

KRISTAL-2002 기반의 학술정보관리시스템의 설계 및 구현

김병규⁰, 강무영, 박재원
한국과학기술정보연구원
(yourovin⁰,kmy,ilontos)@kisti.re.kr

Design and Implementation of KRISTAL-2002 based Academic Content Management System

Byung-Kyu Kim⁰, Moo-Young Kang, Jae-Won Park
Dept. of Domestic information, Korean Institute of Science and Technology Information

요 약

과학기술 분야 학회에서 발생하는 고급 학술정보는 산업,경제적 측면에서 그 효용 가치가 매우 크다. 따라서 국가 경쟁력 강화를 위해 전략적 차원으로 첨단과학기술정보를 디지털화하여 부존자원화 하고 효율적인 서비스체제를 구축하는 것이 매우 중요하다. 이러한 일련의 과정을 수행하는 것이 학술정보관리시스템이다. 한국과학기술정보연구원에서는 학술정보관리를 위하여 관계형 DB와 KRISTAL-II 기반의 관리시스템을 개발하여 운영하고 있다. 하지만 시스템에 종속적인 메타데이터의 형식과 복잡한 관리 프로세스 그리고 다양하지 못한 검색 서비스 제공과 같은 단점을 가지고 있다. 본 논문에서는 데이터의 저장관리와 빠른 검색을 제공하는 KISTI의 최신 IRMS인 KRISTAL-2002만을 사용하여 새로운 학술정보관리시스템(ACMS)을 설계하고 구현하였다. 또한 ACMS는 학술정보의 처리를 메타데이터의 표준인 XML을 사용함으로써 데이터의 유효성 검증과 타 기관의 상호대차 및 원문 교환을 손쉽게 하였다.

1. 서론

과학기술 분야 학회에서 국내,외 학술활동 및 교류를 통해서 생산되는 고급 학술정보는 그 효용 가치가 매우 높으며 이를 디지털화하여 부존자원화 하고 연구자가 쉽고 빠르게 접근 할 수 있는 정보 서비스체제를 구축 및 제공하는 것이 세계적 추세이다.

한국과학기술정보연구원(Korea Institute of Science & Technology Information, KISTI)은 우리나라의 과학기술정보 인프라 구축과 선진 서비스체제의 개발을 위한 대표적 연구기관으로서 학술정보의 체계적인 수집, 가공, 처리, 서비스를 유기적으로 시스템화하고 끈김없는 윈스톱 서비스를 구현하기 위하여 노력하고 있다. 이런 노력의 일환으로 과학기술계 학회들을 위한 온라인 커뮤니티인 학회마을을 구성하고 학회에 대한 정보화 지원과 학회가 생산하는 고급 학술정보를 수집하여 DB 로 구축 인터넷으로 서비스하기 위한 학회정보화사업을 추진하고 있다[1]. 현재 학회정보화사업에는 총 241 개 학회가 참여하고 있으며 학회마을 검색서비스를 통해서 서비스되고 있는 정보자원의 양은 640 종에 28 만건에 이르고 있다.

이렇게 방대한 양의 학술정보를 종합적으로 관리하고 이용자들에게 빠르고 편리한 검색서비스를 제공하기 위한 효율적인 시스템이 절실히 필요하다.

KISTI 에서는 이러한 목적으로 학술정보종합관리 및 검색서비스를 위한 관계형 DB 와 KRISTAL-II 기반의 관리시스템 [2,3,4]을 구축하여 운영하고 있다. 그러나 현재 세가지 측면에서 문제점을 가지고 있는데 첫째는 복잡한 관리 프로세스이고 둘째는 다양한 검색 기능의 부족이며 셋째는 학술정보 즉 논문에 대한 메타데이터의 정의가 고정적이어서 그 확장성이 원천적으로 결여되었으며 타 시스템과의 상호교환 및 운용성에도 상당히 미흡하다는 점이다.

본 논문에서는 다양한 검색기능과 데이터 관리기능을 제공하는 KISTI 의 최신 IRMS 인 KRISTAL-2002[5] 만을 사용하여 새로운 학술정보관리시스템(ACMS)을 설계하고 구현하였다. 또한 학술정보의 메타데이터는 문서 표준인 XML[6]를 사용함으로써 데이터의 유효성 검증과 타 기관의 상호대차 및 원문 교환을 손쉽게 할 수 있게 하였다.

본 논문의 구성은 다음과 같다. 2장에서는 관련연구를, 3장에서는 ACMS의 설계와 개발내용을, 마지막으로 4장에서는 결론 및 향후 연구 방향에 대해 살펴본다.

2. 관련연구

2.1 XML

XML(eXtensible Markup Language)은 1988년 W3C(World

Wid Web Consortium)에 의해서 표준으로 채택된 전자문서의 기술을 위한 메타 언어이다. XML은 SGML(Standard Generalized Markup Language)의 복잡한 면을 제거하고 구조, 검증, 확장성의 특징을 계승하면서 현재 웹 문서 기술 언어로서 널리 사용되고 있는 HTML의 단점을 보완하여 차세대 웹 문서의 표준으로 자리잡았다.

2.2 KRISTAL-2002

KRISTAL-2002는 KISTI에서 개발한 정보검색관리시스템 (Information Retrieval & Management System, IRMS)으로, 윈도우 및 유닉스 기반에서 데이터에 대한 저장,관리,검색에 관한 처리를 효율적으로 수행한다.

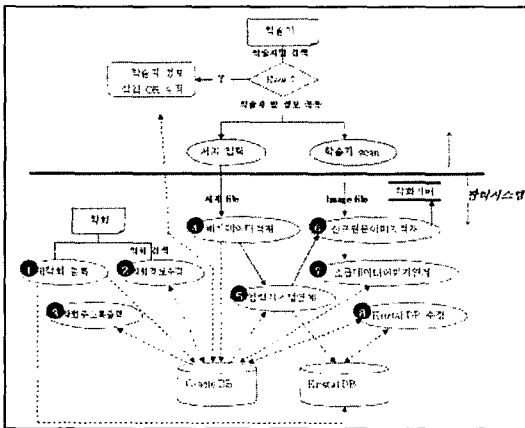
KRISTAL-2002의 특징을 살펴보면 첫째, 데이터의 생성부터 관리, 정보 서비스까지 단일 플랫폼으로 제공하면서 Flat data, XML 정보를 모두 수용할 수 있고 둘째, 소규모 정보서비스부터 대규모 정보 서비스까지 이용가능 하며 셋째로 대용량 정보의 분산 수용 및 검색이 가능 하다는 것이다.

KRISTAL-2000 데이터베이스는 검색엔진(IR)과 데이터베이스(DBMS)를 결합한 확장형 정보검색관리시스템(IRMS)으로서 상용 DBMS와 비교하였을때 대용량 데이터를 신속하게 적재할 수 있으며 더욱 빠르고 정확한 검색을 지원하며 멀티 프로세스를 이용한 다수사용자 지원과 멀티쓰레드를 이용한 동시 검색을 제공하며 트랜잭션 처리를 통한 온라인 데이터관리를 통해 데이터 안정성 및 복구 지원을 제공한다.

2.3 KRISTAL-II 기반의 학술정보관리시스템

KRISTAL-II 기반의 학술정보관리시스템은 학회정보종합관리를 위해 학회관리 모듈, 논문정보관리 모듈,로그 및 통계 모듈로 구성되며 이 모듈들은 필요에 따라 저장관리를 위한 Oracle DB와 검색서비스를 위한 Kristal DB을 이용한다. 또한 일반 사용자에게 학회마을에 소속된 학회들의 논문정보들을 디렉토리 서비스와 키워드검색 서비스를 통해 제공한다.

KRISTAL-II는 객체의 제한과 트랜잭션 처리를 지원하지 않는 단점을 가지고 있다. [그림 1]은 시스템의 관리 프로세스 흐름도이다.



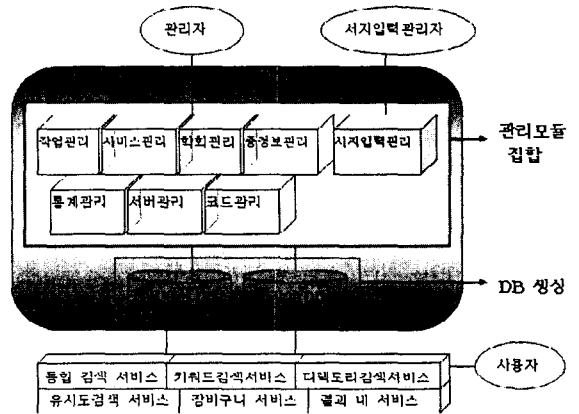
[그림 1] KRISTAL-II기반의 학술정보관리시스템 업무흐름

3. ACMS의 설계 및 구현

ACMS(학술정보종합관리시스템)는 국내 과학기술분야 학회에서 생산하고 있는 학술자료(학회지, 논문지 및 학술회의자료 등)를 망라적으로 수집, 가공하여 학술정보 DB를 구축하고, 정보이용자들이 원하는 학술정보를 언제 어디서든지 인터넷을 통해 서지사항은 물론 디지털 원문까지 획득할 수 있도록 하는 국내 학술정보 토털서비스 체제이다.

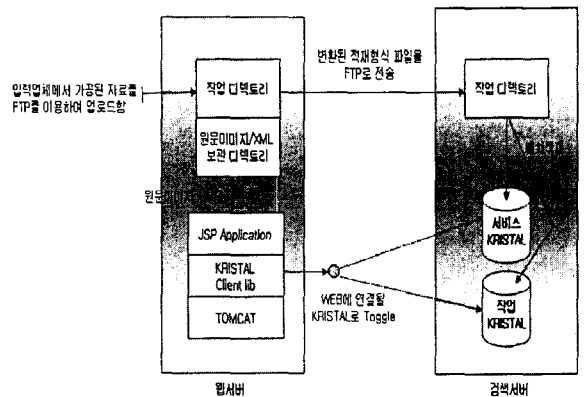
3.1 ACMS의 기능 및 시스템 구성

ACMS는 [그림 3]과 같이 이용자의 레벨이 3가지로 구성되어 있다. 관리자와 서지입력관리자는 학술정보의 입수,가공,DB 저장을 담당하며 일반 사용자는 ACMS에서 제공하는 검색서비스를 이용한다.



[그림 3] 이용자 별 관리 및 검색 기능

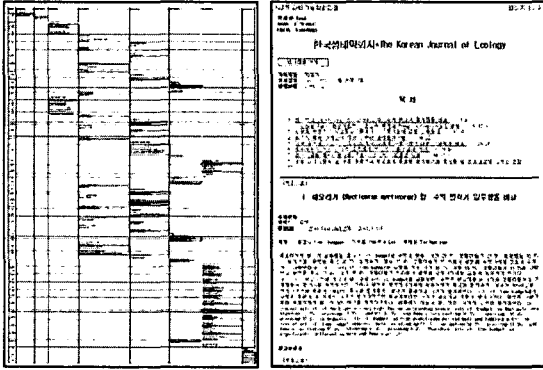
ACMS는 [그림 4]와 같이 웹서버와 검색서버로 분리구성되며 KRISTAL DB를 듀얼로 관리함으로써 시스템 장애 또는 관리실수에 적절히 대응할 수 있다. 또한 킨 단위로 제작되는 대량의 XML문서는 배치적재기능을 통해 KRISTAL DB에 로딩되며 곧바로 검색서비스를 통해 제공된다.



[그림 4] 시스템 구성

3.2 ACMS 동작원리

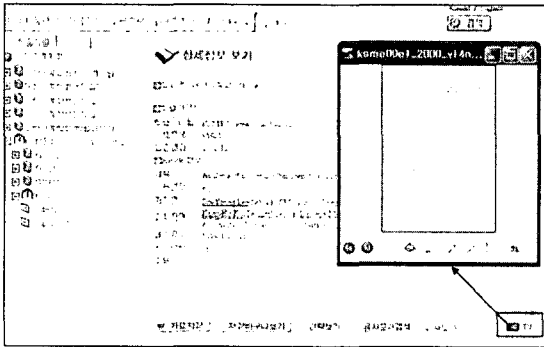
관리자는 입수된 학술정보를 ACMS에 등록하고 서지입력지시자에게 권단위 XML문서로 작성 및 원문이미지 스캔작업 명령을 내린다. 서지입력관리자는 XML데이터와 원문이미지를 제작하여 작업디렉토리에 저장한다. 관리자는 [그림 5]의 학술정보 XML DTD 구조에 맞게 XML문서를 검증하고 XSL를 통하여 논문의 내용확인 및 원문이미지의 가공상태를 확인한다.



[그림 5] 학술정보 XML 문서 구조와 XSL 적용 결과

데이터 가공상태를 확인한 관리자는 XML문서를 작업 크리스탈과 서비스 크리스탈에 배치적재하며 원문이미지를 학회->종->발간년도->권호와 같이 구조적으로 파일시스템에 저장한다.

위와 같은 과정이 끝나면 일반 이용자들은 ACMS를 통하여 다양한 검색 서비스를 제공 받을 수 있다. [그림 6]은 디렉토리 검색 서비스가 이루어지는 화면이다.



[그림 6] 디렉토리 검색 서비스 화면

ACMS는 통합검색 및 유사도 검색과 같은 다양한 검색 서비스와 저장바구니와 같은 편리한 기능들을 제공한다. 또한 검색 시에 한글과 영문으로 검색페이지를 선택할 수 있다.

작업 KRISTAL과 서비스 KRISTAL의 듀얼 체제 하에서 학술정보의 입수부터 배치적재까지의 일련의 과정을 정리하면 [표 1]과 같다.

작업	설명
1. 입수정보등록	입수관리자가 서비스K에 해당 차수의 입수정보를 등록.
2. 작업 지시	해당 차수의 입수정보를 입력관리자가 작업지시를 따라 작업체에서 작업지시
3. 서지입력	입력체에서 해당 차수의 XML생성 및 원문이미지 스캔.
4. 자료 업로드	작업디렉토리에 해당 차수의 XML과 원문이미지를 FTP를 이용하여 업로드.
5. 입수정보 등록	해당 차수의 서비스K에 등록된 입수정보를 작업K로 등록.
6. 오류경로	해당 차수의 XML의 오류경로와 원문이미지의 유효성 조사하여 적재가능한지 확인.
7. 배치적재할일 생성	해당 차수에 대해 오류경로가 잡힌 XML을 KRISTAL 적재할일 변환기당 생성 적재할일을 생성.
7.1.FTP전송	작업디렉토리가 NAS에 존재하지 않을 경우, 검색서버에서 FTP로 웹서버로 전송하여 적재할일을 기저와 검색서버에서 적재.
8. 작업K 배치적재	해당 차수의 배치적재할일을 작업K로 배치적재.
9. 원문자료 복사	적재완료 성공시 XML의 원문이미지를 원문저장디렉토리로 복사.
10. 웹서비스 작업K로 전환	웹서비스 대상 KRISTAL을 서비스K에서 작업K로 전환.
11. 작업K 입수완료 설정	작업K에 해당 차수의 입수완료 설정.
12. 서비스K 배치적재	해당 차수의 배치적재할일을 서비스K로 배치적재.
13. 웹서비스 작업K로 전환	웹서비스 대상 KRISTAL을 작업K에서 서비스K로 전환.
14. 서비스K 입수완료 설정	서비스K에 해당 차수의 입수완료 설정. 성공시와 웹서비스 상태도 동기화.

[표 1] 학술정보 데이터 DB구축 작업 순서

4. 결론

한국과학기술정보연구원에서 개발하여 운영하고 있는 학술정보관리시스템은 시스템에 종속적인 메타데이터의 형식과 관계형 DB와 KRISTAL-II의 공동 사용으로 인한 복잡한 관리 프로세스 그리고 다양하지 못한 검색 서비스 제공과 같은 단점을 가지고 있다. 본 논문에서는 데이터의 저장관리와 빠른 검색을 제공하는 KISTI의 최신 IRMS인 KRISTAL-2002만을 사용하여 새로운 학술정보관리시스템(ACMS)을 설계하고 구현하였다. ACMS는 학술정보의 처리를 메타데이터의 표준인 XML을 사용함으로써 데이터의 유효성 검증과 타 기관의 상호대차 및 원문교환을 손쉽게 되었다. 또한 학술정보의 입수, 가공, 처리의 전 과정을 간단하게 온라인으로 관리할 수 있게 되었으며 KRISTAL-2002에서 제공하는 다양하고 강력한 검색기능들을 이용자들에게 서비스할 수 있게 되었다. 또한 ACMS를 듀얼시스템으로 구축하여 시스템 장애 또는 관리실수와 같은 만일의 경우에 대비하였다.

5. 참고문헌

- [1] 한국과학기술정보연구원, 학회마을 (<http://society.kisti.re.kr>)
- [2] 한국과학기술정보연구원, "정보검색을 위한 효율적인 저장 시스템 개발", 1996.
- [3] 한국과학기술정보연구원, "정보검색을 위한 고성능 검색시스템 개발", 1996.
- [4] 조현양, 최선희, "학회 학술정보시스템 구축을 위한 연구", 한국도서관.정보학회, 제 30권 3호, pp405-425, 한국, 1999년 9월.
- [5] 한국과학기술정보연구원, "사용자 메뉴얼", 2004.
- [6] W3C, "Extensible Markup Language 1.0," <http://www.w3.org/TR/1998/REC-XML-19980210>, 1998.
- [7] W3C, XSL Transformations (XSLT), Recommendation, Oct. 1999. (<http://www.w3.org/TR/WD-xsl>)