

RSS 기반 과학기술정보 배급표준 시스템(STI-RSS)의 개발과 활용방안에 대한 연구

A Study on the Implementation and Service of the RSS Based Science and Technology Information Syndication System(STI-RSS)

권 이 남* · 김 재 수** · 신 동 구*** · 전 성 진**** · 정 택 영*****

Lee-Nam Kwon · Jae-Soo Kim · Dong-Gu Shin · Sung-Jin Jhun · Teak-Young Jeong

차 례

1. 서 론	4. STI-RSS 활용방안
2. RSS와 정보배급 프로토콜	5. 결 론
3. RSS 기반 과학기술정보 배급표준 시스템(STI-RSS) 개발	• 참고문헌

초 록

최근 이메일 서비스의 보완 수단으로서 RSS(Really Simple Syndication)를 이용한 정보 서비스가 활성화되고 있다. 과학기술 분야의 정보는 매우 방대하고 다양하며, 정보의 발생주기 및 서비스 또한 복잡한 특성을 가지고 있다. 단일 정보를 제공하는 기관인 경우 RSS를 이용한 정보 서비스는 정보를 제공하는 하나의 채널을 확보하는 의미에서 구현 시 고려해야 할 요소가 많지 않지만, 다양하고 방대하며 전문적인 과학기술정보의 특성을 반영한 RSS 서비스는 보다 정교한 설계와 구현을 필요로 한다. 본 연구에서는 RSS와 다른 정보배급 프로토콜을 비교하고, 국내외 RSS 동향조사를 수행하였으며, 과학기술정보의 특성을 고려한 RSS 기반 과학기술정보배급 시스템(STI-RSS)의 구현을 고찰하고 이의 활용방안을 제시하고자 한다.

키 워 드

RSS, 과학기술정보, 신디케이션, 상호운용성

* 한국과학기술정보연구원 표준화기술지원실 선임연구원
(Senior Researcher, Technical Support for Standardization Dept., KISTI, ynkwon@kisti.re.kr)
 ** 한국과학기술정보연구원 표준화기술지원실 책임연구원
(Principal Researcher, Technical Support for Standardization Dept., KISTI, jaesoo@kisti.re.kr)
 *** 한국과학기술정보연구원 표준화기술지원실 선임연구원
(Senior Researcher, Technical Support for Standardization Dept., KISTI, lovesin@kisti.re.kr)
 **** 한국과학기술정보연구원 표준화기술지원실 연구원
(Researcher, Technical Support for Standardization Dept., KISTI, sjjhun@kisti.re.kr)
 ***** 한국과학기술정보연구원 표준화기술지원실 책임연구원, 표준화기술지원실장
(Principal Researcher, Director, Technical Support for Standardization Dept., KISTI, tychung@kisti.re.kr)
 • 논문접수일자 : 2005년 11월 14일
 • 게재확정일자 : 2005년 12월 9일

ABSTRACT

Information services adopting RSS(Really Simple Syndication) is on an increase. It is replacing the existing email services. Information in the area of science and technology is of vast amount and various area and their life-cycle and service is complex. Institutions providing information of a single area are not faced with many obstacles in designing information services using RSS in acquiring a channel for the service. However, RSS service in the area of science and technology has to be more specific in designing and implementing the services. In this study, RSS is compared with other information service protocols, domestic and international trends in the area of RSS is analyzed to implement the RSS based Science and Technology information service system(STI-RSS) and its service.

KEYWORDS

RSS(RDF Site Summary, Really Simple Syndication, Rich Site Summary), Science and Technology Information, Syndication, Interoperability

1. 서 론

정보이용자가 개별 웹 사이트를 방문하지 않고도 원하는 시점에 원하는 정보만을 한 화면에 최신 정보만을 제공받을 수 있는 풀(Pull) 방식의 표준화된 포맷의 RSS(Really Simple Syndication) 서비스가 부각되고 있다.

정보 서비스 기관이 이용자에게 정보를 제공하기 위해 알려진 정보배급(Content Syndication) 기술로서 기존의 이메일 푸시(Push) 방식 기술을 들 수 있다. 이용자가 선택한 관심주제나 분야에 대한 정보를 제공하거나, 이용자가 직접 작성한 검색식에 대한 최신 맞춤정보 제공 서비스는 기존 이메일보다 한층 업그레이드된 이용자중심의 서비스라고 볼 수

있지만, 이용자가 원하지 않는 시점에 원하지 않는 정보도 함께 제공받으므로 다분히 공급자적 측면의 기술이라고 볼 수 있다.

RSS는 정보 서비스 기관과 이용자 간의 매일 정보제공 방식 서비스를 대체할 수 있는 새로운 패러다임의 정보배급 기술로서, 이용자가 정보를 요청할 때 제공하는 이용자 측면의 풀 방식 기술이다. RSS를 이용한 RSS 서비스는 정보이용자가 자신이 관심 있는 홈페이지를 일일이 방문하거나 자신의 개인정보를 제공하지 않고도 여러 개의 관심 있는 홈페이지에서 발생하는 신규 정보를 하나의 화면에서 스팸 없이 한꺼번에 볼 수 있는 서비스이다. 즉, RSS가 이용자 측면의 정보배급 기술이라는 것은 RSS 서비스를 통해 이용자에게 정보제공

시점과 정보제공 범위, 개인정보제공 등의 선택을 정보제공기관이 아닌 이용자가 직접 선택할 수 있도록 권한을 부여한다는 점이다.

RSS 기반 과학기술정보 배급표준 시스템(STI RSS : Standard system for the syndication of Science and Technology Information based on RSS, 이하 STI RSS)은 이러한 배경에서 이용자 측면의 과학기술 정보배급 서비스를 제공하기 위해 구현된 시스템이다. 과학기술 분야의 정보는 매우 방대하고 다양하며, 정보의 발생주기 및 서비스 또한 복잡한 특성을 가지고 있다. 단일 정보를 제공하는 기관인 경우 RSS를 이용한 정보 서비스는 정보를 제공하는 하나의 채널을 확보하는 의미에서 구현 시 고려해야 할 요소가 많지 않지만, 다양한 과학기술정보의 특성을 반영한 RSS 서비스는 보다 정교한 설계와 구현을 필요로 한다.

과학기술 분야 지식정보는 다양한 정보 소스가 수시로 발생하며, 이에 따른 복잡하고 다양한 서비스는 상호운용성을 유지하면서 유연하게 대응할 수 있어야 한다. 만약, 각각의 다양한 정보 채널별로 RSS 정보배급 시스템이 표준화되어 있지 않다면 과학기술 분야의 다양한 환경에 따른 서비스별 개발중복의 문제점뿐만 아니라, 각 서비스별 콘텐츠의 통합과 연계 및 상호호환, 채널별 통합 관리를 위해서 많은

시간과 노력을 투자해야만 하는 문제가 발생하게 된다. 다양하고 전문적인 과학기술 분야 지식정보를 RSS 서비스로 제공하기 위한 시스템의 개발은 분야별 다양한 과학기술 분류 및 다양한 정보의 생성주기, 다양한 과학기술정보 서비스의 특성 등을 고려하여야 한다.

본 연구에서는 RSS와 다른 정보배급 프로토콜을 비교하여 RSS의 특징을 도출하고, 국내외 RSS 동향조사를 수행하였으며, 과학기술 정보의 특성을 고려한 STI RSS의 구현을 고찰하고 이의 활용방안을 제시하고자 한다.

2. RSS와 정보배급 프로토콜

2.1 RSS의 개요

RSS는 뉴스나 블로그(Blog)¹⁾ 등 자주 갱신되는 성격의 사이트를 위한 XML 기반의 포맷으로 웹 사이트끼리 간단한 정보를 서로 자료를 주고 받기 위한 정보배급 표준이다(Ben Hammersley 2003). “RDF Site Summary”, “Really Simple Syndication”, “Rich Site Summary” 등의 약자로 사용되며, 다양한 웹 사이트 상의 콘텐츠를 요약하고, 상호교환할 수 있도록 만든 표준 포맷이다(전종홍 2004). 웹 사이트에서 발생하는 모든 종류의 정보를 공유하는 데 사용되고 있으며, 뉴스, 신규 업체

1) 초보자도 쉽게 생성할 수 있는 개인 홈페이지 성격으로 ‘웹에 기록하는 일지’, 작성 내용이나 형식에 제한이 없다(김중태 2004).

이트 정보, 과학기술 지식정보, 이벤트, 콘텐츠 모음, 상품 정보 등 분야에 제한없이 많은 정보들을 표현할 수 있다.

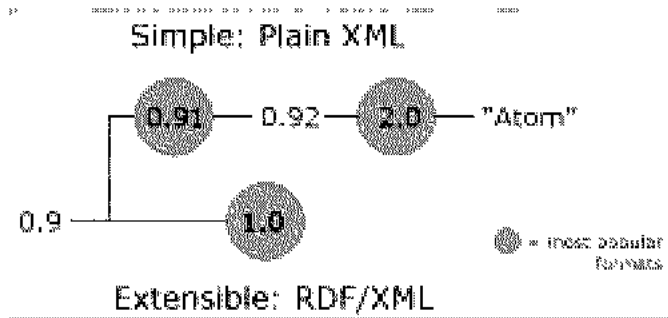
예를 들어, A라는 사이트가 B라는 사이트의 RSS 파일을 정기적으로 수집하면 B사이트의 갱신된 콘텐츠 제목, 링크, 요약 내용을 자동화된 과정에 의해 자신의 사이트에 올려놓을 수도 있고, 이용자는 B사이트를 직접 방문하지 않고도 RSS 구독기 프로그램을 사용해 최신 정보를 편리하게 볼 수 있다. RSS 구독기는 RSS 파일을 읽어 해당 사이트 정보를 표시해주는 프로그램으로 주로 데스크톱 이용자들이 사용한다. 대표적인 RSS 구독기로 국산은 연모, Xpyder, 인터넷 도사 등이 있으며, 외산 제품으로는 FeedException, SharpReader 등이 있다.

RSS 표준은 네스케이프 사에 의해 처음 개발되어, 현재 블로그 커뮤니티에 가장 많이 적용되고 있으며, 가장 널리 활용되고 있는 공개

된 웹 콘텐츠 신디케이션²⁾ 표준으로 인정받고 있다. RSS 파일을 상호교환하게 되면 최신 정보의 신디케이션을 간단하게 구성할 수 있을 뿐만 아니라, RSS 구독기가 있는 웹 이용자라면 즐겨 찾는 사이트를 각각 방문하지 않고도 RSS 채널(channel)의 URL 정보만을 등록하는 것으로 웹 사이트의 최신 RSS 피드(feed) 정보를 하나의 구독기 화면에서 한꺼번에 확인할 수가 있다. RSS 채널은 RSS를 지원하는 사이트의 URL 정보이며, RSS 피드는 RSS 표준 형태로 제공하는 각종 정보를 말한다.

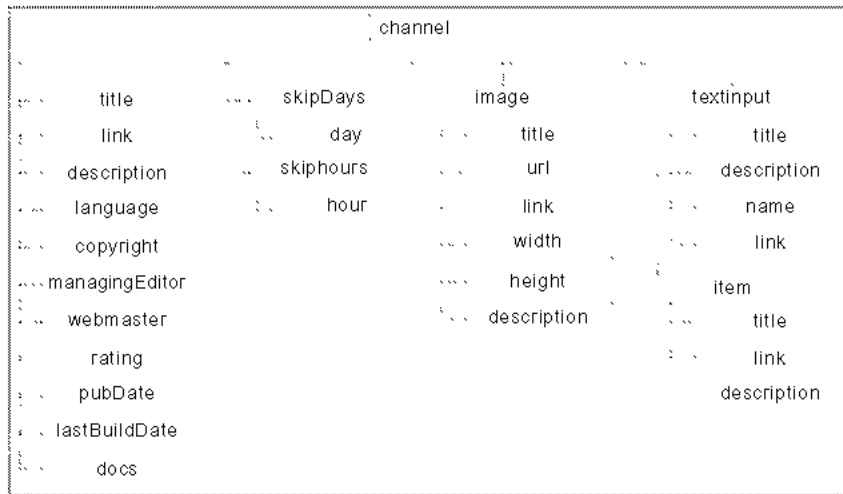
이용 대상자는 주로 일반 데스크톱 구독자와 자신들이 가진 페이지에서 다른 사이트에서 불러온 제목 정보를 추가시키기를 원하는 웹 마스터들이 주로 활용하고 있다.

현재까지 대표적인 RSS 표준은 <그림 1>에 나와 같이 RSS 0.9, RSS 0.91, RSS 0.92, RSS 1.0, RSS 2.0, Atom으로 총 6개의 버전이 있으



<그림 1> RSS의 다른 버전(Tony, Timo, and Ben 2003)

2) 신디케이션(Syndication) : 콘텐츠 제작업체와 최종 이용업체의 중간에서 콘텐츠를 가공·유통하는 설루션을 통해 쌍방의 비용을 절감하는 방식으로 일종의 콘텐츠 B2B 모델이다(백운주 2001).



〈그림 2〉 RSS 0.91의 트리 구조(Ben Hammersley 2003)

며, 가장 많이 활용하고 있는 표준으로는 RSS 2.0, RSS 0.91, RSS 1.0 순이다. 가장 처음 나온 RSS 0.91 표준의 표현 방식은 다음 〈그림 2〉의 트리 구조를 보면 쉽게 이해할 수 있다.

〈그림 2〉에서 보듯이 하나의 RSS 파일을 구성하기 위해 Channel 아래에 title, link, description, language 등의 여러 개의 아이টে็ม으로 구성되며 웹 사이트의 정보를 간략하게 표현할 수 있는 구조이다.

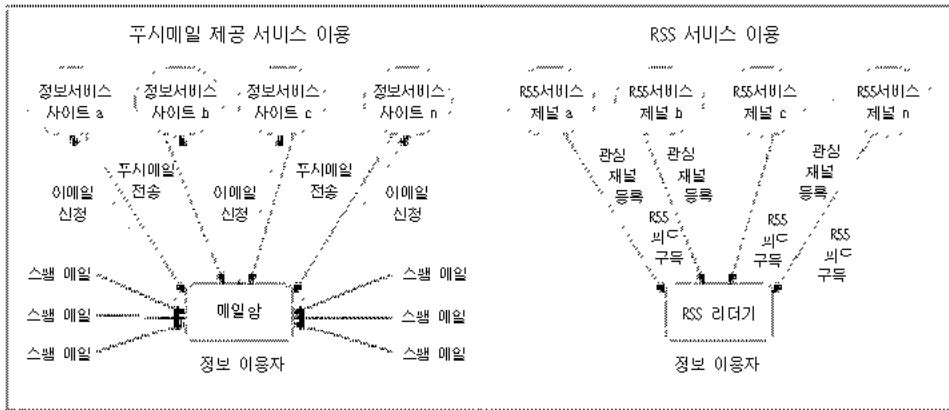
일반적인 RSS의 장점으로는 첫째, 선택적 구독(사용자가 원하는 topic과 정확히 일치하는 채널 선택), 둘째, 빠른 구독(동시에 다양한 채널 소스 접근), 셋째, 이력관리(다양한 채널의 과거 기록들 보관이 가능), 넷째, 자동화된 콘텐츠 연동이 용이, 다섯째, 콘텐츠 재사용성(구조화된 XML 데이터로 손쉬운 변환 및 처리가 가능), 마지막으로 커뮤니케이션 방식의 변화(1:1에서 1:N으로 발전) 등의 장점이 있다(전

종홍 2004).

RSS의 적용은 정보를 제공받는 이용자 뿐만 아니라 정보를 서비스하는 기관 모두에게 유용한 장점을 가지고 있다. RSS 서비스를 통해 정보 서비스 기관은 콘텐츠 제목이나 간단한 요약정보 등을 무료로 제공하는 대신 해당 사이트로 트래픽을 유도할 수 있으며, 서비스 기관의 인지도를 향상시킬 수 있는 수단이 되기 때문이다.

〈그림 3〉은 이용자가 이메일을 통한 푸시 메일 제공 서비스를 이용할 때와 RSS를 통한 서비스를 이용할 때의 차이점을 나타낸 것이다.

이메일이 푸시 방식의 정보전달이라면 RSS 구독 방식은 풀 방식의 정보전달이라고 볼 수 있다. 폐쇄적인 시스템은 대개 내보내기에 역점을 두기 때문에 푸시 방식을 택하는 반면, 공개와 공유를 위한 서비스는 대부분 끌어오기에 역점을 두기 때문에 풀 방식을 취하고 있다. 풀



〈그림 3〉 푸시 메일 제공 서비스 이용과 RSS 서비스 이용의 차이점

방식은 자료의 공유 여부를 이용자의 선택에 맡기는 것이 차이라고 볼 수 있다(김중태 2004). 이용자가 원하는 자료만 끌어오기로 전달받을 수 있게 지원하는 RSS는 이러한 풀 방식의 서비스를 제공하기 위한 정보배급 프로토콜이다.

2.2 RSS와 다른 정보배급 프로토콜

정보배급은 콘텐츠 제작기관과 최종 이용자의 중간에서 콘텐츠를 가공·유통하는 설루션을 통해 쌍방의 비용을 절감하는 방식으로 일종의 콘텐츠 B2B 모델이다. 즉, 콘텐츠 제작기관의 다양한 콘텐츠를 콘텐츠 신디케이터의 수집 및 가공을 통해 적절한 콘텐츠 분배업체로 공급되는 과정이다(백윤주 2001).

RSS는 가장 널리 활용되고 있는 공개된 웹 정보배급 프로토콜이며, RSS를 활용한 애플리케이션으로 Urchin이 있으며, 이외에도 유사

한 ICE(Information Content Exchange) 등이 있다. RSS는 자기 스스로는 아무런 기능을 수행하지 않으나, 그에 따르는 프로세스에 의해 운영되는 가장 기초적인 기술이라 할 수 있다. RSS 기술은 해당 기술에 자료를 맞추는 기존의 기술과는 달리 자유롭게 목적에 맞게 재변형될 수 있다. RSS와 다른 프로토콜의 가장 큰 차이점은 RSS가 B2C 설루션이라는 데 있다. RSS는 비교적 크기 및 기술면에서 가벼우면서 대용량의 자료를 전송하지 않는다. 또한, 브로드캐스트 기능에 기반하고 있으며 형태에 있어서 자유롭게 변형이 가능하여 다른 프로토콜에 비해 자료의 유형에 관대하다.

정보배급 프로토콜로 이미 알려진 ICE는 출판 업계에서 많이 쓰이는 프로토콜이며 OAI PMH(Open Archives Initiative Protocol for Metadata Harvesting)는 디지털 라이브러리 분야에서 많이 쓰는 프로토콜이다. 이 두 가지 프로토콜 모두 데이터 교환을 하기 위한

방법을 지원하며 웹 환경 기반으로 구축되어 RSS와 비슷한 특징을 보인다. RSS와 다른 정보배급 프로토콜의 가장 큰 차이점은 RSS가 B2C 설루션의 형태로서 콘텐츠의 전체 내용에 대한 메타데이터 전송보다는 비교적 가볍고 요약된 메타데이터의 전송을 담당한다는 점에서 ICE 프로토콜과 차이가 있다. 또한 RSS는 브로드캐스트용 포맷으로서 일대일 전송에는 적합하지 않은 형태를 가지고 있다(Tony, Timo, and Ben 2004).

2.2.1 Urchin

NPG에 의해 개발된 서버측 애플리케이션으로서 RSS 피드 및 다른 데이터 소스들의 수집기(aggregator) 및 필터 역할을 수행한다. Urchin은 원래 영국의 JISC(The Joint Information Systems Committee)에 의해 시작되었으나, 이후에 NPG(Nature Publishing Group)의 관할 아래 진행되었다. Urchin은 오픈 소스 애플리케이션으로서 리눅스 기반으로 아파치, MySQL 및 펄(Perl)을 이용하여 개발되었으며, 이후에 다른 유닉스 계열 운영체제에서도 사용이 가능토록 포팅되었다.

Urchin의 기본적인 기능은 RSS, Atom은 물론 기본적인 HTML 페이지 및 DB 등 여러 소스로부터 데이터를 수집하는 데 있다. 이렇게 수집된 정보는 내부적으로 저장되어 있다가 원하는 방식으로 변환되어 출력되게 된다. Urchin은 어떠한 형태의 RSS나 Atom이든 RSS 1.0 형태의 데이터로 표준화하여 변환해준다.

2.2.2 ICE

ICE는 산업 콘텐츠 제공자들과 소프트웨어 벤더들의 협력체인 IDEAlliance에 의해 개발된 프로토콜로서 출판사 및 비영리 콘텐츠 제공자들에게 계획적이고 안정적인 콘텐츠의 재분배를 수행하기 위해 개발된 자동화 시스템이다. Basic ICE는 성능면이나 그 구성에 있어 RSS와 비슷하며 HTTP를 통해 전송받는 형태의 ICE 카탈로그 및 패키징을 지원한다. 반면에 Full ICE는 ICE 패키지의 푸시 및 풀을 모두 지원한다. ICE는 신디케이터와 예약된 구독자 사이에 규정된 비즈니스 룰 및 관계를 관리한다. 또한 여러 종류 콘텐츠의 SOAP(Simple Object Access Protocol) 메시지 포맷으로의 변환을 지원한다. ICE는 B2B 트랜잭션에 맞는 안정성에 초점을 두고 있는 반면, RSS는 보다 쉽게 접근하고 사용할 수 있는 방식이라고 볼 수 있다.

2.3 국내외 RSS 동향

2.3.1 국내

RSS를 이용한 서비스는 해외는 물론, 최근 국내에서도 각 분야에서 다양하고 활발하게 연구 및 응용되고 있다. 여기에서는 최근 국내외의 대표적인 RSS 서비스 현황과 표준화 동향을 살펴보았다. 국내에서는, 초기에 포털 사이트 내에서 블로그끼리 정보교환을 위한 포맷으로 RSS를 활용하는데 그쳤으나, 2004년부터 뉴스 분야 사이트(조선닷컴, 조인스닷컴 등)를

중심으로 RSS 정보배급 서비스를 시작하였다.

2005년에는 보다 다양한 부분에서 RSS를 응용하여 활용하고 있다. 야후코리아의 '야후 피플링'은 RSS 기능을 이용해 여러 블로거들이 자신이 쓰고 있는 글을 주제별로 따로 분류해 등록하고, 해당 주제의 글만 한꺼번에 읽을 수 있도록 지원하고 있다. '다음 RSS넷'은 RSS 기능을 활용해서 정보를 편집하고 구독할 수 있는 일종의 마이페이지 서비스를 제공하고 있으며, SK 커뮤니케이션즈의 '뽕'도 주제에 따라 유사한 관심사를 가진 블로그들을 연결해주는 RSS 서비스의 일종이다. 이는 블로그끼리 정보를 교환하고 연결하기 위한 포맷으로서 RSS를 활용하고 있는 사례이다. 또 다른 RSS 응용사례로는 RSS 피드를 특정 포털 내에서만이 아닌 국내 전체 RSS 피드를 대상으로 검색해주는 '드림위즈 RSS 검색 서비스'가 있다. 이는 RSS 피드들을 서비스 콘텐츠로 규정하고, RSS 피드의 검색 및 신규 피드 정보를 제공하는 일종의 RSS 포털 사이트라고 볼 수 있다.

국내 공공 부문에서의 최근 RSS 서비스로는 단일 정보 소스이기는 하지만, 한국기초과학지원연구원의 최신 연구장비정보의 RSS 피드 제공을 들 수 있으며, 대학도서관으로는 처음으로 성균관대학교 학술정보관의 RSS 피드 서비스 사례를 들 수 있다. 성균관대학교의 RSS 서비스는 신착도서, 게시판, 맞춤형 채널을 제공하며, 특히 검색결과에 업데이트 정보를 RSS로 받기 기능을 통하여 사용자가 지정한 키워드의 맞춤 채널을 등록하고 해당 검색

결과만을 RSS 피드로 제공받을 수 있도록 하고 있다.

과학기술 분야 지식정보를 종합적으로 생산·유통하고 있는 한국과학기술정보연구원(KISTT)은 정보 소스가 다양하고 각 분야의 서비스별로 다양한 기술분류를 활용하고 있는 특성 등을 감안하여 2004년부터 과학기술정보 서비스에 RSS 적용에 관한 활용방안을 연구하였으며, 2005년에 STI RSS를 개발하였고, 이를 다양한 정보 서비스에 활용할 예정이다.

2.3.2 해외

해외에서는, 이미 다양한 분야에서 RSS 서비스를 보편적으로 제공하고 있으며, RSS 피드 및 이용자층에 대한 전문적인 분석을 제공하는 사이트(Syndic8, NewsIsFree, UserLand 등)는 물론, 전문적인 RSS 피드 검색 엔진(Feedster 등)을 제공하고 있다. 즉, 세계적인 RSS 피드의 현황 정보(피드 작성언어 분포, 접속 순위, 인기 순위, RSS 버전 사용분포, 총 피드수, 총 이용자수)를 전문적으로 분석하여 통계 그래프를 제공하여 서비스하고 있다.

〈표 1〉은 해외 분야별로 RSS 피드를 제공하는 주요한 몇 개의 RSS 채널만을 정리한 것이다.

마이크로소프트(MS)는 내년에 출시되는 윈도우 차기 버전인 '비스타'에 RSS 지원기능을 추가할 계획을 발표하였다. 무엇보다, 기존 RSS 버전에서 정의된 아이템 외에 〈msdn : headlineImage〉, 〈category domain〉, 〈msdn

〈표 1〉 해외 분야별 RSS 피드 제공 주요 웹 사이트 목록

분야	RSS 채널 명	RSS 피드 제공 URL
뉴스	Yahoo News	http://rss.news.yahoo.com/rss/topstories
	Google News - World	http://news.google.com/news?ned-us&topic-w&output=rss
	Reuters News	http://www.reuters.com/newsrss.jhtml
	CNET News	http://news.com.com/2009-1090-980549.html?tag=rss
	BBC News	rss://news.bbc.co.uk/rss/newsonline_uk_edition/front_page/rss091.xml
	Washingtonpost	http://www.washingtonpost.com/wp-adv/rss/front.htm
과학기술	Scientific American	http://www.sciam.com/news_directory.cfm
	New Scientist	http://www.newscientist.com/feed.ns?index-online-news
	New York times:Science	http://www.nytimes.com/services/xml/rss/nyt/Science.xml
	PC Magazine	rss://rssnewsapps.ziffdavis.com/pcmag.xml
	Computerworld	rss://www.computerworld.com/news/xml/0,5000,54,00.xml
	Slashdot News	rss://slashdot.org/index.rss
	D-Lib Magazine	http://dois.mimas.ac.uk/rss/dlib.xml
	Nature Journal	http://www.nature.com/nature/current_issue/rss
American Journal of Human Genetics	http://www.journals.uchicago.edu/cgi-bin/rss?jml-AJHG&type-pwr	

: contentType) 등 MS란의 독자적인 아이템을 정의하여 익스플로러 화면에서 RSS 구독기 없이도 피드별로 정의해 놓은 카테고리 정보나 콘텐츠 타입 등을 볼 수 있도록 하고 있다.

해외의 과학기술 분야 전문정보 역시 RSS를 이용하여 보다 다양하게 응용하여 서비스하고 있다. 중국 최대 과학기술 전문정보 사이트인 '중국 과학기술 정보 네트워크'와 중국 내 최대 RSS 콘텐츠 및 서비스 제공 사이트인 '간천하(看天下)'가 장기적인 협력관계를 구축하여 본격적인 RSS 서비스를 준비하고 있다(中國科技信息 2005).

과학기술 분야 대표적 출판사인 NPG(Nature Publishing Group)는 새로운 뉴스, 저널 기사 정보를 RSS로 제공하는 방식은 기

본이고, 최신 저널의 목차(TOC) 정보를 관련 기관 및 이용자에게 RSS로 웹 및 모바일 통신으로 서비스하고 있다. 또한, 기존의 표준 RSS 요소 외에 더블린 코어(Dublin Core)와 프리즘(PRISM) 메타데이터 요소를 RSS 포맷에 추가하여 서비스하고 있다.

NPG 외에도 저널의 TOC 정보를 RSS로 서비스하고 있는 과학기술 분야의 대표적인 출판사로 BioMed, Extenza, OUP(Oxford University Press), American Journal of Neuroradiology, AIP(American Institute of Physics), BMJ(British Medical Journal), Medscape, IOPP(Institute of Physics Publishing) 등이 있다(Tony, Timo, and Ben 2004). 이처럼, RSS 이용이 확산되면서

국제적인 표준 제정의 필요성이 증대되어 'Atom'이라는 프로젝트로 구체화되었고, 현재 IETF의 AtomPub WG에서 출판 프로토콜과 신디케이션 포맷의 확장에 대한 표준화 작업이 진행 중이며, 'Atom 1.0'은 RFC 인터넷 표준이 되기 전 마지막 단계인 "제안 표준(proposed standard)"으로 인정받은 상태이다(IETF 2005).

3. RSS 기반 과학기술정보 배급표준 시스템(STI-RSS) 개발

3.1 STI-RSS 개발 목표

STI RSS는 정보 서비스 기관이 풀 방식의 RSS 서비스 제공을 위해서 표준적인 관리 시스템 형태로 개발하는 것을 목표로 하였다. 과학기술정보의 다양한 형태와 서비스 환경 변화를 유연하게 받아들여 이 시스템을 통해 쉽게 RSS로 배급 및 통합 관리가 가능하도록 하며, 시스템 관리자가 RSS 피드의 배급시간이나 배급량, RSS 버전, 해당 채널, 해당 소스, 채널 관리자 등을 구성하고 운영할 수 있도록 유연하고 표준화된 시스템을 지향하였다.

3.2 STI-RSS 설계

STI RSS의 설계 방향은 다음과 같다.

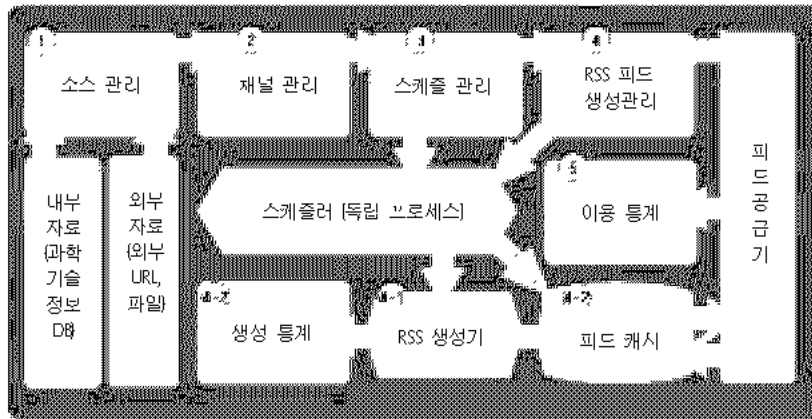
첫째, 과학기술정보의 다양한 정보 소스를 수용할 수 있어야 하고, 둘째, 무수히 많은 과

학기술 분류를 수용하여 처리할 수 있어야 하며, 셋째, 세계적인 표준화의 흐름에 맞춰 채널별로 RSS의 다양한 버전으로 생성할 수 있어야 한다. 넷째, 피드의 배급시간과 배급량을 설정 관리할 수 있어야 하며, 다섯째, 피드의 생성과 이용에 관한 통합적인 통계정보를 제공하고 관리할 수 있어야 한다. 여섯째, 채널별 관리뿐만 아니라 여러 채널의 통합적인 관리가 가능해야 한다.

따라서, 시스템 구성은 기본적으로 각 기능별 모듈화된 방식으로 설계하고, 시스템 운영시 발생하는 모든 중간 생성물을 XML형태로 정의하여 변환과 활용을 극대화하였다. 이러한 주요 설계 방향을 근거로 시스템을 설계하였으며, <그림 4>는 이러한 요구조건을 반영한 STI RSS의 구성도이다.

STI RSS의 활용절차는 다음과 같은 순서로 진행된다.

- ① 소스 관리: RSS 피드를 생성하고자하는 정보 소스를 먼저 선택한다.
- ② 채널 관리: 선택된 소스에서 과학기술 분류별 그룹 채널 또는 단위 채널을 만든다.
- ③ 스케줄 관리: 피드 생성을 위한 스케줄을 생성한다.
- ④ RSS 피드 생성관리: 피드 생성 스케줄에 따라 스케줄러가 피드 생성신호를 보내면 RSS 생성기를 통해 피드가 생성되며, 생성된 피드는 캐시에 저장된다.
RSS 생성기를 통해 생성된 피드는 생성통계를 통해 확인할 수 있으며, 생성관리를 통



〈그림 4〉 STI-RSS 시스템 구성도

해 공급될 피드의 캐시 량을 관리한다. 이용자가 접속하여 요청 시 피드 공급기를 통해 RSS 피드를 제공한다.

- ⑤ 이용통계: 이용자의 이용 내역을 통계관리에서 확인한다.

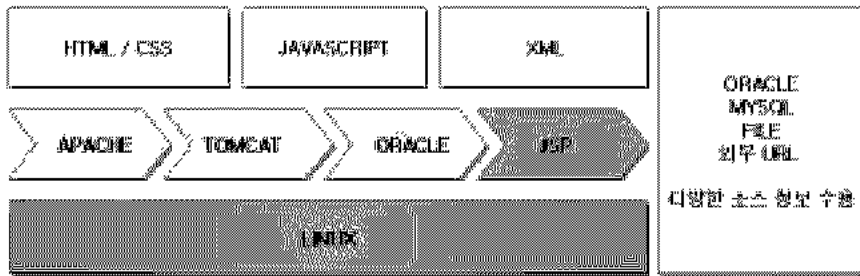
3.3 STI-RSS의 구현

3.3.1 구현 개요

앞에서 제시한 설계 방향에 따라 구현은 내·외부 다양한 정보 소스를 고려하고, 계층



〈그림 5〉 STI-RSS 관리자 초기화면



〈그림 6〉 STI-RSS 시스템 운영환경

적이고 세분화된 기술분류별로 효율적인 서비스를 하기 위해 그룹핑이 가능한 그룹 채널, 피드 생성기간과 배급량의 조절을 위한 피드 생성 스케줄, 피드의 생성과 배급 및 피드캐시의 모니터링, 채널별 또는 그룹 채널별, 전체 통합 관리가 가능한 통계관리, 다양한 관리자 권한 관리 등을 고려하여 구현하였다.

STI RSS의 운영환경은 리눅스 환경에서 JSP와 Java 언어를 사용하여 구현하였으며, DB는 오라클과 My SQL로 구현하였다. 세계적으로 가장 많이 사용되고 있는 Java 기반의 WAS인 TOMCAT과 JSP의 사용을 통해 범용성과 안정성을 확보하였으며, Oracle 9i 이상의 사용으로 콘텐츠의 효율적 관리와 운영을 보장한다. 오라클의 InterMedia를 이용하는 검색을 통해, 별도의 검색 엔진이 없이도 빠르고 안정적인 검색이 가능하다. 가능한 모든 중간 생성물을 XML형태로 정의하여 변환과 활용 가치를 극대화하고, 시스템에서의 호환성을 보장하도록 구현하였다.

3.3.2 구현 전략

RSS 기반 과학기술정보 배급표준 시스템의 구현 시 고려한 주요 구현 전략을 요약하면 다음과 같다.

1) 그룹 채널 및 단위 채널의 개념

다양한 과학기술정보를 수집, 관리하며 공급하여야 하는 기관에서는 이에 대한 채널의 구성이 다소 복잡해 질 수 있고, 보다 세분화된 콘텐츠(피드)를 제공하여야 한다. 이렇게 될 경우 관리상의 효율성은 떨어질 수밖에 없다. 따라서 채널 구성을 통한 서비스의 세분화와 다양화는 유지하고, 관리상의 효율성을 높이기 위한 방안이 필요한데, 이에 따라 구현된 것이 ‘그룹’의 개념이다.

그룹 채널은 원칙적으로 채널의 묶음 개념이지만, 관리적인 측면에서 고려된 것으로 관리자 및 관리 권한의 계층화와 통합이다. 즉, 그룹 관리자는 해당 그룹에 속한 모든 채널에 대한 관리 권한을 자동으로 갖게 되어 관리의 효율성을 높일 수 있다. 또한, 그룹에 대한 각종설정 정보는 그룹에 속한 채널을 신규 생성할 때 기본 값으로 사용되어 채널 생성 시 개별적으로 입력하는 번거로움을 피하고, 그룹

에 속한 각 채널에서의 수정이 불가능하도록 설정하여 채널 관리자의 실수를 미연에 방지할 수 있으며, 각기 다른 채널들에서 제공되는 RSS 기반 정보의 통일성을 유지할 수 있게 된다.

2) 스케줄링

스케줄링은 데이터베이스나 외부 URL과 같은 원천 소스로부터 데이터를 읽어, 채널 정보에 의해 설정된 각종 특성에 맞춰 자료를 재구성하고, RSS 피드의 구독자들에게 공급될 수 있는 형태로 변환 및 생성하는 일련의 작업에 관해 시간과 조건을 배정하는 것을 의미한다. 이와 같은 스케줄링에도 그룹 개념이 적용되어, 그룹 단위 스케줄이 가능하도록 하였다. 즉, 그룹 스케줄을 작성하면, 해당 그룹에 속한 모든 채널의 스케줄로 변환 생성된다. 이렇게 그룹 스케줄을 개별 채널 스케줄로 재생성하는 시스템 구성을 통해 단일사용자가 여러 채널의 피드를 생성관리하고 그 결과 통계를 확인하는 것으로 효율성을 높일 수 있다.

RSS 서비스는 특정 웹 사이트의 신규 내용을 채널 형태로 홍보하여 사이트에 방문을 유도하는 것이 주요 목적이므로 채널 중심의 콘텐츠 구성이 일반적이다. 이렇게 채널 중심의 콘텐츠를 스케줄링하여 제공할 경우, 해당 채널을 통해 제공될 개별 피드 콘텐츠의 내용을 사전에 확인하여 내용을 수정하거나 제거하는 등의 작업이 불가능하다. 따라서 채널에 관한 스케줄에 의해 생성될 피드를 사전에 확인하고 이에 대한 개별 피드를 수정, 삭제할

수 있고, 원천 소스에 포함되어 있지 않은 신규 피드를 즉각적으로 생성하기 위한 '피드 단위 스케줄'의 개념이 필요하다. 따라서 STI RSS 스케줄링 시스템은 채널 스케줄을 중심으로 통합 관리를 위한 그룹 스케줄과 개별 피드의 관리를 위한 피드 스케줄로 구성하였다.

3) 피드 생성 시 중복 생성 방지

해당 콘텐츠 또는 서비스를 즉각적으로 제공하는 일반 RSS 서비스와 달리, STI RSS는 생성된 피드 콘텐츠를 데이터베이스 기반의 캐시(Cache)에 보관하여 제공한다. 이는 피드의 생성 스케줄을 효과적으로 조절하여, 원천 웹 서비스에 영향을 주지 않고 RSS 서비스를 제공할 수 있다. 이와 같이 캐시를 생성하여 제공하는 경우, 스케줄이 중복되거나 단일 스케줄에 의해 생성되도록 설정된 피드의 수가 업데이트된 원천 소스의 레코드 수보다 많을 경우, 같은 피드 콘텐츠를 중복 생성하는 일이 생길 수 있다. STI RSS에서는 별도의 고유 식별값(Unique Identifier)을 원천 소스에 배정하도록 하여 동일 채널에 같은 RSS 피드가 중복 생성되는 것을 방지하였다.

4) 각종 설정값의 XML화 지원

신규 채널이나 그룹의 등록은 다소 많은 설정 정보를 입력해야 하는 일이고, 대부분의 웹 서비스가 입력 폼을 사용하여 일일이 입력하는 것은 그리 쉬운 일이 아니다. 또한 RSS 서비스를 제공하고, 관리해야 하는 관리담당자 대부분이 RSS의 기술적 특성이나 해당 채널의 원

천 소스에 대한 정보를 정확하게 인지하고 있지 않은 것이 현실이다.

따라서, 해당 채널에 대한 각종 정보, 기술적인 설정값들을 XML형태로 제공하여 업로드 또는 다운로드하도록 설계·구현함으로써, 사용자의 편의성을 극대화하였으며, STI RSS가 배포될 때, 함께 배포하여 템플릿 역할을 할 수 있도록 하였다.

5) 표준화의 범위 및 커스터마이제이션

STI RSS는 '표준 시스템' 이므로 이에 따라 외부 원천 소스에 대한 정의 부분 및 통계 정보 등의 제공에 있어서 적절한 표준화 범위를 설정하기 위해 노력하였다.

먼저, 외부 원천 소스가 데이터베이스인 경우, 우선적으로 Oracle™ 및 MySQL™ 에 대해 지원하고, 향후 JDBC기반의 표준 인터페이스를 제공하는 모든 데이터베이스로 확장이 가능하도록 하였다. 또한 표준 시스템에서는 단일 데이터베이스 테이블만을 채널에서 활용할 수 있도록 제한하였으며, 해당 원천 소스에 포함된 여러 개의 테이블을 조인해서 피드 콘텐츠를 생성해야 하는 경우는 개별적인 커스터마이징 과정으로 규정하여, 추가적인 뷰(View)를 작성하고, 그 정보를 참조하는 것을 원칙으로 하였다.

통계정보의 제공 또한 표준 시스템에서는 생성과 공급에 관한 채널별, 기간별 통계치를 누적 제공하여 기존 웹 서비스의 통합 관리환경에서 활용할 수 있도록 하였으며, 이에 따라 함께 제공되는 그래프의 표현방법 또한

Macro Media사의 Flash™ 기술을 활용하게 제공하고 있으며, 커스터마이징의 유연성을 확보하기 위해 XML 방식과 자바 스크립트 방식의 통계정보를 그래프가 표현할 수 있도록 하였다.

6) 스케줄러의 수행을 위한 타이머 구현방안
스케줄 관리에 의해 작성된 스케줄은 우선적으로는 데이터베이스에 보관되어 해당 시간 이과 조건이 되었을 때 수행된다. 하지만, 일반 웹 인터페이스에서 이와 같은 반복적인 작업을 자동으로 할 수는 없기 때문에, 별도의 타이머 서비스 데몬(Timer Demon)을 제작하여 반복적인 작업이 안정적으로 수행되도록 하였다.

타이머의 구현은 타이머 서비스 데몬이 미리 설정된 일정 시간마다 스케줄러로 신호만 전달하는 역할을 수행하고, 피드 생성 프로세스와는 독립적으로 수행되도록 구현하였다. 이것은 커스터마이징 측면에서 어떤 기능의 조건이 변경될 경우 서비스 데몬을 수정하지 않고도 변경된 기능의 수정만하면 된다는 장점이 있다.

3.3.3 주요 기능 및 특징

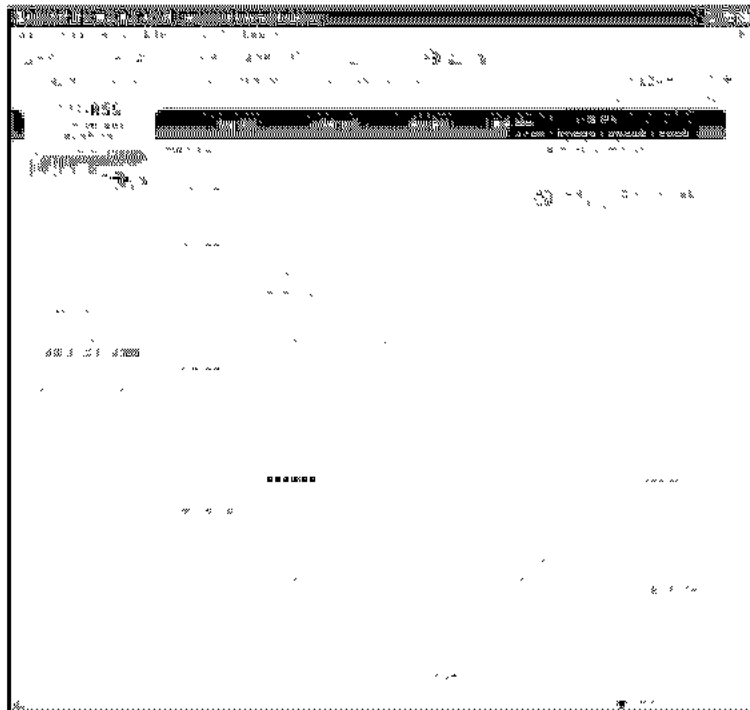
〈표 2〉는 구현전략을 토대로 구성한 STI RSS의 기능구조를 나타낸 표이다.

1) 소스 관리

소스 관리는 RSS 피드 생성을 위한 첫 단계로서 과학기술정보 DB(오라클, My SQL), 파일, 외부 URL의 신규 소스 등록과 함께 이미 등록된 소스 목록을 관리한다. 해당 소스에 대

〈표 2〉 STI-RSS 기능구조

주기능	서브기능	설명
소스 관리	소스 목록	등록된 소스 목록 관리
	신규 소스 등록	소스(DB, 파일, RSS주소) 등록
채널 관리	채널 목록	채널 생성 및 관리
	그룹 목록	그룹 채널 생성 및 관리
스케줄 관리	그룹 스케줄	그룹 채널 스케줄 생성 및 관리
	채널 스케줄	단위 채널 스케줄 생성 및 관리
	피드 스케줄	피드 단위 스케줄 생성 및 관리
	시스템 관리	시스템 타이머 운영 관리
생성 관리	피드 생성 현황	피드 생성 현황 모니터링
	피드 공급 현황	피드 공급 현황 모니터링
	캐시 운영 관리	채널 캐시 현황 확인 및 관리
통계 관리	통계 보기	그룹별, 채널별, 스케줄별 피드 생성 및 공급 통계 확인
권한 관리	관리자 목록	그룹 및 채널 관리자 등록 및 관리
	권한 목록	등록된 관리자의 메뉴별 권한 관리



〈그림 7〉 STI-RSS 소스 관리 화면

한 접속정보의 유효성검사를 통해 오류를 최소화하고, 소스설정의 편의성을 높이기 위해 기술적인 설정값을 미리 XML 파일로 만들어놓고, 지정하면 소스등록을 위한 전체 설정에 반영되도록 하였다. 피드의 중복 생성을 방지하기 위해 별도의 고유 식별값(Unique ID)을 소스에 배정하여 동일 채널에 같은 피드가 생성되는 것을 방지한다.

2) 채널 관리

채널 관리는 단위 채널의 그룹핑이 가능한 그룹 채널과 개별 채널, 피드 단위 채널의 목록을 생성하고 관리한다. 그룹 채널과 단위 채널 간의 종속관계를 부여함으로써 그룹에 대한 각종 설정정보가 해당 그룹에 속한 채널을 신규

생성할 때 기본값으로 사용되도록 하여 채널별로 재설정해야 하는 번거로움 없이 일관성을 유지하도록 한다.

세계적인 표준화의 흐름에 맞춰 다양한 RSS 버전으로 생성할 수 있도록, 채널 생성 시 채널별로 채널 관리자, RSS 버전(0.91, 0.92, 1.0, 2.0, ATOM 0.2, 1.0) 선택이 가능하다. 특정 조건만을 만족하는 피드를 생성할 수 있도록 피드 생성조건 설정을 통해 동일한 소스더라도 해당 조건만을 만족하는 정보를 선별하여 피드로 생성할 수 있고, 유효성 검사를 위해 '미리보기' 기능을 제공한다. 과학기술정보 RSS 피드가 한꺼번에 많은 양이 발생할 수 있으므로 피드의 생성과 배급량 조절을 위한, 피



〈그림 8〉 STI-RSS 채널 설정 화면

드 생성 시간, 피드 생성량 및 캐시 운영 방법 (캐시 기간, 캐시 건수)을 설정하고 관리한다.

3) 스케줄 관리

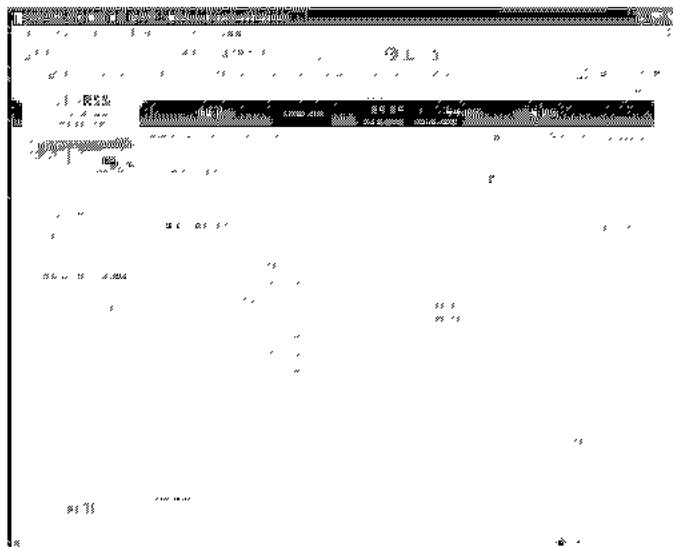
스케줄 관리는 그룹 채널 스케줄, 개별 채널 스케줄, 피드 단위 스케줄의 신규 생성 및 관리와 함께 스케줄러의 동작시간과 간격을 관리한다. 그룹 스케줄을 작성하면, 해당 그룹에 속한 모든 채널의 스케줄도 동시에 생성된다. 피드 단위 스케줄은 특정 조건에 해당하는 개별 피드를 추가하거나 제외하는 기능을 제공한다. 스케줄러의 수행을 위한 타이머를 서비스 데몬으로 독립 구성하여, 타이머의 주기적인 신호에만 피드 생성 프로세스를 수행하도록 구현함으로써 지속적인 스케줄 체크의 부하를 줄일 수 있다.

4) 생성관리

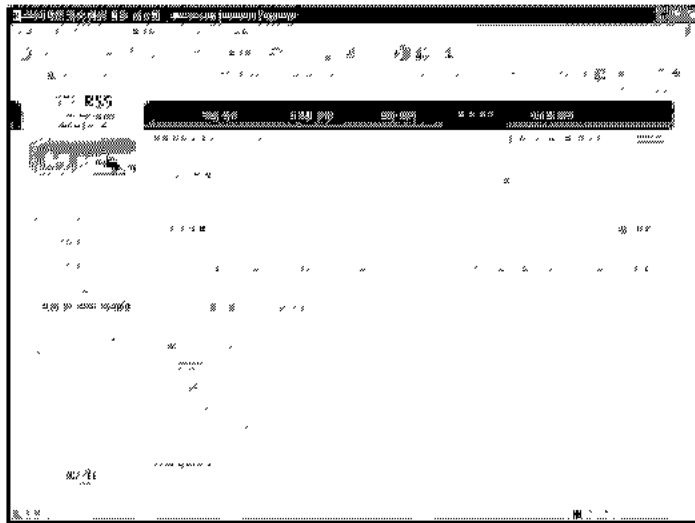
생성관리는 생성된 피드의 배급을 위한 캐시 양을 조정하도록 피드의 생성현황 및 공급현황, 캐시 사용현황을 확인하고 운영관리하며, 지정한 기간별로 그래프를 제공한다. 매일 발생하는 정보도 있지만 비정기적으로 발생하는 과학기술정보의 양과 주기를 고려하여 피드 캐시 기간과 양을 조정하도록 채널별 피드의 모니터링 정보를 제공한다.

5) 통계관리

통계관리는 그룹 채널별, 개별 채널별, 스케줄별 피드 생성과 공급 통계뿐만 아니라, 전체 통합 피드의 생성과 공급, 여러 통계를 확인한다. 통계는 각 기간별로 구분하여 그래프로 표현하고 엑셀 파일로 다운로드 할 수 있다. 통계는 그룹별 채널별 스케줄별 주간, 월간, 연간 피드 통계를 제공한다.



〈그림 9〉 STI-RSS 채널 스케줄 관리 화면



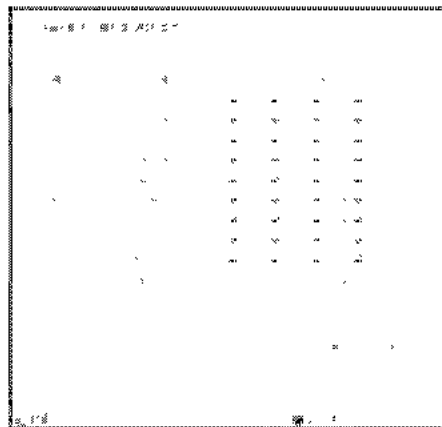
〈그림 10〉 STI-RSS 통계관리 화면

통계 그래프 생성기능은 플래시 기술을 활용하였으며, 시스템 연동시 XML 파일로 입력하는 방식과 자바 스크립트로 직접 콜(call)하는 방식을 제공한다.

6) 권한관리

권한관리는 관리자목록에서 신규관리자를

등록하고, 권한목록에서 등록된 관리자가 그룹 채널 및 단위 채널별 각 서브 메뉴별로 읽기, 쓰기, 삭제, 수정 등의 다양한 권한을 부여하고 관리한다. 관리자 그룹별(그룹 관리자, 채널 관리자, 시스템 관리자)로 묶어서 한 번에 권한을 부여하고, 개별 관리자별로도 권한을 부여할



〈그림 11〉 STI-RSS 채널 관리자 권한관리 화면

수 있도록 하기 위해서 관리자별 메뉴별 권한 테이블을 <그림 11>과 같이 매트릭스 형태로 정의하여 구현하였다.

4. STI-RSS 활용방안

4.1 정보 서비스의 STI-RSS 활용

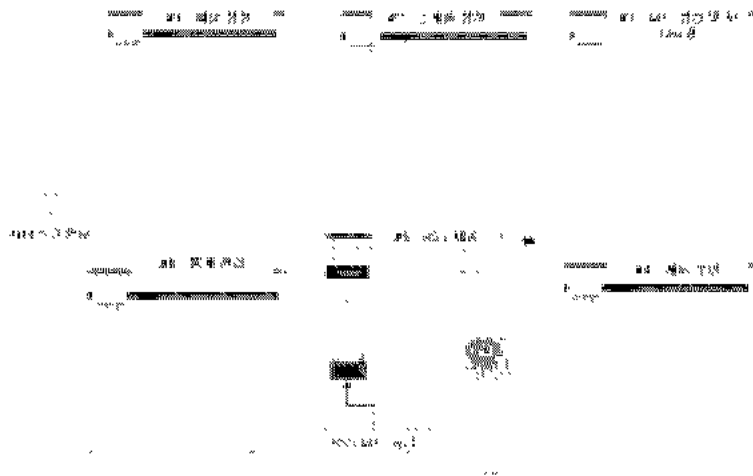
STI RSS는 기존 정보이용자들의 정보 서비스에 영향을 주지 않으면서, 서비스 운영자가 간단한 설정만으로 추가적인 RSS 정보 서비스를 제공할 수 있는 시스템이다.

<그림 12>는 일반 정보 서비스에서 STI RSS를 단계적으로 적용하는 절차를 순차적으로 나타낸 그림이다.

즉, 정보 서비스 운영자가 STI RSS를 적용하여 RSS 서비스를 제공하기 위한 일반적인 절차의 화면이며, 시스템 관리자가 아닌 개별

정보 서비스 운영자인 경우 채널 관리자의 권한을 부여받아 이용할 수 있는 활용 사례 화면이다.

- ① 채널 생성: 서비스 운영자는 기존에 서비스 하고 있는 정보의 소스를 선택하고 RSS 피드로 서비스할 항목을 선택한 다음 RSS 채널 정보(RSS 버전, 채널 명, 채널 주소, 채널 그룹 등)를 설정하고, 신규 채널을 생성한다.
- ② 스케줄 생성: 신규 채널이 생성되면 채널의 피드 생성 스케줄(생성일정, 생성개수, 반복수행 등)을 설정하고 생성한다.
- ③ 피드 생성 및 모니터링: 서비스 운영자가 피드 생성을 수행하거나 스케줄에 따라 피드가 생성되면 캐시의 피드 양을 모니터링(신규 생성건수, 전체 캐시 건수 등) 한다.
- ④ 캐시 운영: 캐시의 RSS 피드를 추가하거나 수정하여 캐시 양을 조정하고 서비스 준비



<그림 12> STI-RSS 적용 일반 절차 화면예시

를 완료한다.

- ⑤ RSS 제공: 기존 정보 서비스 웹 사이트에서 RSS 채널 주소를 공개하면 RSS 서비스가 제공된다.
- ⑥ 통계관리: 이용자가 해당 채널의 RSS 피드를 이용한 통계도 STI RSS의 통계관리를 통해서 확인할 수 있다.

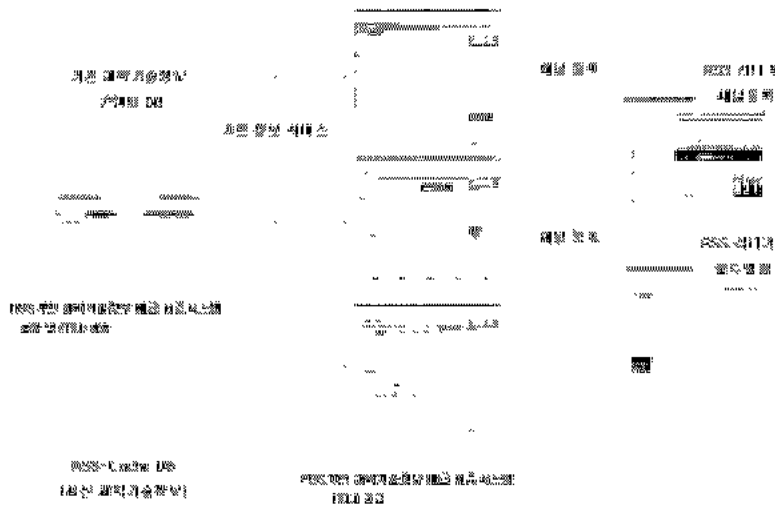
4.2 과학기술정보 서비스의 STI-RSS 활용

과학기술정보 서비스에서 STI RSS 활용은 기본적인 순서는 동일하지만 서비스의 특성과 정보속성에 맞춰 보다 더 세분화된 기능을 활용하면 효과적인 운영이 가능하다.

〈그림 13〉은 기존에 서비스 중인 다양한 과학기술정보 서비스별로 RSS 채널을 생성하고 정보이용자가 RSS 리더기 내에 서비스 채널

주소를 등록하여 활용하는 것을 전제로 한 Client Server 방식의 그림이다. 즉, STI RSS을 활용하여 정보제공기관이 다양한 서비스별로 채널과 피드를 생성하고 RSS 서비스를 제공할 때 정보이용자의 관점에서 채널 등록을 한 후에 RSS 피드를 열람하는 흐름도이다.

과학기술정보의 특성상 수많은 과학기술 분야의 세부 분야별로 RSS 채널을 제공하고자 할 때는 서비스 운영자의 관리적인 측면뿐만 아니라 정보를 이용하는 정보이용자에게도 관심 채널을 모두 등록해야만 하는 불편을 가중시킬 수 있다. STI RSS는 과학기술정보 서비스 운영자를 위하여 개별 채널을 그룹화하여 생성하고 관리할 수 있는 그룹 채널을 제공하고 있으며, 정보이용자를 위하여 채널 목록을 XML 형식으로 정리해 놓은 OPML(Object Playlist Markup Language)을 제공하여 단



〈그림 13〉 과학기술정보 서비스의 STI-RSS 활용

위 채널을 과학기술 분류로 그룹화한 목록을 제공할 수 있다. 이것은 정보이용자가 세분화된 개별 채널을 너무 많이 제공받게 됨으로써 일일이 채널 등록을 해야 하는 번거로움을 줄이고 이용자가 원하는 해당 분류의 채널만을 받아 등록할 수 있도록 지원한다.

방대한 양이 비정기적으로 생산되는 과학기술정보의 특성을 고려하여 서비스 운영자는 피드 생산량과 생산 스케줄링을 보다 정교하게 설정하고, 이용자가 많지 않은 시간대로 조정하여 피드를 생성함으로써 시스템 운영을 효율적으로 관리한다.

다양한 과학기술정보의 서비스별 정보 생산 주기를 고려한 스케줄링을 하려면 매번 다양하고 복잡한 설정값을 지정해야하는 번거로움이 발생할 수 있으므로, 비슷한 성격의 서비스나 정보 생산주기를 가진 채널의 설정값은 XML로 다운로드 받아놓고 필요할 때 XML 설정값을 업로드하여 활용하도록 한다.

각 채널별 RSS 피드의 생산과 이용에 대한 통계는 개별 채널 관리자의 권한으로 지정되어 있지만 서비스 전체의 피드 생산과 이용에 대한 통계는 시스템 관리자의 권한이므로 필요시에는 채널 관리자나 그룹 관리자가 시스템 관리자에게 특정 권한을 부여받아 활용하도록 한다.

4.3 자동분류 에이전트를 적용한 STI-RSS의 활용방안

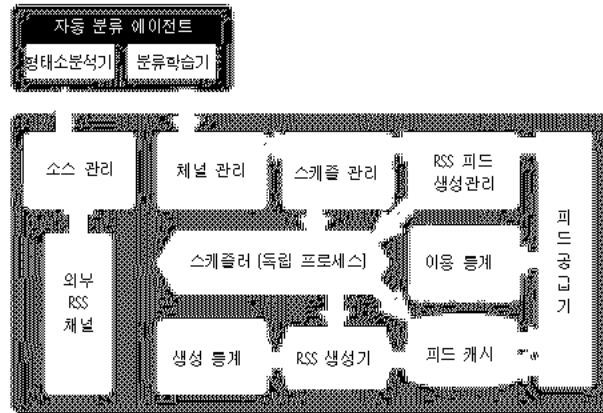
STI RSS의 정보 소스는 과학기술 분야의

DB나 파일, 외부 URL 뿐만 아니라 외부의 RSS 채널에서 발생하는 RSS 피드 정보에 대해서도 그룹핑을 하면 새로운 채널의 RSS 피드로 생성이 가능하다. 하지만, 이러한 외부 RSS 피드의 보다 정확하고 자동화된 그룹핑을 하기 위해서는 채널을 그룹핑하는 과정에 자동분류기능을 적용하는 것이다. 즉, 외부 과학기술 분야의 RSS 채널에서 수집한 RSS 피드에 대해서 과학기술 분류로 재분류한 채널별 자동분류가 가능해진다.

2002년부터 KISTI가 자체 개발하여 문서 자동분류기능을 가진 SWING(Science Web Information Nimble Guide)은 인터넷의 과학기술정보를 통합 검색하여 주제별로 자동분류하여 제공해주는 시스템으로, 외부 RSS 피드의 자동분류를 위해 STI RSS와 연계가 가능할 것이다.

SWING의 '자동분류 에이전트'는 수집된 정보를 형태소 분석기와 분류 학습기를 이용하여 해당 분류에 할당해 주며, 형태소분석기에 의해 각각의 형태소가 추출되고 빈도수가 계산되면, 분류 학습기에서는 추출된 형태소별로 각각의 기존 분류 디렉터리에서의 중요도에 따른 가중치와 정확도를 계산하여 해당 분류 디렉터리에 할당하고 분류 디렉터리별 가중치를 재조정한다(황성하 2005).

SWING의 자동분류 에이전트를 STI RSS의 일부 내부 모듈과 연동하면 재분류된 외부 RSS 피드는 미리 설정해놓은 분류별 채널 관리와 스케줄 관리를 통해서 자동으로 분류별 채널



〈그림 14〉 자동분류 에이전트를 적용한 STI-RSS 구성도

의 피드 캐시에 저장되고 서비스 할 수 있게 된다. 〈그림 14〉는 자동분류 시스템의 자동분류 에이전트를 적용한 STI RSS 구성도이다.

SWING의 자동분류 에이전트와 STI RSS 가 연동하려면 STI RSS의 일부 기능이 확장되어야 한다. 소스 관리에서는 외부 RSS 채널을 정보 소스로 추가하여야 하며, RSS 생성기는 외부에서 수집된 RSS 피드를 동일한 버전으로 변환하기 위한 변환기능이 추가되어야 한다. 또한 생성 통계도 외부 RSS 피드의 수집통계로 집계되어야 할 것이다.

5. 결 론

정보 서비스 기관의 정보를 제공하기 위한 이메일 제공방식의 푸시 서비스 대체수단으로 B2C 기반 정보배급 기술인 RSS가 새로운 패러다임으로 급속히 진행되고 있다. 공급자적인 측면의 정보제공 방식에서 이용자측면의 최신

정보 구독에 대한 선택과 권한을 이용자에게 부여하는 것은 바람직하다고 본다. 하지만, 모든 이용자가 이러한 정보제공 방식을 반기는 것은 아니라고 보며, 다양한 계층의 이용자 수요를 감안하면 RSS 서비스는 전통적인 이메일 제공방식과 함께 양립하여 제공하는 것이 바람직할 것이다. 또한, RSS 서비스가 이용자에게 정말 편리한 서비스가 되려면 단순히 RSS 서비스를 추가제공하는 것으로는 부족하며, 각 RSS 채널별로 제공받는 피드의 적절한 양과 정보의 품질 면에서 모두 만족되어야 한다고 본다.

본 연구에서는 RSS와 다른 정보배급 프로토콜과의 특징을 분석해보고, 국내외 RSS 동향을 살펴보았다. 국내 과학기술 분야의 RSS 서비스를 위해 과학기술 분야 지식정보의 다양한 정보 소스와 환경을 고려한 STI RSS의 개발내용과 함께 활용방안에 대해서도 제시하였다.

STI RSS의 개발은 과학기술 분야의 다양

한 서비스별 RSS 시스템의 개발중복 가능성을 미연에 방지할 수 있고, 무엇보다 표준화된 포맷을 사용함으로써 연계 되는 콘텐츠 재사용 및 시스템 간 상호운용성, 통합관리 측면의 효과에서 매우 의미가 있다고 본다. STI RSS는 현재 RSS 서비스 대상 정보 소스가 데이터베이스인 경우, 오라클 및 MySQL에 대해서만 지원하지만 모든 데이터베이스로 확장이 필요하고, 실제 다양한 환경에 적용할 때 이에 대한 시스템 성능평가도 필요하다.

현재는 기존 정보제공 서비스와 함께 RSS 서비스를 추가제공하는 단계에 있지만, 과학기술 분야의 전문정보에 대해 외부에서 발생하는 각종 분야별 RSS 피드를 수집한 후 피드의 분류별 재분배 서비스를 제공하는 RSS 포털 서비스가 더욱 향상된 서비스로 대두될 것으로 본다. 향후 이러한 국내외 과학기술정보 RSS 포털 서비스를 위해 STI RSS를 확장하여 적용하게 되면 매우 유용할 것으로 예상된다. NPG의 사례처럼 과학저널의 기사 정보를 RSS로 제공시 정보의 식별을 위해 DC와 PRISM의 메타데이터를 추가하여 발급하듯이, 과학기술정보 RSS 서비스에도 과학기술 분야별 정보의 특성을 고려한 표준화된 메타데이터 항목의 적절한 추가 및 응용도 필요하리라고 본다.

참고문헌

권이남, 김명일, 김재수, 신기정. 2004. 효율적

인 맞춤형정보 서비스를 위한 RSS 기반 개인화 지원 연구. 『KOSTI Workshop 2004 과학기술정보』, 69-82.

김중태. 2004. 『나는 블로그가 좋다』. [서울: 이비컴]

백윤주. 2001. 디지털 콘텐츠 신디케이션. 『정보처리학회지』, 8(5): 103-110.

성균관대학교 과학기술정보관. 2004. 성균관대학교 과학기술정보관 RSS 서비스. [인용 2005. 11. 22].

〈<http://skksl.skku.ac.kr/DLiWeb20/components/sites/skku/suwon/rss/front.aspx>〉.

전종홍. 2004. 콘텐츠 신디케이션 표준화 동향. 『IT Standard Weekly』, 40권.

中國科技信息. 2005. 中國科技信息 Homepage. [cited 2005. 5. 10].

〈http://www.chinainfo.gov.cn/data/200505/1_20050510_109960.html〉.

한국기초과학지원연구원. 2002. KEOL(연구장비정보망) 홈페이지. [인용 2005. 11. 10. 16].

〈<http://www.keol.net>〉.

황성하, 최광남, 이대규, 이상호. 2005. 인터넷 문서의 자동분류 서비스 시스템에 관한 구현. 『한국콘텐츠학회 추계학술대회』, 66-71.

Ben Hammersley. 2003. *Content Syndication with RSS*. O'Reilly & Associates, Inc.

- Ben Hammersley, 2005, *Development Feeds with RSS and Atom*, O'Reilly & Associates, Inc.
- Danny Ayers, and Andrew Watt, 2005, *Beginning RSS and Atom Programming*, Wiley Publishing, Inc.
- IETF, 2005, Atom Publishing Format and Protocol(atompub) Homepage, [cited 2005, 8, 15].
<<http://www.ietf.org/html.charters/atompub-charter.html>>.
- James Lewin, 2000, *An introduction to RSS news feeds Using open formats for content syndication*, [cited 2005, 5, 22].
<<http://www-128.ibm.com/developerworks/library/wrss.html?dwzone=web>>.
- Peter Cliff, 2002, *RSS Sharing Online Content Metadata*, [cited 2005, 6, 12].
<<http://www.cultivate-int.org/issue7/rss/>>.
- Rael Dornfest, 2000, *RSS: Lightweight Web Syndication*, [cited 2005, 5, 14].
<<http://www.xml.com/pub/a/2000/07/17/syndication/rss.html>>.
- Tony Hammond, Timo Hannay, and Ben Lund, 2004, "The Role of RSS in Science Publishing," *D Lib Magazine*, 10(12)
- Tony Hammond, Timo Hannay, and Ben Lund, 2003, "TOC Syndication using RSS," *Crossref Technical meeting September 17*.