

KRISTAL기반의 FSTA 구축을 위한 효율적인 관리 시스템 구현

윤희준, 이석형*, 윤화목**, 김한기***
 한국과학기술정보연구원

Efficient Management System Implementation for FSTA Database based on KRISTAL

Yoon hee-jun, Lee seok-hyoung*, Yoon hwa-mook**, Kim han-gi***
 Korea Institute of Science and Technology Information

요약

국내의 정보서비스 업체나 기관들은 세계적인 과학기술정보의 최신 동향과 연구 정보를 제공하기 위해 해외에서 발행된 데이터베이스를 도입, 구축한 정보를 서비스로 제공하고 있다. 여러 데이터베이스 중에서 FSTA(Food Science and Technology Abstracts)는 전 세계의 식품과학, 식품공학, 식품기술 및 음식과 관련된 식품 영양학에 관한 광범위한 초록 정보를 제공하는 데이터베이스이다. 식품과학에 관련된 정보를 유일하게 제공하는 FSTA 데이터베이스를 한국과학기술정보연구원(KISTI)에서 개발한 정보검색관리 시스템인 KRISTAL를 기반으로 데이터베이스를 구축하여 서비스를 제공할 수 있도록 데이터베이스 도입, 구축, 검색, 수정, 재구축, 통계 등의 여러 과정이 필요하다. 본 논문은 국내 이용자들에게 FSTA 데이터베이스의 정보를 서비스하기 위해 FSTA 데이터베이스를 입수하는 단계부터 KRISTAL 정보검색관리 시스템에 구축하는 단계까지의 과정을 체계적으로 구축하여 효율적으로 관리 할 수 있도록 하는 시스템을 구현하는데 있다. 보다 빠르고 정확한 FSTA 데이터베이스 구축을 통해 사용자들은 최신의 FSTA 데이터베이스를 검색하여 원하는 정보를 얻을 수 있다.

Abstract

Domestic information service authority and company provide database services to introduce the newest scientific and technological information, trend, and research data. Among many databases, FSTA((Food Science and Technology Abstracts)) is a bibliographic database providing worldwide coverage of all scientific and technological aspects of the processing and manufacturing of human food products. To service FSTA database through KRISTAL search engine developed by KISTI(Korea Institute of Science and Technology Information) need many process such as data obtainment, building, search, modification, rebuilding, and statistics. In this paper, we present implementation of management tool used to build FSTA database effectively. The management tool provides graphical user interface and controls all process from data obtainment to building of KRISTAL information searching system. By building FSTA database fast and accurately, user can get information that they want to find from the newest FSTA database.

I. 서론

과학기술의 중요성은 지식기반사회에서 더욱 강조되어 과학기술의 지식공유와 지식창출을 통해 국가의 과학기술 발전에 큰 역할을 하고 있다. 과학기술 정보의 양은 하루하루

기하급수로 증가하고 있으며 과학기술 정보에 대한 접근은 보다 쉬워지고 더욱 효율적으로 이루어지고 있는 반면 과학기술 정보는 원시적이고 비효율적인 방법으로 구축하고 있어 과학기술 정보의 품질이 저하되고 이용자들의 신뢰도가 낮아지는 문제점이 발생하고 있다. 단순한 데이터의 저장소의 역할만이 아닌 데이터의 검색과 이용을 위한 체계적인

시스템을 구축함으로써 이를 해소할 수 있다. 본 논문은 과학기술의 한 분야인 FSTA데이터베이스에 대한 구축과 관리를 위한 시스템을 구현하는데 있어 2장에서는 FSTA데이터베이스를 소개하고 3장은 DB의 구성과 시스템 기능 및 구조를 설명하며 4장은 FSTA 자료관리시스템에 대한 구축에 대해 서술하고 마지막으로 결론을 맺는다.

II. FSTA 데이터베이스

FSTA(Food Science and Technology Abstracts)는 전 세계의 식품과학과 식품기술 및 음식과 관련된 영양학에 대한 정보를 제공한 유일한 데이터베이스이다. 영국의 IFIS(International Food Information Service)에서 제공하며 영국의 CAB International, 미국의 Institute of Food Technologists(IFT) 그리고 독일의 Deutsche Landwirtschafts-Gesellschaft(DLG) 총 3개의 국가 기관이 대표로 있는 Board of Governors에 의해 운영되고 있다. 학술저널, 특허, 책, 회의록, 보고서, 논문, 표준, 법률에 대한 문헌 정보를 수록하고 있으며 40여 개국 이상의 다른 언어로 출판된 문헌의 abstracts를 모두 영문으로 제공하고 있다 [1].

FSTA장점	영국의 IFIS가 제공하는 식품관련분야만을 수록한 유일한 데이터베이스
FSTA수록범위	1969년부터 현재까지 75개국에서 수집한 1,800여종의 학술잡지의 기사가 83%, 20여 개국의 특허가 14%, 단행본 2%, 39개국의 규격이 1%를 수록
FSTA수록분야	Basic food science, Food microbiology, Food, engineering 등
갱신주기/건수	약1,500건/월
KISTI구축건수	708,011건(2006년 4월 30일 기준)

표 1. FSTA 데이터베이스 요약

III. DB 구성 및 시스템 기능 및 구조

1. FSTA 데이터베이스 구성

FSTA 데이터베이스는 입수된 원시데이터를 파싱하여 중

복체크와 필수데이터 유무 등의 절차에 따라 정상데이터 데이터베이스와 오류가 있는 데이터는 오류 데이터베이스로 분류되며, 입수와 관련된 정보를 저장하는 입수 데이터베이스, 데이터베이스 구축에 대한 정보를 관리하는 관리 데이터베이스 그리고 구축과 관련된 여러 통계 정보를 저장하는 통계 데이터베이스로 구성되어 있다. 그림 1은 FSTA 데이터베이스의 전체 구성도를 보여주고 있다.

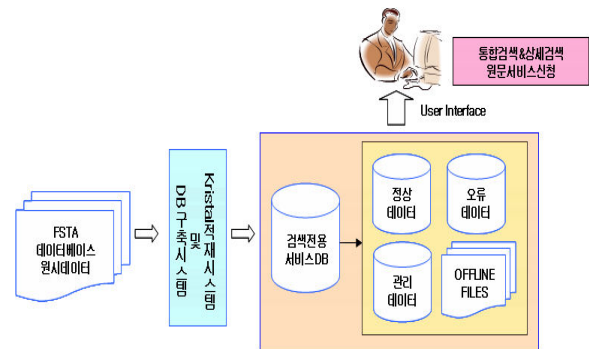


그림 1. FSTA 자료 관리 시스템의 DB구성

2. FSTA 시스템 기능 및 구조

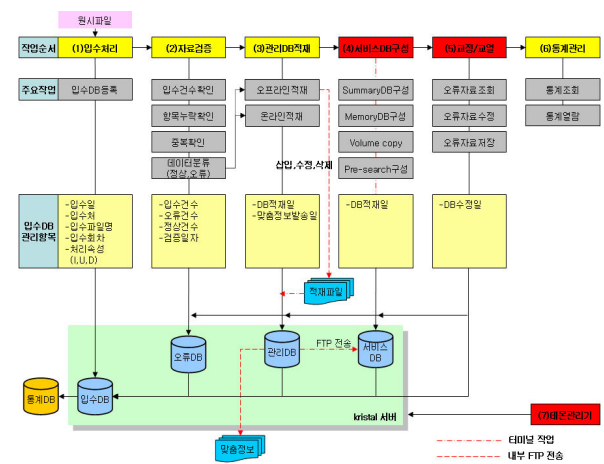


그림 2. KRISTAL 기반 FSTA 자료관리시스템 구조

KRISTAL 기반의 FSTA 자료관리구축시스템의 주요 기능은 입수된 원시데이터에 대해 입수일, 입수자, 입수화일명, 입수회차 등의 정보와 입수된 파일에 대해 자동으로 규칙적인 새로운 이름을 부여하도록 하는 원시 데이터 입수 등록 기능과 등록된 원시 데이터에 대해 입수건수확인, 필수항목 누락확인, 중복확인 등의 절차를 통해 정상데이터와 오류데이터로 분류하는 자료검증기능이 있다. 그리고 자료검증을 거쳐 적재방법에 따라 온라인 및 오프라인으로 적재할 수 있도록 하는 데이터베이스 적재기능과 오류데이터로 분류된 데이터에 대해 잘못된 부분을 수정할 수 있도록 하는 교정,

교열 기능 그리고 데이터베이스의 구축된 내용과 진행상황을 볼 수 있는 통계관리기능이 있다. 또한 FSTA 데이터베이스 자료관리시스템을 이중화된 구조로 구성하였기에 다른 FSTA 데이터베이스를 동기화시키는 서비스 DB 구성의 기능을 가지고 있다. 그림 2 는 전체적인 시스템의 구조를 보여준다.

IV. FSTA 자료관리시스템 구축

1. FSTA 데이터베이스 구축

전체적인 FSTA 데이터베이스 구축 절차는 먼저 IFIS의 FTP서버에 접속하여 FSTA 원시데이터를 입수한다. 입수된 원시데이터는 태그 형식의 일반적인 텍스트화일이다.

10	30	31	32
	35	36	40
	41	42	50
	52	55	56
	57	58	80
	81	301	302
	303	304	
2005-Ac0421 Nagarale, R. K.; Gohil, G. S.; Shahi, V. K.; Trivedi, G. S.; Thampy, S. K.; Rangarajan, R.			
2004 Studies on transport properties of short chain aliphatic carboxylic acids in electrodialytic separation. Journal Article			
Desalination 0011-9164 -#c 171-#c (2)			
195-204 14 ref. En [Cent. Salt & Marine Chem. Res. Inst., Bhavnagar 364 002, India, Tel. +91 (278) 256-9445, Fax +91 (278) 256-7562, E-mail salt(a)csir.res.in]			
An electrodiolysis process for the concentration of short chain aliphatic carboxylic acids, such as formic acid, propionic acid and acetic acid (which is used in foods and beverages, and as a substrate for the production of dietary supplements) is reported. Interpolymer-type ion-exchange membranes were used due to their high chemical stability and durability. Effects of concn., electric current and time on electrodiolysis were studied. A mathematical equation representing the water transport behaviour of the electrodiolysis process for concentrating acids under the influence of dif			

그림3. FSTA 원시데이터

```

@DATA.VIEW
#LAF=English
#SN=0011-9164
#DTV=Journal Article
#LA=en
#PFI=T
#ABF=T
#PY=2004
#JTE=Desalination
#IMF=#
#CP=FSTA
#ETN=20060504-1548
#ETD=20060504
#NT=171 (2) p 195-204
#IND=20060504-1548
#CN=FSTA200600000364
#LOF=3
#DK=studiesontransportpropertiesofshortchainaliphaticcarboxylicacidsinelectrodiolyticseparation2004
#ABE=An electrodiolysis process for the concentration of short chain aliphatic carboxylic acids, such as formic acid, propionic acid and acetic acid (which is used in foods and beverages, and as a substrate for the production of dietary supplements) is reported. Interpolymer-type ion-exchange membranes were used due to their high chemical stability and durability. Effects of concn., electric current and time on electrodiolysis were studied. A mathematical equation representing the water transport behaviour of the electrodiolysis process for concentrating acids under the influence of dif
#DT=J
#AIE=Nagarale, R. K.; Gohil, G. S.; Shahi, V. K.; Trivedi, G. S.; Thampy, S. K.; Rangarajan, R.
#ETF=FSTAFSTA-200603000
#SJT=A (Food Sciences): Chemistry and biochemistry
#PCN=2005-Ac0421
#CC=#c
#DBT=FSTA
#CSE=Cent. Salt & Marine Chem. Res. Inst., Bhavnagar 364 002, India, Tel. +91 (278) 256-9445, Fax +91 (278) 256-7562, E-mail salt@csir.res.in
#OPF=#
#TIE=Studies on transport properties of short chain aliphatic carboxylic acids in electrodiolytic separation.
    
```

그림4. 원시데이터로부터 변환된 KRISTAL 형식의 데이터

이 파일을 적절히 파싱하여 KRISTAL 검색 서버의 관리 DB에 적재될 수 있는 형식으로 변환시켜서 새로운 파일을 만든다. 그림4는 원시데이터로부터 KRISTAL 검색서버로 적재하기 위해 형식이 변환된 파일의 데이터의 화면이다.

그리고 온라인 적재나 오프라인 적재의 방법을 선택한 후 적재한다. 이 때 데이터의 오류 유무에 따라 FSTA 자료 관리 시스템의 정상 DB와 오류 DB에 적재된다. 적재가 끝난 후 정상 DB는 이중화된 다른 서버의 정상 DB와 동기화를 시켜주기 위해 서비스 DB를 구성하여 볼륨을 복사함으로써 FSTA 데이터베이스가 구축된다. 또한 오류 DB는 관리자가 일부 데이터를 수정하면 정상 DB로 이동되고 그 자료는 다시 이중화된 다른 서버의 정상 DB와 다시 동기화가 되어 관리자의 수정이 서비스 이용자들에게 바로 반영되도록 하였다.

2. FSTA 자료관리시스템 구축

FSTA 자료관리시스템은 자바 JDK 1.4를 기반으로 SWING 컴포넌트를 사용하여 Client/Server 구조로 구현되었다. 주요 클래스들의 설명은 표2와 같다.

Class	설명
mainDMS	Main Class
mainFrame	시스템의 환경변수와 메뉴 별기능호출
InsertDBProcessFSTA	FSTA 원시데이터를 파싱하여 KRISTAL 검색 서버에 적재
InsertRegisterData	원시자료를 등록
updateEDBDlg	
updateEDBFrame	에러 DB를 수정
updateEDBEditDlg	
updateMDBDlg	
updateMDBFrame	정상 DB를 수정
updateMDBEditDlg	
viewLoadingInfoDlg	구축현황 및 진행상황의 통계
xKristal2002	KRISTAL 검색서버와 통신하기 위한 API
CreateMemory	
CreateService	이중화된 서비스 DB를 구성
CreateSummary	
KdbServer	KRISTAL 서버구성
LoginDialog	프로그램실행시의 인증

표2. FSTA 자료관리시스템의 주요 클래스

원시데이터의 입수부터 검색 서버까지 적재를 위한 모든

절차를 GUI 환경으로 구현하였다. 특히 관리자들의 오류가 빈번한 부분을 GUI로 자동화하여 오류를 최소화시켰다.

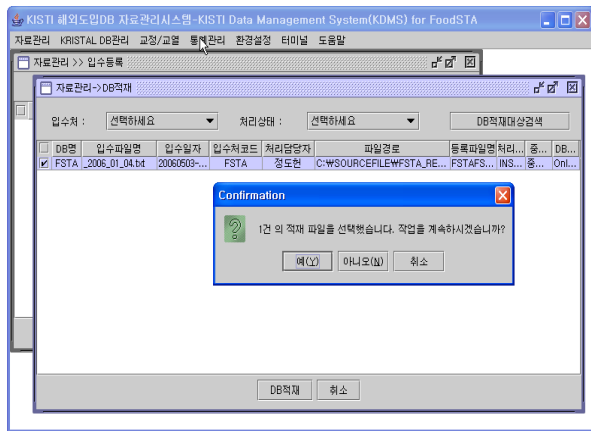


그림 5. FSTA 자료관리시스템의 입수등록 및 적재

그림5는 입수된 원시데이터 파일을 네이밍 규칙에 따라 새로운 파일명 바꾸고 파일명이 바뀐 원시데이터 파일을 KRISTAL 검색서버에 적재하기 위해 확인을 하는 화면이다.

또한 KRISTAL 검색 서버에 적재된 자료를 수정해야 할 경우는 원하는 자료를 몇 가지 조건으로 검색하여 온라인으로 직접 수정할 수 있다. 그림6은 제목을 study라는 단어로 검색한 결과와 그중의 첫 번째 자료를 수정하기 위해 더블 클릭했을 때의 화면을 나타낸다. 이 팝업화면에서 원하는 부분을 관리자가 수정하여 갱신할 수 있다.

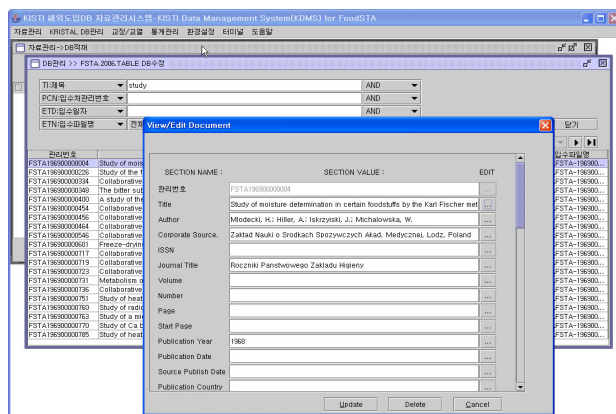


그림 6. 자료 검색과 온라인 수정

V. 결론

본 논문에서는 해외에서 도입되는 FSTA 데이터베이스를 서비스하기 위해서 원시데이터를 입수하여 등록하는 과정부터 데이터의 적절한 검증을 통해 정상데이터를 적재한 후

온라인으로 수정하도록 하는 FSTA 자료관리시스템을 구축하였다. 이 시스템을 통해 관리자가 해야 할 많은 일들을 자동화함으로써 체계적으로 관리가 가능하게 되었으며 데이터를 관리하는데 따른 알지 못하는 에러를 최소화시켰다. 또한 서비스되는 데이터를 온라인으로 직접 수정할 수 있도록 하여 효율적인 관리가 가능해졌다. 본 시스템을 좀 더 보완하고 기능을 추가하여 과학기술 정보의 지식공유와 연구활동을 지원하는데 이바지 할 것이라 판단된다.

참고 문헌

- [1] 한국과학기술정보연구원, “한국과학기술정보연구원 데이터베이스 백서”, 2003.
- [2] www.ifs.org
- [3] www.yeskisti.net