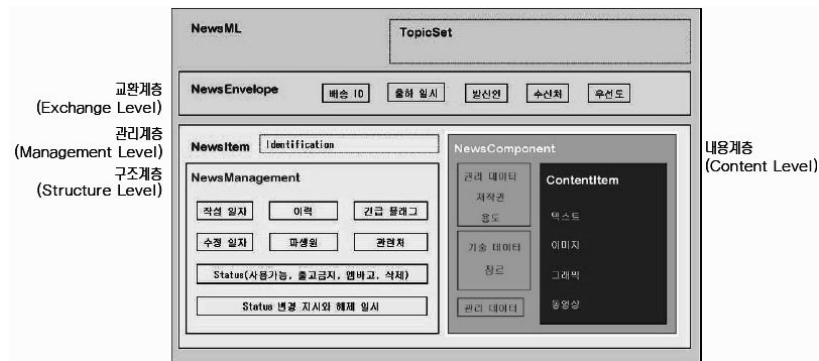
셋째, EventsML은 IPTC가 새롭게 개발하기 시작한 표준이며, 이 표준은 이벤트에 관한 정보를 전달하기 위한 것이다. EventML에 관련한 프로젝트는 보도가치가 있는 이벤트 정보를 위한 정보 교환 표준을 정의하는데 제한되어 있다.

넷째, WeatherML은 IPTC가 새롭게 추진하고 있는 프로젝트 중 하나이며, WeatherML 프로젝트의 임무는 과거, 현재 예상되는 기상정보를 제공하기 위한 WDD(Weather Data Definition)를 제안하는 것이다. WDD는 데이터 수집, 기록보존, 정보 배포를 지원한다 최종적으로 WeatherML에는 기상과 직·간접적으로 영향이 있는 지면, 수면 대기 등에 관한 정보를 포함한다.

3. 3 NewsML의 요소

NewsML은 <그림 1>과 같이 ContentItem, NewsComponent, NewsItem, NewsEnvelope 등으로 이루어진 계층구조를 이루며, 네 가지 계층이 각각 다른 기능을 담당하고 있다(한국 언론재단 2004).

<그림 1>에서 가장 안쪽에 있는 ContentItem은 내용계층으로 뉴스 콘텐츠의 내용을 기술한다. ContentItem은 뉴스 환경에서 다루어지는 정보의 기본단위이며, 여기에는 뉴스 본문과 함

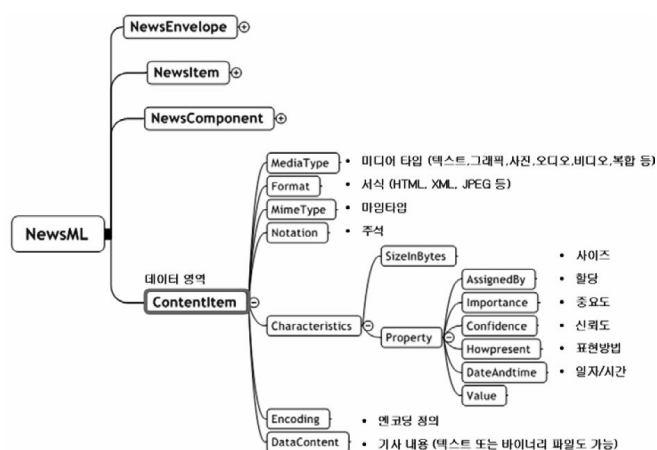


〈그림 1〉 NewsML의 구조

께 이것이 텍스트인지, 멀티미디어인지를 알려 주는 부가정보를 포함한다. NewsML에서는 텍스트문서, JPEG 사진, EPSF 그래픽, Flash 애니메이션 등과 같이 다양한 종류의 미디어를 효과적으로 표현할 수 있다(그림 2 참고).

구조화 계층인 NewsComponent는 뉴스의 개체를 담는 컨테이너 기능을 한다. NewsComponent는 여러 개의 ContentItem 즉 다수의 콘텐츠를 포함할 수 있다. 일례로, 텍스트 기사와 이

기사와 관련된 여러 개의 사진(섬네일, 인터넷 용, 출판용), 그리고 사진에 대한 별도의 설명문 등으로 구성될 수 있다. NewsComponent는 구성요소가 포함하거나 참조한 콘텐츠의 생성이나 뉴스 개체들의 묶음에 관한 정보를 제공하는 AdministrativeMetadata, 구성요소가 포함하거나 참조한 콘텐츠의 저작권이나 사용권한에 관한 정보를 제공하는 RightsMetadata, 구성요소가 포함하거나 참조한 콘텐츠에 관한 정

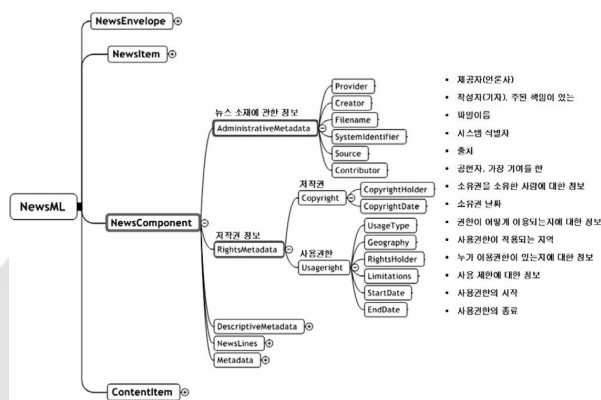


〈그림 2〉 NewsML의 ContentItem 요소
(출처: 한국언론재단, 2004)

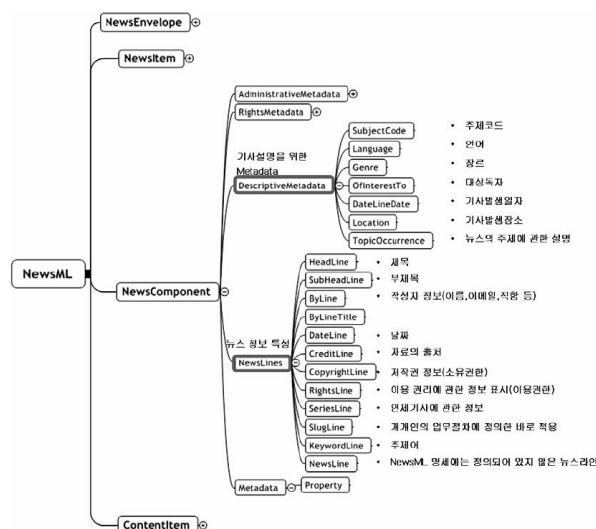
보를 제공하는 DescriptiveMetadata, 구성요소에 관한 정보를 사람들이 읽을 수 있도록 하는 NewsLines 등으로 구성된다. NewsComponent는 해당 콘텐츠의 제목, 제공자, 작성자, 출처, 파일이름, 작성일자, 저작권, 사용권한, 주제, 장르, 키워드 등과 같은 다양한 정보를 기록할 수

있다(그림 3, 4 참고).

관계체인 NewsItem은 뉴스를 관리하는 가장 중요한 단위이다. NewsItem은 개별적으로 식별이 가능하고 출고할 수 있는 뉴스의 단위 조각이다. 정보제공자는 뉴스를 생성하고 보관하고 관리하고 재사용하며, 다른 NewsItem



〈그림 3〉 NewsML의 NewsComponent 요소_1
(출처: 한국언론재단, 2004)



〈그림 4〉 NewsML의 NewsComponent 요소_2
(출처: 한국언론재단, 2004)

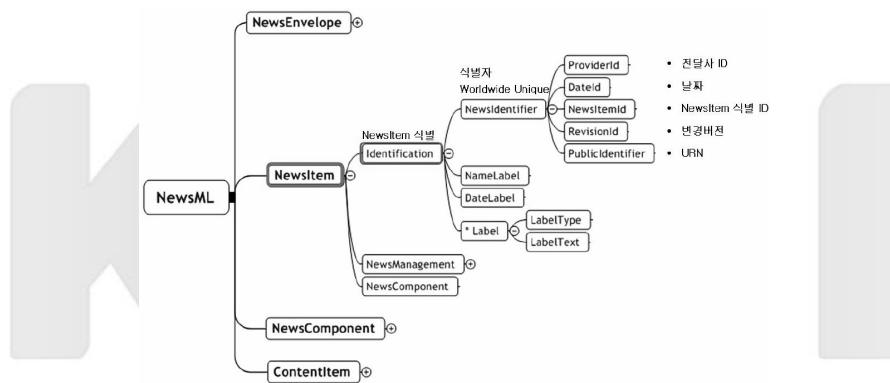
을 참조할 수 있다. NewsItem은 식별정보를 포함해 제공된 뉴스에 대한 변경이력, 긴급도 등처럼 뉴스콘텐츠 관리에 필요한 메타데이터로 이루어진다(그림 5, 6 참고).

NewsML 문서의 가장 바깥쪽에 위치한 교환계층인 NewsEnvelope는 뉴스 콘텐츠를 배포하는 역할을 한다. 해당 뉴스 콘텐츠를 누가, 누구에게 제공하는지와 제공되는 날짜와 시간, 그리고 서비스나 상품에 대한 정보 등을 포함한

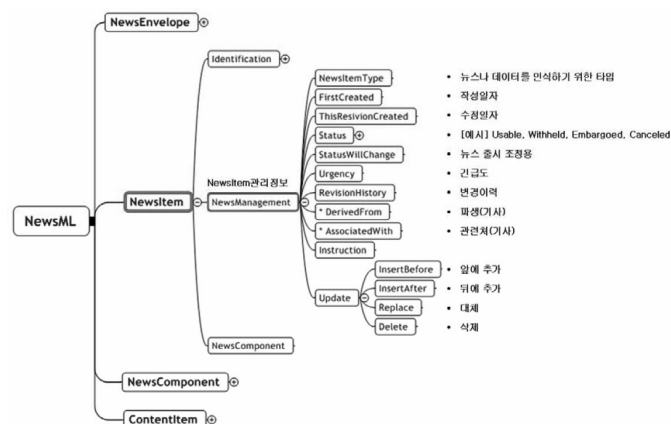
다. NewsEnvelope에는 전송 ID, 전송처, 송신자, 전송일자/시간, 우선도 등을 기술할 수 있다(그림 7 참고).

이와 같이 뉴스콘텐츠의 내용 외에 이에 대한 다양한 정보를 제공함으로써 콘텐츠의 활용도를 높이고, 관리의 효율성도 높일 수 있다.

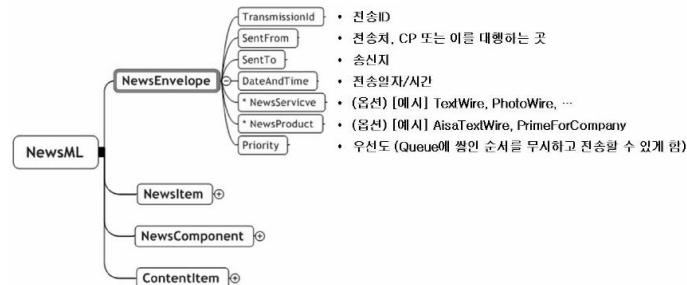
NewsML의 또 다른 특징은 콘텐츠 분류와 용어에 대한 표준을 함께 제공한다는 것이다. 뉴스 콘텐츠의 원활한 교환을 위해서는 콘텐츠



〈그림 5〉 NewsML의 NewsItem 요소_1 (출처: 한국언론재단, 2004)



〈그림 6〉 NewsML의 NewsItem 요소_2 (출처: 한국언론재단, 2004)



〈그림 7〉 NewsML의 NewsEnvelope 요소 (출처: 한국언론재단, 2004)

형식의 표준화뿐만 아니라 콘텐츠를 설명하는 각종 용어에 대한 표준화도 동시에 이루어져야 한다. 그렇지 않으면 비록 NewsML로 제공했다 하더라도 부가정보를 인식할 수 없거나 다른 의미로 받아들여 오류가 발생할 수 있다. 일례로 A사가 콘텐츠의 주제 분류값을 '경제'로 분류해 전송한 것을 B사는 'economy'로 표현하고 있다면 시스템적 오류가 발생하거나 사람이 인식하기에도 불편이 따르고 나아가 별도의 변환이 필요하게 된다.

NewsML에서는 이러한 문제를 해결하기 위해 반복적으로 자주 사용되거나 정해진 표현으로 명기해야 하는 각종 용어를 미리 정의해 두었는데 이런 용어를 '통제용어'라고 부른다

NewsML에서는 통제용어를 유형별로 묶어 토픽세트(TopicSet)를 제공하고 있다. IPTC는 통제용어가 등록된 토픽세트를 각국 언어로 번역해 사용할 수 있도록 웹사이트에 공개하고 있다. 각국의 모든 뉴스 공급자와 이용자들이 공통적으로 토픽세트를 사용함으로써 콘텐츠의 유통을 손쉽게 하고 관리의 효율을 높일 수 있도록 하고 있다.

토픽세트와 여기에 등록된 통제용어는 사전의 역할을 하는 동시에 뉴스 콘텐츠를 분류하는 기능을 하기 때문에 토픽세트는 뉴스 콘텐츠 분류계통도(Taxonomy)이기도 하다. NewsML 최신 버전에 포함된 토픽세트의 종류는 〈표 2〉와 같다(한국언론재단 2004).

〈표 2〉 NewsML 토픽세트의 종류

• 신뢰도 • 중요성 • 공개 지시 • 주석 • 미디어 타입 • NewsItem 타입 • 전달 방법 • NewsItem 라벨 • 국가명 • 지(역)명	• 긴급도 • 메타데이터의 중요성 • 주제 코드 • NewsLine 타입 • 토픽 타입 • 룰(Role) • 독자층 • 서비스의 상표 • 통화(단위) • (미) 산업분류코드	• 전달우선순위 • 독자와의 관련성 • 포맷 • 마임(Mime) 타입 • 메타데이터 • 장르 • 언어 • 단체명 • 나스닥 회사 코드
--	---	--

출처: 한국언론재단, 2004.

NewsML의 내용과 적용사례, 요소들을 검토한 결과 이 연구의 기술동향정보 메타데이터 선정에 NewsML을 고려해야 되는 이유는 다음과 같다.

첫째, 기술동향정보는 속보성이라는 특성 때문에 대부분 뉴스의 형태를 띠고 있다. 따라서 뉴스의 형태를 띠고 있는 기술동향정보의 요소 선정시 뉴스분야의 경험이 반영된 NewsML을 고려하는 것이 바람직하다.

둘째, 뉴스 메타데이터의 대표적인 표준으로 세계적인 통신사와 신문사에서 사용되고 있는 NewsML은 국내에서도 한국언론재단을 중심으로 그 사용을 독려하고 있고, NewsML의 적용사례를 검토한 결과 향후 뉴스관련분야의 표준으로 확산될 것으로 예상된다. 이러한 상황에서 뉴스분야의 표준인 NewsML을 검토하여 유사한 요소를 유지함으로써 보다 신속하고 효율적인 서비스를 제공할 수 있고, 향후 신문사와 같은 뉴스제공기관이 NewsML의 체제로 전면 개편된다면 기술동향정보도 뉴스제공기관으로부터 직접 데이터를 제공받아 활용할 수 있어 그 효과는 배가 될 것으로 예상되기 때문이다.

셋째, 기술동향정보 메타데이터는 기술동향정보를 충분히 기술하여 효과적으로 검색할 수 있는 요소를 마련하고, 특히 저작권 등을 고려한 체계적인 메타데이터를 제공해야 하는데, NewsML에서 제공되는 일련의 메타데이터는 이러한 요구에 적합하기 때문이다.

넷째, NewsML에서 제공되는 다양한 토픽 세트를 활용하여 뉴스 검색과 관리가 더 효과적으로 이루어질 수 있고, 효율적인 데이터 공유가 가능하기 때문이다.

다섯째, 뉴스산업 분야에서 NewsML의 표

준규격에 의한 메타데이터(콘텐츠와 이를 설명하는 모든 정보)를 활용하여 멀티미디어에 콘텐츠를 제공하는 등 새로운 상품 개발에 활용하듯이, 향후 기술동향정보 서비스 개선을 위한 기초자료로 활용하기 위한 것이다.

4. 국내 기술동향정보와 NewsML의 요소 비교·분석

4.1 국내 기술동향정보와 NewsML의 요소 비교

이 절에서는 국내의 기술동향정보 메타데이터 요소를 선정하기 위하여, 앞 절에서 살펴본 국내 기술동향정보 요소와 NewsML의 요소를 비교분석함으로써 요소선정의 근거자료로 활용하고자 한다.

이를 위해 먼저 뉴스의 작성, 편집, 관리, 전송 등의 전 영역을 커버하는 NewsML 요소 중 뉴스의 교환계층인 NewsEnvelope의 요소와 뉴스출고 등의 관리계층인 NewsItem의 요소는 기술동향정보를 기술하기 위한 요소가 아니기 때문에 고려대상에서 제외하고, NewsComponent와 ContentItem을 중심으로 살펴 보았다.

국내 기술동향정보에서 사용된 요소와 NewsML 요소의 매핑 결과는 <표 3>과 같다. <표 3>의 왼쪽은 앞에서 조사된 국내 기술동향정보 서비스 제공 기관에서 사용하고 있는 요소이고, 오른쪽은 NewsML의 요소 중 NewsComponent와 ContentItem에 해당하는 요소들이다.

<표 3>의 비교 결과 국내 기술동향정보와 NewsML에서 대응하게 매핑되는 요소로는 제

〈표 3〉 국내 기술동향정보 요소와 NewsML 요소의 비교

국내 기술동향정보 요소	NewsML 요소
제목	HeadLine
기사작성일(등록일)	DateLine
본문키워드	KeywordLine
기사날짜(원문출판일)	DateLineDate
국가코드	Location
본문내용	DataContent
기자명(작성자)	Creator
작성자직책	ByLineTitle
작성자소속기관	ByLine
작성자이메일	ByLine
정보출처	Source
번역원본파일	Source_NewsItemHref*
그림파일	ImageHref*
원문파일	DataContentHref*
저작권자	Copyright_CopyrightHolder
주제분야(주제코드)	SubjectCode
원문언어	
자료유형	
작성자최종학위	
작성자전문분야	
	SubHeadLine
	SeriesLine
	Language
	Genre
	MediaType
	Format
	MimeType
	Notation
	Characteristic_SizeInBytes
	Contributor
	Provider
	SystemIdentifier
	Copyright_CopyrightDate
	UsageRight_UsageType
	UsageRight_Geography
	UsageRight_RightHolder
	UsageRight_Limitations
	UsageRight_StartDate
	UsageRight_EndDate
	Role
	BasisForChoice
	RightsLine
	CopyrightLine
	SlugLine
	CreditLine
	FileName
	Encoding
	OfInterestTo
	TopicOccurrence
	Property
	Newsline

* NewsML에서는 속성으로 정의되어 있음

목, 기사작성일, 본문키워드, 기사날짜, 국가코드, 본문내용, 작성자, 작성자직책, 작성자소속기관, 작성자이메일, 정보출처, 번역원본파일, 그림파일, 원문파일, 저작권자, 주제분야(주제코드) 등이 있다.

서로 매핑되지 않는 요소 중에서 국내 기술동향정보에서 사용되는 요소로는 원문언어, 자료유형, 작성자최종학위, 작성자전문분야 등이다. NewsML에서만 정의된 요소로는 SubHeadLine, SeriesLine, Genre, Language, MediaType, Format, MimeType, Notation, Characteristic_SizeInBytes, Contributor, Provider, SystemIdentifier, Copyright_CopyrightDate, UsageRight_UsageType, UsageRight_Geography, UsageRight_RightHolder, UsageRight_Limitations, UsageRight_StartDate, UsageRight_EndDate, Role, BasisForChoice, RightsLine, CopyrightLine, SlugLine, CreditLine, FileName, Encoding, OffInterestTo, TopicOccurrence, Property, Newslane 등으로 나타났다.

4. 2 기술동향정보 요소 선정을 위한 분석

기술동향정보 메타데이터는 신속성, 신뢰성, 검증가능성, 검색성의 특징을 충족시켜주어야 하며, 기술동향정보 요소의 선정에서는 기술동향정보의 특성이 잘 표현될 수 있도록 요소를 선정할 필요가 있다.

1) 매핑되는 요소

기술동향정보의 요소 중 NewsML에 매핑된 제목, 기사작성일, 본문키워드, 국가코드, 그림파일, 본문내용 요소는 거의 동일한 의미로 사용되어 일대일로 매핑 되었다. 기술동향정보 메

타데이터와 NewsML에서 공통으로 사용된 이러한 요소들은 제안될 메타데이터에서도 수용되는 것이 바람직하다.

특히 기사작성일과 등록일, 원문출판일 등을 관리하는 것은 기술동향정보의 신속성을 담보하기 위한 것이다. 이외에도 NewsML에서는 뉴스에서 취급된 사건의 발생일자를 기술하며, 날짜와 관련된 코드로 ISO8601: 2000 basic format을 사용함으로 보다 통일된 관리는 물론 일시에 대한 정확한 기술과 함께 신속하게 제공해줄 수 있도록 하고 있다.

주제코드와 관련해서, 현재 국내 기관에서 사용하는 주제코드는 각 개별기관의 특성을 반영한 주제코드이다. NewsML에서도 SubjectCode를 정의하여 제공함으로써 NewsML에 기반하여 기사를 작성할 때 사용하도록 하고 있다. 그러므로 NewsML을 기반으로 작성된 기사를 연계하여 검색하기 위해서는 정의한 NewsML의 SubjectCode를 수용하여 활용하는 것이 바람직하다. 그러나 NewsML에서 제공되는 SubjectCode 이외에 개별 기관에서 특수하게 사용될 수 있는 별도의 분류코드가 필요한 경우가 있을 수 있다. 이런 경우 각 기관의 SubjectCode에서 발생하는 모든 경우를 수용하기 위한 통합분류체계를 생성하기보다는 예외적으로 기관의 특수성을 반영할 수 있는 '로컬주제코드'를 추가하는 것이 효과적이다.

2) 국내 기술동향정보에만 있는 요소

국내 기술동향정보에서 사용되는 요소 중 '자료유형'과 '원문언어', '작성자최종학위', '작성자전문분야'는 NewsML의 요소와 매핑되지 않았지만 제안될 기술동향정보에서는 유지되어

야 할 요소이다. 이는 기술동향정보는 언론사의 뉴스 기사뿐만 아니라 학술지, 잡지 등의 최신 동향정보도 포괄하는 것이므로 이를 구분해 주기 위해 요소 '자료유형'을 포함시킬 필요가 있다. 또한 기술동향정보는 이미 작성된 기사를 선별하여 게시하거나 번역하는 별도의 과정을 거치므로 가공전 원뉴스의 '원문언어'를 기술해 줄 필요가 있고, 번역 등의 가공을 하는 경우 원문언어와 작성 기사의 언어, 혹은 직접 기사를 작성시에 사용하는 언어를 명확하게 구분할 필요가 있다.

기술동향정보를 여러 기관에서 상호운영하거나 공유할 경우를 고려해서, 어떠한 질적 기준이 보장되는 신문사에 속한 기자와는 달리 작성자에 대한 전문성을 판단할 수 있는 요소인 '작성자최종학위', '작성자전문분야'를 기술해 줄 필요가 있다.

3) NewsML 에만 있는 요소

(1) 포함시킬 요소

• SubHeadLine, SeriesLine, Genre, Language, MediaType, Format, MimeType, Characteristic_SizeInBytes, SystemIdentifier

제목과 관련하여, SubHeadLine은 제목 이외의 추가적인 제목을 제공하는 것으로, NewsML 에서는 제목의 역할을 하는 HeadLine이라는 요소와는 별도로, 부수적인 제목의 역할을 하는 SubHeadLine을 정의하고 있다. 기술동향정보의 가공 대상이 되는 뉴스에서도 부제목을 사용하고 있고, 이는 정확한 검색을 위해 사용될 수 있기 때문에 부제목을 유지하는 것이 바람직하다. 이외에도 NewsML에서 정의된 요소 중 SeriesLine, Genre, Language는 추가적인 검색을 위해,

MediaType, Format, MimeType, Notation, Characteristic_SizeInBytes, SystemIdentifier 역시 검색과 통합 관리를 위해 활용될 수 있을 것이다.

• Source URL, Creator 상세 정보, Contributor, Provider

기술동향정보는 신뢰성이라는 중요한 특징이 확보되도록 해야 한다. 신뢰성의 판단은 우선 어떠한 근거로 작성되었느냐의 출처와 신뢰도의 질을 판단하는데 있어서 문서 자체를 가지고 신뢰성이나 질적 보장이 어려운 경우 이를 보완하기 위해 작성자 혹은 제공자에 대한 충분한 정보를 제공하도록 하는 요소가 마련되어야 한다.

출처와 관련해서 정보출처로 바로 이동할 수 있는 주소에 관한 요소를 유지하는 것이 바람직하다. 그리고 현재 일부 서비스 기관에서 제공하는 펄글은 저작권 침해의 소지가 있으므로 NewsML의 Source 에서와 같이 제공기관과 출처를 직접 볼 수 있는 주소를 제공하는 것이 저작권 보호의 관점에서 바람직하다. 그러나 주소 변경과 소멸의 우려가 있는 경우 이에 대비한 방법도 강구되어야 할 것으로 보인다.

또한 작성자의 Profile 관리를 통해 작성자에 대한 더 많은 정보를 제공함으로써 기술동향정보의 신뢰성을 확보하기 위한 요소로 활용할 수 있다. 또한 이는 관리뿐만 아니라 최종이용자의 입장에서 보면 작성자에 대한 정보를 확인할 수 있도록 하는 것이 필요하다. 예를 들어 작성자 직책, 소속기관, 최종학위, 전문분야, 이메일 등에 관한 정보를 제공할 필요가 있다.

또한 기사에 주된 책임을 지는 최종작성자 이외에도 부차적인 기여자를 기술하여 신뢰성

을 확보할 필요가 있다. 예를 들어 의견 첨부 번역을 통해 기사를 작성하는 경우 번역과 의견 첨부를 통해 새롭게 작성된 기사에 부가적인 공헌을 한 사람과 공헌의 종류를 기술할 수 있다. 이를 위해 NewsML의 요소 Contributor를 활용할 수 있을 것이다.

이와 함께 뉴스를 발행한 제공자에 대한 정보를 제공할 필요가 있다. 이 요소는 여러 기관이 자원을 공유할 때 필요하다. 예를 들어 KISTI에서 새로운 기사를 작성할 때 KISTI가 제공자가 되는 것이다. 이를 위해 NewsML에서는 Provider라는 요소를 정의하고, 명확하게 기술하도록 하고 있다. 이것은 또한 최종이용자에게는 신뢰성 판단의 근거가 될 수 있다.

• Copyright와 UsageRight 관련 요소

기술동향정보의 각각의 뉴스에 대한 저작권이 관리되어야 한다. 이는 기술동향정보의 신뢰성과 검증성을 확보하기 위한 것이며, 향후 저작권 보호가 강화될 것이 예상되므로 제안될 기술동향정보의 요소는 저작권과 관련된 요소를 충분히 고려해야 한다.

기술동향정보 작성 기관에서 제공하는 뉴스 콘텐츠에 대해서 저작권 정보가 명시되지 않으면, 콘텐츠 제공시점에 이미 저작권 침해의 단초를 제공하고 있는 것이다. 또한 기술동향정보 기관이 작성된 뉴스를 외부로 제공하는 경우, 해당 콘텐츠에 대한 소유권이 누구에게 있으며, 그것이 발효된 날짜, 그리고 해당 콘텐츠를 이용할 수 있는 권한에 대한 상세한 정보를 명확히 제공해야 한다. 이러한 정보를 제공함으로써 도용과 무단 전제 등을 예방하는 역할을 할 수 있다. 이를 위해 NewsML에는 저작권을 위한

Copyright의 CopyrightHolder, CopyrightDate 요소와 사용권한을 위한 UsageRight의 RightHolder, UsageType, Geography, Limitations, StartDate, EndDate 요소가 있으며 이러한 요소들을 도입할 필요가 있다.

(2) 제외할 요소

기술동향정보 요소에서는 NewsComponent와 ContentItem에 속하는 요소 중에서도 전송과 취재한 기사의 작성, 편집 및 관리를 위한 요소, 사용자 임의사항을 위한 요소 등은 고려하지 않는 것이 바람직하다. 이를 구체적으로 살펴보면 다음과 같다.

• NewsLines에 포함된 일부요소

구조화계층인 NewsComponent의 기능은 뉴스 개체를 담는 수용체 역할을 하며, 구성요소들과 연관있는 메타데이터를 첨부하여 공유하도록 하며, 구성요소에 관한 정보를 사람들이 읽을 수 있도록 하는 NewsLines를 제공한다.

NewsLines의 설계목적은 콘텐츠와 함께 제공되는 메타데이터의 일부를 사람이 읽고 이해할 수 있는 것으로 제공하지는 취지이다. 여기서 사람은 최종이용자뿐만 아니라 뉴스가 생산되어 유통되는 전과정에 참여하는 사람들이 포함된다. NewsComponent에 속하는 요소 중 실제 관리를 위한 콘텐츠에 관한 메타데이터는 AdministrativeMetadata, RightsMetadata, DescriptiveMetadata, ContentItem이라고 할 수 있다. 따라서 NewsLines의 요소는 다른 메타데이터 요소와 중복되는 요소를 가지고 있다.

기술동향정보 관점에서는 이러한 취지의 NewsLines를 유지해서 중복값을 유지하는 것보다는

하나의 고유값을 유지하는 것이 바람직하기 때문에 NewsLines를 원칙적으로 제외하되 다른 범주와 겹치지 않는 HeadLine, SubHeadLine, KeywordLine, SeriesLine, DateLine 등의 요소를 검토대상으로 하는 것이 바람직하다.

한편 NewsLines에 속하는 요소 중 Slug-Line은 다른 범주의 메타데이터와 중복되지는 않지만, 통신사에서 신문사에 기사를 전신타자기(teletypewriter)로 송신할 때 사용되는 요소이다. 통신사에서 통신기자가 취재하는 대로 즉시 내용을 기사로 작성하여 본사나 지국에 제1보를 알리고 나서, 계속 취재를 하면서 제2, 3보를 보내게 된다(차배근 1991). 그러나 기술동향정보에서는 통신사로부터 직접 제1,2,3보의 기사를 전달받아 기사를 작성하기보다는 이미 작성이 완료된 기사를 주로 대상으로 하기 때문에 이 요소는 기술동향정보의 성격에 부합되지 않는다.

• Role, BasisForChoice, FileName, Encoding OfInterestTo, TopicOccurrence, Comment, Property

취재한 뉴스의 작성과 편집에 초점을 둔 Role, BasisForChoice는 기술동향정보 서비스의 성격에 부합되지 않기 때문에 제외되어야 한다. Role은 뉴스 개체가 어떤 역할을 하는지를 기술하기 위한 요소이며, BasisForChoice는 대체가 가능한 뉴스나 뉴스 구성요소사이에서 선택의 기준으로 사용되는 정보를 기술하기 위한 요소이다. 예를 들어 하나의 뉴스에 대해 송신된 여러 장의 사진 중 뉴스 작성시 어느 것을 선택할 것인가에 대한 정보를 제공한다.

FileName과 Encoding 역시 작성된 뉴스의

물리적인 위치와 데이터 파일의 부호화의 알고리즘을 관리하기 위한 요소이다. 일반적인 독자를 대상으로 하는 NewsML에서 OfInterestTo 요소는 나이에 의해 대상 독자를 구분하기 위한 것이지만, 기술동향정보는 과학분야의 특정 이용자를 주로 대상으로 하기 때문에 이러한 구분이 부적합하다. 또한 TopicOccurrence는 주제 이외의 인물, 사건 등 토픽의 수를 통제하기 위한 요소이다. 또한 기사 작성과 편집, 관리를 위한 추가 설명을 기술하는 Comment, 임의사용을 위한 Property 요소도 제외하는 것이 바람직하다.

5. 기술동향정보 메타데이터

이 절에서는 이 연구에서 선정한 기술동향정보의 메타데이터 요소와 스킴을 설명하고자 한다. 이 연구에서 선정한 메타데이터는 앞 절의 국내 기술동향정보의 적용 요소와 NewsML 요소의 비교 분석 결과가 반영된 것이다.

선정된 요소의 요소명, 설명, 필수/해당시 필수/선택, 반복가능여부는 국내 기술동향정보와 NewsML의 사례를 참고하였다. 또한 인코딩 스킴은 국내기술동향정보와 NewsML, NewsML-KS에서 제안한 스킴을 제시하거나, 필요한 경우 부분적으로 보완하거나 새롭게 설계하였다. 기술동향정보 메타데이터에서 제공된 스킴은 ISO8601: 2000 basic format, ISO3166-alpha2, ISO639-1, RFC 3066, Genre(IPTC NewsML), MediaType(IPTC NewsML), IMT, Subject-Code(NewsML-KR), 본 연구에서 정의한 MaterialType 등이다.

이 연구에서 제안된 기술동향정보 요소는 설명 메타데이터(DescriptiveMetadata) 범주, 내용 메타데이터(ContentItemMetadata) 범주, 관리 메타데이터(AdministrativeMetadata) 범주, 권한 메타데이터(RightsMetadata) 범주 등으로 나눌 수 있고, 각 범주에 대해 자세히 살펴보면 다음과 같다.

5. 1 설명메타데이터 범주

설명메타데이터(DescriptiveMetadata) 범주는 기술동향정보의 내용을 설명하기 위한 정보를 제공한다. 설명메타데이터 범주에 포함되는 요소들은 기술동향정보의 내용에 대한 요소

로서, 이용자들에게 기술동향정보에 대한 상세한 정보를 제공할 뿐만 아니라 검색과 통합을 위해 활용될 수 있다. 설명메타데이터 범주의 요소는 <표 4>와 같다.

설명메타데이터 범주에서 정의된 요소 중 주제코드는 멀티코드를 유지한다. 주제코드에는 NewsML-KR에서 정의한 주제코드(Subject-Code)를 사용하여 기술하는데, 주제코드에서 정의된 과학/테크놀러지, 보건위생, 경제/산업 분야를 주로 사용한다. 또한 로컬에서 사용하는 주제코드는 로컬주제코드(SubjectCodeLocal)에 기술할 수 있다. 예를 들어 KISTI는 주제코드와 함께 현재 사용하고 있는 주제코드, 차세대 기술코드 등을 로컬주제코드에 기술할 수 있다.

<표 4> 설명메타데이터 범주의 요소

요소	필수/선택	설명	스킴
표제 (HeadLine)	필수	기술동향정보에 부여된 이름(표제)을 기술한다.	
부표제 (SubHeadLine)	해당시필수	기술동향정보에 부여된 표제이외의 부차적인 이름(표제)을 기술한다.	
주제코드 (SubjectCode)	필수	기술동향정보에서 취급된 주제를 기술한다. 이 요소는 인코딩 스킴으로 NewsML-KR의 SubjectCode를 사용한다.	NewsML-KR의 SubjectCode
로컬주제코드 (SubjectCodeLocal)	해당시필수	기술동향정보에서 취급된 주제를 NewsML-KR의 주제코드가 외에 로컬에서 별도로 부여하는 주제코드에 의해 기술한다.	
키워드 (Keyword)	필수	기술동향정보를 설명하는 키워드를 기술한다.	
연재물 (Series)	해당시필수	기술동향정보가 연재기사인 경우, 시리즈명과 시리즈 번호를 기술한다.	
언어 (Language)	선택	기술동향정보의 본문에 사용된 언어를 기술한다. 이 요소는 인코딩 스킴으로 ISO639-1, RFC 3066을 사용한다.	ISO639-1, RFC 3066
장르 (Genre)	선택	기술동향정보의 성질이나 특징을 기술한다. 예를 들어 분석기사, 특집 등이 해당된다. 이 요소는 인코딩 스킴으로 IPTC NewsML의 Genre를 사용한다.	IPTC NewsML의 Genre
사건발생일자 (DateEvent)	선택	기술동향정보에서 취급된 기술동향 사건이 언제 발생했는지에 대한 정보를 기술한다.	ISO8601: 2000 basic format
사건발생장소 (LocationEvent)	선택	기술동향정보에서 취급된 기술동향 사건이 어디서 발생했는지에 대한 정보를 기술한다. 이 요소는 인코딩 스킴으로 ISO3166-alpha2를 사용한다.	ISO3166-alpha2

5. 2 내용메타데이터 범주

내용메타데이터(ContentItem) 범주에는 요소 식별자를 이용해서 기술동향정보에 직접 접근하거나 식별하기 위한 정보를 기술한다. 또한 이것이 텍스트인지 아니면 멀티미디어 데이터 인지를 인식하기 위한 부가정보를 포함한다. 미디어유형, 형식, 크기는 디지털 형식의 기술동향정보에 대한 한정된 검색을 가능하게 한다. 내용메타데이터 범주에서 해당시필수요소로는 식별자, 이미지식별자, 미디어유형, 형식이다. 이런 요소들이 해당시필수요소인 이유는 개별기관에서 기술동향정보 서비스가 단순링크만을 제공하는 경우, 이런 요소들의 값을 기술하지 않고 다음 범주에 나올 출처에 관련 정보를 기술하기 때문이다. 내용메타데이터 범주에 속하는 요소는 <표 5>와 같다.

5. 3 관리메타데이터 범주

관리메타데이터(AdministrativeMetadata) 범주는 기술동향정보의 작성자와 출처에 관한 정보를 관리하기 위해 사용한다. 이 범주에서 제공되는 작성자, 기여자, 발행자, 출처와 관련된 요소들은 대개 기술동향정보의 신뢰성을 보장하기 위한 요소들이다. 관리메타데이터 범주에 대한 요소는 <표 6>과 같다

5. 4 권한메타데이터 범주

권한메타데이터(RightsMetadata) 범주는 기술동향정보의 저작권과, 저작권 소유자가 다른 단체에게 위임하였거나 사용할 수 있는 권한에 관련된 정보를 제공한다. 이는 기술동향정보의 신뢰성과 검증성을 확보하기 위한 것이며, 저작

<표 5> 내용메타데이터 범주의 요소

요소	필수/선택	설명	스킴
식별자 (Identifier)	해당시필수	기술동향정보에 직접 접근하거나 식별하기 위한 정보를 기술한다. 이 요소는 해당 기술동향정보에 직접 연결하기 위한 연결정보, 즉 URL, URN을 기술하기 위한 요소이다.	
이미지식별자 (IdentifierImage)	해당시필수	이미지와 같이 기술동향정보의 구성 요소에 직접 접근하거나 식별하기 위한 정보를 기술한다. 이 요소는 해당 기술동향정보의 구성요소에 직접 연결하기 위한 연결정보, 즉 URL, URN을 기술하기 위한 요소이다.	
미디어유형 (MediaType)	해당시필수	기술동향정보의 미디어유형을 기술한다. 예를 들어 텍스트, 사진 등이 올 수 있다. 이 요소의 인코딩 스킴은 IPTC NewsML의 MediaType을 사용한다.	IPTC NewsML의 MediaType
형식 (Format)	해당시필수	기술동향정보에 대한 기술적인 데이터 유형을 기술한다. 이 요소는 기술동향정보에 접근하기 위해 필요한 소프트웨어를 식별하는데 사용한다. 이 요소의 인코딩 스킴은 IMT(Internet Media Type) 리스트로부터 선택하여 기술한다.	IMT 확장
크기 (Size)	선택	기술동향정보의 바이트 수를 기술한다. '0'에서 '9'까지만 사용한다. MBytes, GB 등은 허용하지 않는다. 이 요소는 기술동향정보의 압축 버전의 크기가 아니라 실제크기를 말한다.	

〈표 6〉 관리메타데이터 범주 요소

요소	필수/선택	설명	스킴
작성자 (Creator)	필수	기술동향정보의 내용을 작성한 개인이나 단체 등의 정보를 기술한다. 단 기술동향정보가 단순링크 서비스인 경우 원뉴스의 작성자를 기술한다.	
작성자직책 (CreatorTitle)	선택	작성자의 신분에 관한 정보를 기술한다. 예를 들어 기자, 교수 등이 해당된다.	
작성자소속기관 (CreatorOrganization)	선택	작성자의 소속기관을 기술한다.	
작성자최종학위 (CreatorDegree)	선택	작성자가 취득한 최종학위를 기술한다. 예를 들어 학사, 석사, 박사가 해당된다.	
작성자전공분야 (CreatorMajor)	선택	작성자의 전공분야를 기술한다. 작성자가 취득한 학위의 전공분야 혹은 현재의 전공분야를 기술한다.	
작성자의 이메일주소 (CreatorEmail)	해당시필수	작성자의 이메일 주소를 기술한다.	
기여자 (Contributor)	선택	Creator 요소와 관련을 가지며 Creator 요소가 일차적인 책임을 갖는 중요한 작성자를 기술하기 위한 요소라면, 이 요소는 부차적 책임을 갖는 기여자를 기술하기 위한 요소이다.	
기여자기여유형 (ContributorContribution)	선택	기술동향정보 향상에 기여한 종류를 알려주는 요소이다. 예를 들어 편집, 감수, 개발, 출판 등이 올 수 있다.	
발행자 (Publisher)	필수	기술동향정보를 발행한 기관을 기술하기 위한 요소이다. 특정 기관에서 관리하는 모든 기술동향정보를 찾을 수 있도록 한다.	
출처제공자 (SourceProvider)	필수	기술동향정보 작성시 참고한 자료(예: 원뉴스, 통계, 그래프 등)를 공급한 개인이나 단체를 기술하기 위한 요소이다.	
출처명 (SourceName)	필수	참고한 자료명을 기술하기 위한 요소이다.	
출처식별자 (SourceIdentifier)	해당시필수	참고한 자료가 온라인 자료인 경우, 참고한 자료에 직접 접근할 수 있는 정보를 기술한다.	
출처일자 (SourceDate)	필수	참고한 자료의 발행일자를 기술한다.	ISO8601: 2000 basic format
출처언어 (SourceLanguage)	필수	참고한 자료의 언어를 기술한다.	ISO639-1, RFC 3066
출처발생장소 (SourceLocation)	필수	참고한 자료의 발생장소를 기술한다.	ISO3166-alpha2
출처자료유형 (SourceMaterialType)	선택	참고한 자료의 유형을 기술하는 요소이다. 예를 들어 학술지, 뉴스 등이 해당된다. 이 요소의 인코딩 스킴은 Material-Type을 사용한다.	MaterialType
작성일자 (DateCreated)	필수	기술동향정보의 작성일자를 기술하기 위한 요소이다.	ISO8601: 2000 basic format
수정일자 (DateModified)	해당시필수	기술동향정보의 수정일자를 기술하기 위한 요소이다.	ISO8601: 2000 basic format

〈표 7〉 권한 메타데이터 범주의 요소

요소	필수/선택	설명	스킴
저작권자 (CopyrightHolder)	필수	기술동향정보의 저작권자를 기술한다. 즉 누가 기술동향정보 개체를 소유하고 있는가를 알려준다.	
저작권일자 (CopyrightDate)	필수	저작권의 일자를 기술하기 위한 요소이다.	ISO8601: 2000 basic format
이미지 저작권자 (ImageCopyrightHolder)	해당시필수	기술동향정보에 포함된 이미지의 저작권자를 기술한다.	
이미지 저작권일자 (ImageCopyrightDate)	해당시필수	기술동향정보에 포함된 이미지의 저작권 일자를 기술한다.	ISO8601: 2000 basic format
사용권자 (UsageRightHolder)	해당시필수	기술동향정보의 사용권한을 누가 가지고 있는지를 기술한다.	
사용형태 (UsageRightType)	선택	사용권한의 사용 형태를 기술한다. 예를 들어 TV 용이 해당된다.	
사용권한지역 (UsageRightGeography)	선택	사용권한이 있는 특정 지역을 기술한다.	ISO3166-alpha2
사용권한제한 (UsageRightLimitations)	선택	기술동향정보의 내용을 사용하는데 적용되는 어떤 제한이 있는지를 기술한다.	
사용권한시작일자 (UsageRightStartDate)	선택	사용권한의 사용기간(시작일자)을 기술한다	ISO8601: 2000 basic format
사용권한종료일자 (UsageRightEndDate)	선택	사용권한의 사용기간(종료일자)을 기술한다	ISO8601: 2000 basic format

권을 보호하기 위한 요소이다. 권한 메타데이터 범주에 해당되는 요소는 〈표 7〉과 같다

6. 결 론

기술동향정보는 연구개발 활동과 과학기술 정책의 결정과정에서 매우 중요한 요소이다. 그러나 국내에서 서비스되고 있는 기술동향정보의 사례를 보면, 기술동향정보의 검색을 위한 충분한 메타데이터가 마련되어 있지 않으며, 기관마다 다양한 메타데이터가 존재한다. 이러한 이유로 기술동향정보의 검색이 비효율적이고, 데이터 공유가 어려워지는 결과를 초래하며, 비용낭비로 이어지고 있다.

이 연구는 기술동향정보를 위한 메타데이

터 요소와 스키마를 제안하기 위한 것이다. 이 연구에서는 기술동향정보의 특성을 반영하기 위하여 국내 서비스 사례에서 사용된 요소와 뉴스분야에서 표준적 위치에 있는 NewsML 을 고려하여 기술동향정보 메타데이터를 제안하였다.

이 연구에서 제안된 기술동향정보 메타데이터가 NewsML을 고려한 이유는 먼저 기술동향정보가 속보성이라는 특성 때문에 대부분 뉴스의 형태를 띠고 있기 때문이다. 따라서 기술동향정보의 요소 선정시 뉴스분야의 경험이 반영된 NewsML을 고려하는 것이 바람직하다. NewsML 이 향후 뉴스관련분야의 표준으로 확산될 경우 유사한 요소를 유지함으로써 보다 신속하고 효율적인 서비스를 제공할 수 있을 것이다.

다음으로, 기술동향정보 메타데이터는 기술

동향정보를 충분히 기술하여 효과적으로 검색할 수 있는 요소를 마련하고, 특히 저작권 등을 고려한 체계적인 메타데이터를 제공해야 하는데, NewsML에서 제공되는 일련의 메타데이터는 이러한 요구에 적합하기 때문이다. 또한 NewsML에서 제공되는 다양한 토픽셋을 활용하여 뉴스 검색과 관리가 더 효과적으로 이루어질 수 있고, 효율적인 데이터 공유가 가능하기 때문이다.

제안된 기술동향정보 메타데이터는 국내기술동향정보 서비스에서 제공되는 요소는 유지하면서, NewsML의 전송과 업무처리 요소를 제외한 요소들을 추가하였다. 기술동향정보 메타데이터는 설명메타데이터(DescriptiveMetadata), 내용메타데이터(ContentItemMetadata), 관리메타데이터(AdministrativeMetadata), 권한메타데이터(RightsMetadata) 등의 범주로 구분하였다.

이 연구에서 제안된 기술동향정보 메타데이터는 국내 기술동향정보가 다루는 자료 유형과 서비스 유형을 고려해서 제안되었기 때문에, 동향정보에 관한 새로운 기사의 작성, 이미 작성된 기사의 링크제공, 분석을 첨부한 번역과 같은 가공 기사 등 다양한 관점에서 사용가능하도록 제안되었다.

또한 기존 적용 메타데이터에 작성자, 출처를 비롯한 다양한 정보에 대한 요소를 추가함으로써 검색의 효율을 높이고, 정보의 신뢰성과

정확성, 검증가능성을 확보하도록 하였다 특히 국내 기술동향정보 메타데이터는 저작권 측면에서 미비한 점이 많았으나, 본 연구에서 제안한 저작권 관련 요소들은 기술동향정보의 활용도를 높이면서 기술동향정보의 저작권을 보호하도록 제안되었다.

이 연구에서 제안된 기술동향정보 메타데이터는 기술동향정보의 검색을 위해 효과적으로 활용될 수 있으며, 기술동향정보 관련 데이터를 통합적으로 공유할 수 있는 체제 마련에 기초를 제공하였다고 볼 수 있다.

이 연구 결과는 앞으로 국내의 실정에 맞는 기술동향정보 메타데이터 표준을 개발하기 위한 기초자료로 활용될 수 있을 것이다. 표준화된 기술동향정보의 메타데이터를 사용하게 되면, 기술동향정보의 계량적 분석, 평가, 예측 등에 활용이 가능하고, 동향정보 메타데이터 표준화를 통한 데이터 모델 재활용 및 중복 구축 방지 등 효율적이고 다각적인 R&D 정보 활용 및 공유 체제의 구축이 가능할 것이다. 궁극적으로 국가 R&D 메타데이터 레지스트리의 효율적 구축, 이용 및 관리에 활용할 수 있을 것으로 기대된다. 향후 기술동향정보 서비스를 위해 기술동향정보의 메타데이터의 표준화와 함께, 저작권과 이용권리에 대한 문제를 보다 명확하게 할 필요가 있으며, NewsML을 활용한 기술동향정보 구축과 제공 전반에 대한 개선방향을 검토해 볼 필요가 있다.

참 고 문 헌

- 오수정. 2004. 인터넷: NewsML기반의 뉴스 전송 포맷 표준화 모델 개발. 『신문과방송』, 11: 155-159.
- 오수정. 2005. 인터넷 뉴미디어: 카인즈 이렇게 달라졌다. 『신문과방송』, 3: 145-149.
- 차배근. 1991. 『커뮤니케이션學概論, 下』 서울: 世英社.
- 최성배, 한선화. 2001. 해외과학기술동향 정보서비스 및 자료평가시스템. 『지식정보인프라』, 7: 94-104.
- 한국언론재단. 2004. 『뉴스전송 표준화 모델 개발 연구보고서』. 한국언론재단: 동재단 와이즈미디어(주).
- 한국온라인신문협회. 2005. “디지털뉴스 이용규칙.” [cited 2005.8.4].
 <<http://www.donga.com/docs/intro/2004/FAQv2.pdf>>
- 황정근. 2004. 류쿠신보의 NewsML 사진관리 시스템 구축. 『Media Worldwide』 6: 69-73.
- 『미디어오늘』. 2005. 인터뷰 업그레이드 카인즈 개발자 김명기 와이즈미디어 사장. 1월 26일. [cited 2005.6.20].
 <<http://www.mediatoday.co.kr>>.
- Compton, D. 2005 “NewsML: Markup in the real world.” [cited 2005.7.15].
 <<http://www.newssummit.org/2005/presentations/0505-NSS-NewsML-Reuters.pdf>>.
- EventsML Home. 2005. [cited 2005.6.30].
 <<http://www.iptc.org/EventsML/>>.
- Fujiwara, T. and K. Tanaka. 2003. “NewsML in Japan: how it is used and issues.” [cited 2005.7.18].
 <<http://www.newssummit.org/2003/slides/fujiwara&tanaka.ppt>>.
- IPTC NewsNL Home. 2005. [cited 2005.8.29]. <<http://www.iptc.org/>>.
- Le Meur, L. 2005. “AFP: NewsML case study...Magazine Forum”. [cited 2005.6.8].
 <<http://www.newssummit.org/2005/presentations/AFP2005-06-NewsMLCaseStudy.pdf>>.
- Lorenzen, J. 2005. “NewsML and Consumer Acceptance: A case study: BUSINESS WIRE.” [cited 2005.7.1].
 <http://www.newssummit.org/2005/presentations/20050524_nss_bw_case_study.pdf>.
- ProgramGuideML Home. 2005. [cited 2005.7.20].
 <<http://www.programguideml.org/>>.
- SportsML Home. 2005. [cited 2005.8.3].
 <<http://www.sportsml.org/>>.
- Wong, K. 2004. “Turning innovation into a reality”. [cited 2005.7.20].
 <<http://cnewsml.org/mission/Mission.ppt>>.

КСІ