

EXOD

작성자 : 최지훈

생성일자 : 2008-03-05

문서 관리번호 : 문서 버전 : v2.5.0

문서 변경 내역

| 날짜 | 작성자 | 버전 | 수정사항 내용 |
|------------|-----|--------|-----------------|
| 2007.07.25 | 최지훈 | v2.0.0 | 추가 기능 및 기존내용 보강 |
| 2008.03.05 | 최지훈 | V2.5.0 | 추가 기능 및 기존내용 보강 |

<< 목 차 >>

| 1. EXOD | – 1 – |
|-------------------------------------|--------|
| 1.1. 특장점 | 1- |
| 1.2. 구성도 | 2- |
| 1.3. EXOD 관리범위 | 2- |
| 1.3.1. 데이터의 수집(SCD 생성) | 2 - |
| 1.3.2. 데이터(SCD)의 색인 | 2 - |
| 1.3.3. 색인에 대한 검색 | 2 - |
| 1.3.4. 공유메모리 및 프로세스 현황 확인 | 2 - |
| 1.4. EXOD 기능 리스트 | 3 - |
| 2. EXOD 설치 | – 4 – |
| 2.1. 설치 요구 사항 | 4- |
| 2.2. EXOD BUILD/SETUP 파일이 있는 디렉토리 | 4- |
| 2.3. EXOD BUILD/SETUP 옵션 | 5- |
| 2.4. EXOD BUILD | 6- |
| 2.5. EXOD SETUP | – 7 – |
| 2.6. EXOD DIRETORY 구조 | 8- |
| 2.7. EXOD 실행준비 | – 10 – |
| 2.7.1. CONF 디렉토리 | 10 - |
| 2.7.2. BIN 디렉토리 | 11 - |
| 2.8. EXOD 실행시키기 | – 14 – |
| 2.8.1. 색인 SERVER와 검색 SERVER가 하나인 경우 | 14 - |
| 2.8.2. 색인 SERVER | 14 - |
| 2.8.3. 검색 SERVER | - 15 - |
| 3. 수 집 | 16 - |
| 3.1. DB설정 | – 16 – |
| 3.1.1. 데이터베이스 등록 | 16 - |
| 3.1.2. 데이터베이스 등록 화면 | 16 - |
| 3.1.3. 데이터베이스 관리 | 17 - |
| 3.2. 수집기 | – 18 – |
| 3.2.1. 수집기 등록 | - 18 - |
| 3.2.2. 데이터베이스 미리보기 | - 18 - |
| 3.2.3. 수집기 관리 | - 22 - |
| 4. INDEX | 26 - |
| 4.1. COLLECTION | – 27 – |

| 4.2. | Collection(자동완성/동의어) | - 31 - |
|---|--|--|
| 4.3. | INDEXSCHEDULER | - 32 - |
| 4.4. | LOGINQUIRE | - 34 - |
| 4.5. | SHAREDMEMORY | - 35 - |
| 4.6. | REALTIMEINDEX | - 35 - |
| 4.7. | INDEXCHECK | - 38 - |
| 5. S | EARCH | 40 – |
| 5.1. | SEARCHSETUP | - 40 - |
| 5.2. | SEARCHINQUIRE | - 43 - |
| 5.3. | UsageInquire | - 45 - |
| 5.4. | 템플릿 | - 46 - |
| 5.4 | .1. 템플릿 불러오기 | 46 - |
| 5.4 | .2. 템플릿 변수 리스트 | 46 - |
| 5.4 | .3. 템플릿 등록 | 48 - |
| 5.4 | .4. 템플릿 검색 | 53 - |
| 5.5. | LOGSTATISTICS | - 54 - |
| 5.5 | .1. 통계 등록 | 54 - |
| 5.5 | .2. 통계 조회 | 54 - |
| 5.5 | .3. 상세 통계 | 55 - |
| 5.5 | .4. | - 55 - |
| | | 55 - |
| 5.6. | LogScheduler | |
| 5.6. 5.7. | | - 56 - |
| | SEARCH OPTION | - 56 - - 57 - |
| 5.7. 6. 人 | SEARCH OPTION | - 56 - - 57 - 59 - |
| 5.7. 6. A | SEARCH OPTIONㅡ | - 56 - - 57 - 59 - - 59 - |
| 5.7. 6. 人 6.1. 6.1 | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - |
| 5.7. 6. 人 6.1. 6.1 | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - 59 - |
| 5.7. 6. 人 6.1. 6.1 6.1 | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - 59 - 60 - |
| 5.7. 6. 人 6.1. 6.1 6.1 6.2. | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - 59 - 60 - - 60 - |
| 5.7. 6. 人 6.1. 6.1 6.1 6.2. | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - - 59 - 59 - 60 - - 60 - 60 - |
| 5.7. 6. 人 6.1. 6.1 6.1 6.2. 6.2 | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - - 59 - - 59 - - 60 - - 60 - - 60 - 60 - |
| 5.7. 6. 人 6.1. 6.1 6.1 6.2. 6.2 6.2 | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - - 59 - - 59 - - 60 - - 60 - - 60 - 60 - 60 - |
| 5.7. 6. 人 6.1. 6.1 6.1 6.2. 6.2 6.2 6.2 | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - - 59 - 59 - 60 - 60 - 60 - 61 - - 61 - |
| 5.7. 6. 人 6.1. 6.1 6.1 6.2. 6.2 6.2 6.2 6.3 6.3 | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - - 59 - - 59 - - 60 - - 60 - - 60 - - 61 - - 61 - - 61 - |
| 5.7. 6. 人 6.1. 6.1 6.1 6.2. 6.2 6.2 6.2 6.3 6.3 6.3 | SEARCH OPTION - USER DICTIONARY. 1. 사용자 사전 등록 2. 사용자 사전 별드 PREANALYSIS DICTIONARY. 1. 기분석 사전 등록 2. 기분석 사전 등록 3. 기분석 사전 편집 4.1. 가동완성 사전 등록 | - 56 - - 57 - 59 - - 59 - - 59 - - 59 - - 60 - - 60 - - 60 - - 61 - - 61 - 61 - 61 - |
| 5.7. 6. 人 6.1. 6.1 6.1 6.2. 6.2 6.2 6.2 6.3 6.3 6.3 6.3 | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - - 59 - - 60 - - 60 - - 60 - - 61 - - 61 - - 61 - - 61 - - 62 - |
| 5.7. 6. 人 6.1. 6.1 6.1 6.2. 6.2 6.2 6.3 6.3 6.3 6.4. 6.4 | SEARCH OPTION | - 56 - - 57 - 59 - - 59 - - 59 - - 59 - - 60 - - 60 - - 60 - - 61 - - 61 - - 61 - - 61 - - 62 - 62 - |

| 6.5.1. 사전 배포 등록 | 63 - |
|--------------------------|--------|
| 6.5.2. 사전 배포 실행 | - 63 - |
| 7. 부가기능 | 64 - |
| 7.1. System Monitor | – 64 – |
| 7.2. Process Monitor | – 65 – |
| 7.3. OTHERS SCHEDULER | – 65 – |
| 7.4. Morph Analyzer | – 66 – |
| 7.5. DAEMON CONTROLER | – 66 – |
| 7.6. IMAGE VIEWER | – 67 – |
| 7.7. 영한 자동 변환 | – 67 – |
| 8. REAL TIME INDEX | 68 - |
| 8.1. RUN SET에 바로 적용 | – 69 – |
| 8.1.1. INSERT | 69 - |
| 8.1.2. DELETE | 70 - |
| 8.2. ONLINE SET을 만들어서 적용 | – 70 – |
| 8.2.1. INSERT | 70 - |
| 8.2.2. DELETE | 71 - |
| 9 ADD ON | - 71 - |

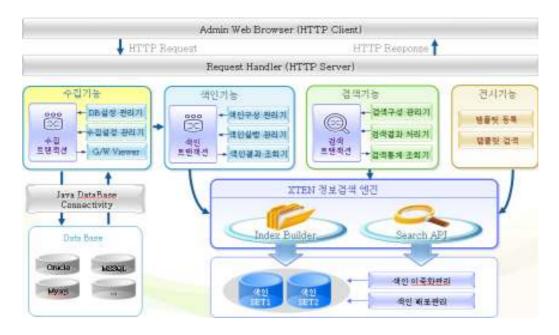
1. EXOD

EXOD 관리도구는 XTEN 정보검색엔진과 결합하여 문서 검색의 기본이 되는 일련의 과정 [수집], [색인], [배포], [검색], [전시] 에 대한 관리기능 및 기타 유용한 기능들을 웹 인터페이스를 통해 제공하는 XTEN 검색엔진 통합 관리모듈

1.1. 특장점

- 웹 인터페이스를 통한 관리기능
 - 쉽고 편리한 웹 인터페이스를 통하여 검색 서비스 개발 활용도 향상
- DB Gateway Viewer
 - 데이터베이스의 테이블, 필드정보, 데이터 등을 확인 할 수 있는 G/W Viewer를 사용하여 효율적인 데이터 분석과 데이터 가공내역 확인 가능
- 색인 데이터의 다중 배포
 - 색인배포 기능을 이용하여 색인 데이터를 다수의 서버에 전달 함으로써 멀티서버 운용 시 강점
- 결과출력을 위한 템플릿 에디터 제공
 - html, xml, xls 등 다양한 형태로 검색 결과물 생성
- 다양한 색인, 검색 방법 지원
 - XTEN 검색 엔진에서 제공하는 다양한 색인, 검색 방법들을 human readable 환경에서 모두 지원
- 실시간 색인 기능 강화
 - 여러 문서의 삽입, 삭제 기능이 강화 되었으며 멀티 서버 운영 시 실시간 색인의 동기화 기능 지원
- 기능별 스케쥴 관리 시스템
 - 각 기능별 스케쥴러를 이용하여 자동색인주기 설정, 실시간색인주기 설정, 검색결과 로그 통계주기 설정 및 기타 스크립트 주기 설정 등에 대한 효율적인 관리
- 사전관리를 통한 재현율 및 정확률 향상
 - 기분석 및 사용자 사전 관리 기능제공
 - 자동완성 및 동의어 사전 관리 기능 제공
- 유니코드 지원
 - 각기 다른 언어들을 하나의 코드 형식으로 처리
- 자동완성 기능 제공
 - 추천 키워드 및 이슈가 있는 키워드를 상위로 조절하여 사이트 특성에 맞는 검색을 유도
- 동의어/유사어 검색 기능 제공
 - 한번의 질의어 입력으로 동의어 및 유사어에 대한 검색
 - 영어->한글 입력 시 유사발음에 대한 보정 기능

1.2. 구성도



1.3. EXOD 관리범위

1.3.1. 데이터의 수집(SCD 생성)

EXOD에서 제공하는 crawler를 통해서 my-sql, oracle, ms-sql에서 데이터를 읽어서 SCD를 생성한다. 또는 사용자가 검색 데이터에 맞는 G/W, 필터, web crawler를 통해서 SCD를 생성한다.

1.3.2. 데이터(SCD)의 색인

검색 데이터에 맞는 G/W, 필터, web crawler 또는 EXOD crawler를 통해 추출된 SCDM 대한 색인의 설정, 실행의 일련 과정이다.

1.3.3. 색인에 대한 검색

- 색인 컬렉션에 대한 다양한 검색 설정을 할 수 있다.
- 검색 결과는 기본으로 제공되는 TXT, XML, XLS의 형태로 볼 수 있고, 템플릿 메뉴에서 해당 컬렉션 의 템플릿을 생성한 경우 사용자가 작성한 템플릿으로 결과를 확인할 수 있다.

1.3.4. 공유메모리 및 프로세스 현황 확인

- 공유메모리를 관리 할 수 있다.
- 검색 프로세스 점유 현황을 확인 할 수 있다.

1.4. EXOD 기능 리스트

| 대메 | 소 메 뉴 | 기 능 |
|--------|------------------|---|
| 수 | DB설정 | DB 접속 정보 등록/수정, 삭제 |
| | 수집기 | sql query를 이용하여 수집 설정파일 생성 필드별 데이터가공, DB 미리보기, 수집 데이터 미리보기 |
| 색 | 인 컬렉션설정 | 수집기 및 자동완성/동의어 사전을 이용하여 새로운 컬렉션을 등록 Index, Partition, Date, Sort 볼륨 정보 설정 형태소 분석 방식, 옵션, 중요도, 가중치공식 등을 적용하여 색인 등록 |
| | 색인스케줄러 | 배치/실시간 색인 예약실행 등록 및 강제실행/종료 수행 |
| | 색인로그조회 | 색인 프로세스에 대한 실시간 상태 조회 |
| | 공유메모리제어 | 검색처리를 위한 공유메모리의 수동관리 |
| | 실시간색인 | 실시간 색인 등록 및 INSTRUCTION 테스트 |
| | 색인확인 | 문서 내 색인어 확인 및 색인어가 포함된 문서 확인 |
| UH : | 또 배포서버 지정 | 배포 서버의 IP와 전송포트 번호 지정 |
| | 검색서버 지정 | 검색 서버의 IP와 전송포트 번호 지정 |
| 검 | 백 구성설정 | 컬렉션 정보를 기반으로 한 검색설정 구성 분야/범위 등에 대한 제한검색 설정 및 분야/범위/가나다 정렬설정 동의어 검색을 위한 질의어 확장 설정 |
| | 검색결과조회 | HTML, XML, XLS형태로 검색결과 조회 및 출력 |
| | 사용량 조회 | 일, 시, 분 단위의 검색처리통계 및 그래프 제공 |
| | 검색로그 통계 설정 | 통계 등록, 통계 조회, 통계 상세 정보보기, XML형태로 결과 출력 기능 |
| | 검색로그통계스케줄러 | 로그 통계 예약실행 등록 및 강제실행/종료 수행 |
| 전 . | 시 템플릿설정 | 템플릿 폼에 소스 코드 등록 및 검색 템플릿 페이지 화면 출력 |
| 사 | 전 사용자사전 | 사용자 사전 등록, 편집 및 빌드 (XTEN3.5 이상) |
| | 기분석사전 | 기분석 사전 등록, 편집 및 빌드 (XTEN3.5 이상) |
| | 자동완성사전 | 자동완성 사전 등록 및 편집 |
| | 동의어사전 | 동의어 사전 등록 및 편집 |
| | 사전배포 | 사전 배포 |
| 부가기: | 5 전체 등록정보 | 수집, 색인, 검색 단의 통합 등록 정보 확인 |
| | 시스템 모니터 | 유닉스 커널, UPTIME, USER, CPU, MEMORY, DISK 정보 및 상태 출력 |
| | 프로세스 모니터 | EXOD 프로세스 테이블의 내용을 출력 |
| | 기타 스케줄러 | 기타 스크립트 예약실행 등록 및 강제실행/종료 수행 |
| | 형태소 분석 테스트 | 사용자 입력 쿼리를 이용하여 형태소 분석 테스트 결과 출력 |
| | 데몬관리 | EXOD에서 사용하는 데몬 관리 |
| Add Or | | EXOD외에 타 웹 인터페이스 관리기 및 사이트 링크 |

2. **EXOD** 설치

2.1. 설치 요구 사항

● H/W 권장 사항

- CPU: 2.0Ghz 이상 - RAM: 1Gbytes 이상

- HDD: 10,000RPM, 150Gbytes 이상

● 지원 O/S

- Linux, Sun Solaris, HP-UX, IBM AIX, Microsoft Windows NT/2000/2003

● 지원 DBMS

- Oracle, MS-SQL, MySQL 등

● 소프트웨어 권장사항

- java 1.4 이상
- gcc 4.0 / Visual Studio 6.0 이상

2.2. EXOD build/setup 파일이 있는 디렉토리

```
jhchoi@search:~/SubVersion/exod2.0$ Is

Documents exod.build lib src utf_src_gen.sh
jhchoi@search:~/SubVersion/exod2.0$
```

[그림 2.1] EXOD build 폴더

```
jhchoi@search:~/SubVersion/exod2.0/EXOD_2.5_EUC-KR_Linux$ ls

bin cgi exod.setup g.w imgs tpls xten

jhchoi@search:~/SubVersion/exod2.0/EXOD_2.5_EUC-KR_Linux$
```

[그림 2.2] EXOD setup 폴더

2.3. EXOD build/setup 옵션

● prefix: EXOD가 build/setup 될 디렉토리

● xten : XTEN의 위치

● lang : Server의 \$LANG (XTEN3.5 이상)

● mode: daemon의 구동모드

[그림 2.3] EXOD build 옵션

```
jhchoi@search:~/SubVersion/exod2.0/EXOD_2.5_EUC-KR_Linux$ ./exod.setup

+ Warning: Configuring EXOD with default settings.
+ This is probably not what you really want.
+ first or at least run './exod.setup --help' for
+ a compact summary of available options.

Usage: ./exod.setup [options]
Options: [defaults in brackets after descriptions]
General options:
    --prefix=/home/openbase/exod install target path (must complete path)
    --mode=cgi cgi | daemon (default cgi)
    --help,-h print this message

Jhchoi@search:~/SubVersion/exod2.0/EXOD_2.5_EUC-KR_Linux$
```

[그림 2.4] EXOD setup 옵션

2.4. EXOD build

● ./exod.build --prefix=/EXOD가 build될 경로 --xten=/XTEN의 위치 (XTEN3.5 이상 : 유니코드 사용시 --lang=UTF-8 or --lang=EUC-KR 로 언어옵션 설정)

```
./exod.build --prefix=./ --xten=./xten --mode=EUC-KR
```

[그림 2.5] EXOD build 명령과 옵션

● prefix와 xten의 경로가 이상이 없을 경우 EXOD build가 시작된다. XTEN엔진의 정보와 EXOD가 컴파일될 디렉토리를 보여준다. 필요한 directory를 생성한 후 gcc를 setting하고, 설치되는 서버의 운영체제의 종류를 확인하고 Compile에 들어간다.

```
______
 EXOD 2.5 Solutions Build Info:
 + EXOD
                   lang: EUC-KR
            Platform: i686-whatever-linux22
  index run: ./EXOD_2.5_EUC-KR_Linux/+INDEXER
search info: ./EXOD_2.5_EUC-KR_Linux/+SRCHINF
index dbs: ./EXOD_2.5_EUC-KR_Linux/_INDEXDB
     create install directory at .
./EXOD_2.5_EUC-KR_Linux
./EXOD_2.5_EUC-KR_Linux/bin
./EXOD_2.5_EUC-KR_Linux/cgi
./EXOD_2.5_EUC-KR_Linux/g.w
./EXOD_2.5_EUC-KR_Linux/g.w/classes
./EXOD_2.5_EUC-KR_Linux/g.w/jdbc_driver
./EXOD_2.5_EUC-KR_Linux/g.w/jdbc_driver/oracle
./EXOD_2.5_EUC-KR_Linux/g.w/jdbc_driver/mysql
./EXOD_2.5_EUC-KR_Linux/g.w/jdbc_driver/mssql
./EXOD_2.5_EUC-KR_Linux/g.w/jdbc_driver/mssql
./EXOD_2.5_EUC-KR_Linux/tpls
./EXOD_2.5_EUC-KR_Linux/tpls
./EXOD_2.5_EUC-KR_Linux/tpls
setting C compiler to gcc
 + create install directory at .
 + setting C compiler to gcc
 + configured for Linux platform
    Copy tpls & imgs
   Copy Xten Search Engine
 + Compile programs
       doing..
      please check compile.log
Build Complete
```

[그림 2.6] EXOD build 과정

2.5. EXOD setup

● ./exod.setup --prefix=/EXOD가 setup 될 경로 -mode=daemon or -mode=cgi

```
./exod.setup --prefix=/home/jhchoi/exod2.0 --mode=daemon
```

[그림 2.7] EXOD setup 명령과 옵션

● prefix가 이상이 없을 경우 EXOD setup이 시작된다. EXOD가 설치될 디렉토리를 생성하고, 스크립트, 설정파일, 템플릿 파일, EXOD 바이너리 파일, 사전, XTEN엔진 등을 복사한다.

[그림 2.8] EXOD setup 과정

2.6. EXOD diretory 구조

| 폴 더 | 하위폴더 | 역 할 및 내 용 |
|----------|---------------|---|
| xten | bin | XTEN binary 파일 폴더 |
| | lib | XTEN shared library 폴더 |
| | lib/DICT | XTEN 사전 폴더 |
| bin | | EXOD에서 사용되는 스크립트 및 명령실행 파일들이 들어있다. 이 폴더에 |
| | | 서 EXOD를 실행시키고 종료시킴 |
| conf | | EXOD 설정 파일들 |
| cgi | | EXOD의 개별 CGI module |
| data | | scheduler 정보와 process 정보 |
| logs | | EXOD의 log 와 sspd, itpd, synd의 pid가 들어있는 폴더 |
| imgs | | EXOD에서 사용되는 image들이 들어있는 폴더 |
| +CRAWLER | G.W | 수집기에서 사용될 jdbc driver가 복사되고 class파일 |
| | dbinf | 수집할 Database의 정보를 가진 설정 파일 |
| | inf | 수집기 정보를 가진 설정 파일 |
| | log | 수집기 log를 저장 |
| | qry | 수집기 query를 저장 |
| | run | 수집기 스크립트 |
| +INDEXER | crl | 수집기를 실행시키는 컬렉션.crl.sh 스크립트가 생성 |
| | Inf | 컬렉션 정보를가지고 있는 설정 파일들이 생성 |
| | log | 색인, 실시간색인, 수집 및 DOCID를 얻어오는 과정에 대한 log와 실시간 색인에 대한 pid를 저장 |
| | pst | 색인 server가 검색 server에 배포할 파일이 있을 경우 색인 server의 정 보와 배포할 파일정보가 들어있는 컬렉션.syn 파일이 생성 |
| | swt | 현재 service되고 있는 SET의 번호가 저장되는 컬렉션.swt파일 |
| | run | 컬렉션.idx.sh, 컬렉션.seed.sh, 컬렉션.onidx.sh 스크립트가 생성 |
| | rtinstruction | 실시간 색인을 실행하기 위한 폴더이다. 하위구조로는 rtinstruction/컬렉션/ 폴더가 컬렉션이 생성되면 생기고, 실시간 색인 instruction을 파일로 집어넣는 폴더. 만약 검색 server가 설정되었으면 실시간 색인 instruction을 분석해서 SCD와 DocID List로 배포한다. 컬렉션.seed.sh가 스케쥴러에 의해서 주기적으로 실행. 컬렉션.seed.sh가 instruction에 대한 처리를하고 SCD와 DOC ID List 파일들을 생성한 후 배포를 하고 파일들을 rtrun/컬렉션 폴더로 이동 |

| | rtransaction | 실시간 색인을 실행하는 파일들을 포함하는 폴더. 하위구조로는 rtrun/컬 | | | | |
|----------|--------------|---|--|--|--|--|
| | rtiansaction | 레션/ 폴더가 컬렉션이 생성되면 생기고, 실시간 색인에서 instruction에 | | | | |
| | | 서 얻은 SCD와 DOC ID List. 컬렉션.onidx.sh가 실시간 색인 thread에 | | | | |
| | | 의해서 주기적으로 실행. 컬렉션.onidx.sh가 SCD와 DOC ID에 대한 처리. | | | | |
| | tpl | run 폴더와 crl 폴더에서 사용되는 스크립트들을 생성하기위한 templet. | | | | |
| _INDEXDB | db1 | 하위폴더 구조로는 db1/컬렉션/dat, idx, loc, tmp 폴더를 가지고 있다. | | | | |
| | | SCD및 색인파일, Location정보들을 포함 | | | | |
| | db2 | db1과 같은 구조 | | | | |
| | set1 | db1에 대한 symbolic Link | | | | |
| | set2 | db2에 대한 symbolic Link | | | | |
| | setonline | 실시간 색인에 사용되는 SET이다. 다른 SET들과 같은 구조 | | | | |
| | spec | 실제 service되고 있는 SET의 specfile에 대한 symbolic Link인 컬렉 | | | | |
| | | 션.spec과 실시간 색인 SET의 specfile에 대한 symbolic Link인 컬렉 | | | | |
| | | 션.spec.on 이 생성 | | | | |
| +SRCHINF | inf | 검색설정에서 대한 설정 파일 | | | | |
| | qry | 하위폴더구조로는 qry/검색명/ 이고, 각 검색에 대한 날짜별 검색Log들을 | | | | |
| | | 포함 | | | | |
| | tpl | 하위폴더구조로는 tpl/검색명/ 이고, 각 검색에 대한 templet이 생성 | | | | |
| +STATIST | dat | 키워드 통계의 데이터가 저장 | | | | |
| | inf | 키워드 통계 설정 파일 | | | | |
| | log | 키워드 통계 log 파일 | | | | |
| | run | 키워드 통계 데이터를 생성하는 실행 스크립트 | | | | |
| | tpl | 키워드 통계 실행 스크립트를 생성하기위한 templet | | | | |
| DICT | | 사전을 EXOD에서 관리하기 위한 폴더. | | | | |
| | | XTEN엔진의 /lib/DICT/폴더를 복사 (XTEN3.5 이상) | | | | |
| | TEMPALTE | XTEN엔진에서 제공하는 형태소 분석방법들에 대한 TEMPLATE | | | | |
| | inf | 사전 배포에 대한 inf 파일 | | | | |
| | usr | 사용자 사전을 관리하기 위한 폴더 (XTEN3.5 이상) | | | | |
| | pre | 기분석 사전을 관리하기 위한 폴더 (XTEN3.5 이상) | | | | |
| | autocom | 자동완성 사전을 관리하기 위한 폴더 | | | | |
| | synonym | 동의어/유사어 사전을 관리하기 위한 폴더 | | | | |
| | | | | | | |

2.7. EXOD 실행준비

● EXOD가 설치된 디렉토리로 이동하면 아래와 같은 구조의 폴더들을 확인 할 수 있다.

+CRAWLER +INDEXER +SRCHINF +STATIST DICT _INDEXDB bin cgi conf data imgs logs xten

[그림 2.9] EXOD 설치폴더

2.7.1. conf 디렉토리

● EXOD의 설정파일과 system specfile이 들어가 있다.

```
addon.conf
            itpd.conf spec.sys sspd.conf synd.conf user.conf
<u>ihchoi@sea</u>rch:~/exod2.0/conf$ cat sspd.conf
PORT=7592
PMAX=120
TMAX=120
<u>|hchol@sea</u>rch:~/exod2.0/conf$ cat itpd.conf
[ITPD]
PORT=7280
inchoi@search:~/exod2.0/conf$ cat synd.conf
[USER]
admin=admin
Theholmsearch:~/exod2_0/conf$_cat_spec_sys
                         4096
page_size
extent_size
                         16
                         256
num_of_buffers
num_of_openvolums
                         128
num_of_openfiles
                         1024
num_of_wanted
                         500
volctl_filename
                         /home/jhchoi/exod2.D/xten/VOLUME.CTL
key_filename
                         /home/jhchoi/exod2.0/xten/bin/xten.key
ihchoi@search:~/exod2.U/conf$ cat addon.conf
[addon]
name=Add0n
About Xten Exod=http://xten.openbase.co.kr
```

[그림 2.10] conf 폴더

sspd.conf

- PORT : EXOD가 service되는 port번호, 기본으로 **7578**로 설정되어 있다.

- PMAX: MAX Process 개수, 기본으로 60개를 잡고 있다.

- TMAX : 검색의 최장시간, 기본으로 15로 되어 있다.

itpd.conf

- PORT: itpd daemon이 사용할 port번호, 기본으로 **7280**으로 설정되어 있다.

user.conf

- EXOD 사용자 ID와 PASSWORD를 등록한다. ex) admin=admin

spec.sys

- system specfile이다. (XTEN manual 참조)

addon.conf

- addon 메뉴에 보여질 링크이름=링크주소

2.7.2. bin 디렉토리

| exod_DeleteDoc | exod_Statis | itpctl | sspd |
|----------------------|------------------------|---------|-------------|
| exod_Dict2SCDAutoCom | exod_TXT2BINARYAutoCom | itpd | sspd.cgi |
| exod_Dict2SCDSynonym | exod_TXT2BINARYSynonym | moses | sspd.daemon |
| exod_GetDocID | exod_TXT2DictPre | shmp | synctl |
| exod_InsertDoc | itpc | sspctl | synd |
| exou_filserroc | TEPC | 22hcr I | Syliu |

[그림 2.11] bin 폴더

sspctl

- sspd daemon을 실행, 종료. 궁극적으로 EXOD를 실행시키는 명령어

```
[ SSPD SERVICE HELP ]
- usage : ./sspctl (start|stop|restart|help)
start - start sspd with load shm memory
stop - stop sspd with unload shm memory
restart- restart sspd only
help - this screen
```

[그림 2.12] sspctl 사용법

itpctl

- Itpd daemon을 실행, 종료. Itpd는 itpc가 특정 instruction과 path1과 path2를 보냈을 때 instruction에 해당하는 동작을 수행

```
[ ITPD SERVICE HELP ]
- usage : ./itpctl (start|stop|restart|help)
start - start itpd with load shm memory
stop - stop itpd with unload shm memory
restart- restart itpd only
help - this screen
```

[그림 2.13] itpctl 사용법

itpc

- Itpc는 itpd(server)에게 특정 instruction과 path1과 path2를 전송하는데 사용

```
./itpc server_ip:port instruction:|get|put|pst path1 path2
pst pst_path pst_name
get remote_path local_path
put local_path remote_path
ins remote_path instruction_name
```

[그림 2.14] itpc 사용법

synctl

- synd daemon을 실행, 종료. +INDEXER/pst/컬렉션이름.pst파일이 있을 경우 색인 server로부터 변경 사항을 받아오는 daemon

```
[ SYND SERVICE HELP ]
- usage : ./synct! (start|stop|restart|help)
start - start synd with load shm memory
stop - stop synd with unload shm memory
restart- restart synd only
help - this screen
```

[그림 2.15] synctl 사용법

shmp

각 색인 파일들을 공유메모리에 올리고 내리는 명령어

```
./shmp init|final idx_dir
```

[그림 2.16] shmp 사용법

exod_InsertDoc

실시간 색인시 문서를 삽입하는 명령어

```
usage : ./exod_InsertDoc sysspec usrspec scd_fname doc_start [pid_fname]
```

[그림 2.17] exod_InsertDoc 사용법

exod_DeleteDoc

실시간 색인시 DOCID로 문서를 삭제하는 명령어

```
usage : ./exod_DeleteDoc syspec usrspec D|Q|DF|QF DocID|Query|DocIDFile|QueryFile
ex)
delete from docid : sysspec usrspec D 13
delete from query : sysspec usrspec Q "idxname <indexed_boolean>abc"
delete from docid file : sysspec usrspec L docid_fname"
delete from query file : sysspec usrspec F query_fname"
```

[그림 2.18] exod_DeleteDoc 사용법

exod_GetDocID

실시간 색인시 지울문서의 DOC ID를 얻어오는 명령어

```
usage : ./exod_GetDocID syspec usrspec QIF QueryStr|QueryFile
ex)
getdoc from query : sysspec usrspec Q "idxname <indexed_boolean>abc"
getdoc from file : sysspec usrspec F getdoc.qry"
- contents of getdoc.qry : idxname <indexed_boolean>abc
```

exod_Dict2SCDAutoCom

자동완성 사전의 내용을 SCD로 변환

usage : ./exod_Dict2SCDAutoCom dat_fname scd_fname

[그림 2.20] exod_Dict2SCDAutoCom 사용법

exod_Dict2SCDSynonym

동의어/유사어 사전의 내용을 SCD로 변환

usage : ./exod_Dict2SCDSynonym dat_fname scd_fname

[그림 2.21] exod_Dict2SCDSynonym 사용법

exod_TXT2BINARYAutoCom

자동완성 사전을 구축하기 위한 명령어. 자동완성에 반영할 단어들을 한줄씩 적어서 TXT파일로 만든 파일을 자동완성 사전으로 변환

usage : ./exod_TXT2BINARYAutoCom autcom_txt_fname autocom_dat_fname

[그림 2.22] exod_TXT2BINARYAutoCom 사용법

exod_TXT2BINARYSynonym

동의어 사전을 구축하기 위한 명령어. 자동완성에 반영할 단어들을 한줄에 "키워드 동의어 동의어..." 로 적어서 TXT파일로 만든파일을 자동완성 사전으로 변환

usage : ./exod_TXT2BINARYAutoCom autcom_txt_fname autocom_dat_fname

[그림 2.23] exod_TXT2BINARYSynonym 사용법

moses

EXOD 전용 스크립트(moses) 구동 바이너리

usage : ./moses filename.moses

[그림 2.24] moses 사용법

2.8. EXOD 실행시키기

2.8.1. 색인 server와 검색 server가 하나인 경우

● 색인하는 server와 검색하는 server가 하나일 때는 sspctl만 실행

```
jhchoi@search:~/exod2.0/bin$ ./sspctl start

[ SSPD SERVICE START ]
  - sspd service daemon load : ,.0K : port(7592)
  - dict shm memory load : ,.0K
  - index shm memory load set1 : ..0K
  - index shm memory load set2 : ,.0K
```

[그림 2.25] 단일 server 실행 방법

2.8.2. 색인 server

● 색인 server의 경우 EXOD를 실행시키는 sspctl과 색인파일에 변경이 있을 경우 알리고 파일을 전송할 수 있는 itpctl을 실행

[그림 2.26] 색인 server 실행 방법

2.8.3. 검색 server

● 검색 server의 경우 EXOD를 실행시키는 sspctl과 색인 server의 색인파일에 변경이 있을 경우 메시지를 받고 파일을 받아올 수 있는 itpctl과 +INDEXER/pst/컬렉션이름.pst가 생성되었을 때 컬렉션이름.pst파일에 적힌 정보를 보고 색인 server로부터 변경사항을 받아오는 synctl을 실행

```
Jhchoi@search:~/exod2.0/bin$ ./sspctl start

[ SSPD SERVICE START ]
- sspd service daemon load : ...0K : port(7592)
- dict shm memory load : ...0K
- index shm memory load set1 : ...0K
- index shm memory load set2 : ...0K

Jhchoi@search:~/exod2.0/bin$ ./itpctl start

[ ITPD SERVICE START ]
- itpd service daemon load : ...0K : port(7291)

Jhchoi@search:~/exod2.0/bin$ ./synctl start

[ SYND SERVICE START ]
- synd service daemon load : ...0K
```

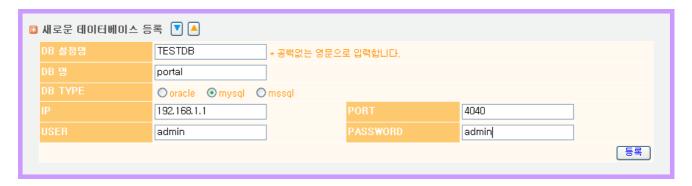
[그림 2.27] 검색 server 실행 방법

3. 수 집

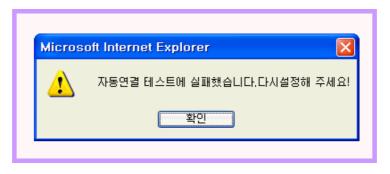
3.1. DB설정

3.1.1. 데이터베이스 등록

- 데이터를 수집하기 위한 설정 파일 생성 작업으로 새로운 데이터 베이스를 등록 할 수 있다.
- 등록 가능한 DB 타입은 기본적으로 Oracle, Mysql, Mssql 3가지를 제공한다.
- 정보를 입력후 등록 버튼을 누르면 설정된 정보를 이용하여 자동으로 연결 테스트를 하게 되며 그 결과를 화면에서 모니터링 할 수 있다.(정상: OK, 비정상: NO)



[그림 3.1] 데이터베이스 등록 화면



[그림 3.2] 데이터베이스 등록시 연결테스트 실패 화면

3.1.2. 데이터베이스 등록 화면

● DB 설정명 : 등록할 데이터베이스의 파일명

● DB 명 : 접속할 데이터베이스의 이름(ORACLE에서는 SID를 의미함)

● DB TYPE: 데이터베이스의 타입 (ORACLE, MYSQL, MSSQL)

● IP: 데이터베이스 서버의 접속 주소 입력 ex) 211.239.160.172

● PORT : 데이터베이스 서버의 접속 포트번호 입력

● USER: 데이터베이스 서버의 접속 사용자 계정 입력

● PASSWORD : 데이터베이스 서버의 접속 사용자 암호 입력창

3.1.3. 데이터베이스 관리

- 저장된 데이터베이스의 설정 정보를 확인할 수 있다.
- 선택된 설정에 한하여 적용(수정), 삭제가 가능하며 데이터베이스 설정 리스트를 통해 전체 데이터베이스를 확인 할 수 있다.
- 설정 리스트의 DB_NAME을 선택하면 각각의 DB설정 정보 창으로 이동이 가능하다.



[그림 3.3] 데이터베이스 관리 화면

- 데이터베이스 관리 화면(상세 내역중 '데이터베이스 등록'과 <mark>중복</mark>되는 설명은 제외한다.)
 - 연결 테스트: 자동연결 테스트를 통하여 정상적으로 데이터베이스에 연결 되었는지 여부 판단 (정상 : OK, 비정상: NO)

- 등록일자 : 최초로 설정된 날짜

- 수정일자 : 설정이 변경된 날짜

3.2. 수집기

3.2.1. 수집기 등록

- 수집기 등록의 1차 작업으로 기본 정보를 등록 한다.
- 기본정보를 입력하고 등록 버튼을 누르면 상세 정보를 입력하기 위한 세부설정 페이지로 이동한다.



[그림 3.4] 수집기 등록 화면

● 수집기 등록 화면

- 수집기명 : 등록할 수집기의 파일명 입력

- DB선택 : 데이터베이스에서 등록항 DB설정을 선택

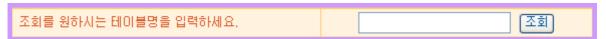
- SQL 쿼리 입력 : 수집 SQL 쿼리를 입력한다. (DB의 보안을 위하여 CREATE, INSERT, UPDATE, DELETE, DROP등의 키워드 들을 입력할 수 없게 처리)

3.2.2. 데이터베이스 미리보기

● 설정된 데이터베이스나 수집기를 이용하여 데이터를 가져올 데이터베이스의 정보를 확인 할 수 있다.



[그림 3.5] 수집기 미리보기 Main 화면



[그림 3.6] 테이블 조회 입력창 화면



[그림 3.7] 쿼리 조회 입력창 화면

- 해당 데이터 베이스의 전체 정보를 확인할 수 있다.
- 상단의 테이블 조회 기능을 이용하여 LIKE 문을 이용한 테이블 조회가 가능하다.
- 상단의 쿼리 입력폼을 이용하여 원하는 쿼리를 입력하여 결과 확인이 가능하다.

| - 테이블 리 | 스트 전체 테이블 건수 : [19 | 5] | | |
|---------|--------------------|-------|------|-------|
| index | 테이블명 | 테이블타입 | 필드확인 | 데이터확인 |
| 1 | CDIR | | 보기 | 보기 |
| 2 | CDIRSITE | | 보기 | 보기 |
| 3 | CREQ | | 보기 | 보기 |
| 4 | CREQUSER | | 보기 | 보기 |
| 5 | CSITE | | 보기 | 보기 |
| 6 | DIRECT | | 보기 | 보기 |
| 7 | JGROUPS | | 보기 | 보기 |
| 8 | JUSERS | | 보기 | 보기 |
| 9 | JWORKLOG | | 보기 | 보기 |
| 10 | MUSIC_ALBUM | | 보기 | 보기 |
| 11 | MUSIC_ARTIST | | 보기 | 보기 |
| 12 | MUSIC_SONG | | 보기 | 보기 |
| 13 | black | | 보기 | 보기 |

[그림 3.8] 테이블 리스트

- 테이블 조회시 테이블명 없이 조회를 하게 되면 전체 테이블 리스트를 확인할 수있다.
- 각각의 테이블 명과 전체 건수 , 그리고 테이블타입을 확인 가능하다.
- 테이블 리스트의 **보기** 버튼을 이용하여 필드확인, 데이터 확인 기능을 사용할 수 있다.

- 필드 확인 : 테이블 스키마 확인

- 데이터 확인 : 테이블 전체 데이터 미리보기

| index | Field | Туре | Null | Key | Default | Extra | |
|-------|-------------|--------------|------|--------|---------|----------------|--|
| | c_num | int(11) | | PRI | | auto_increment | |
| | c_parent | int(11) | | MUL | 0 | | |
| | c_link | int(11) | | MUL | 0 | | |
| | c_name | varchar(50) | | | | | |
| | c_keyword | varchar(255) | YES | | | | |
| | c_comment | varchar(255) | YES | | | | |
| | c_autoclass | char(1) | YES | | | | |
| | c_depth | tinyint(4) | | MUL | 0 | | |
| | c_level | tinyint(4) | | | 0 | | |
| n | 1. | :_+/11\ | | Lat II | 0 | | |

[그림 3.9] 테이블 필드 확인

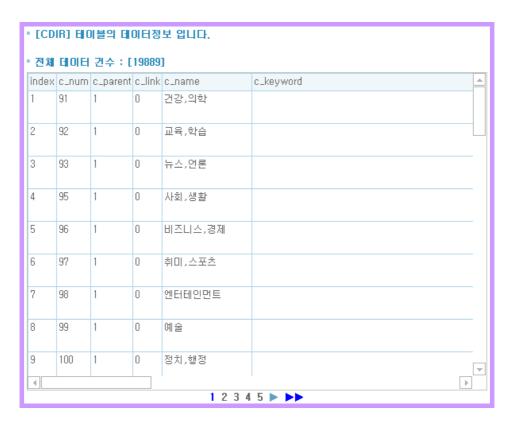
● ExodGate 필드정보 화면

- Field : 선택된 테이블의 필드명

- Type : 필드의 데이터 타입

- Null : 필드가 NULL값의 지원 유무

- Key : 필드의 인덱스 정보 - Default : 필드의 기본값 정보 - Extra : 필드의 기타 정보

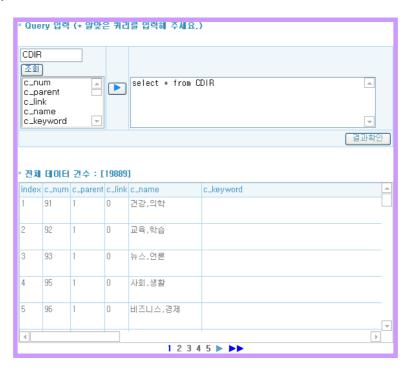


[그림 3.10] 테이블 데이터 확인

● ExodGate 데이터 확인 화면

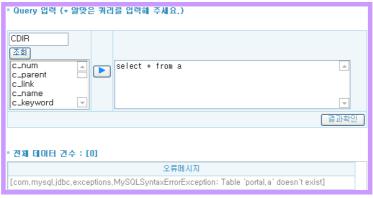
- 해당 테이블의 전체 데이터 건수 출력
- 전체 필드명과 필드 값 출력
- 페이지 네비게이션 기능을 이용하여 1~N 번의 필드 전체를 확인 가능

● 색인등록, 수집데이터 미리보기와 연동되어 데이터베이스의 데이터 정보가 아닌 수집할 데이터를 미리 확인 할 수 있다.



[그림 3.11] ExodGate 화면(쿼리 테스트)

- ExodGate 쿼리 테스트 화면
 - 필드 가져오기 : 테이블을 조회하여 쿼리문 작성을 돕기위한 선택된 테이블의 필드명 제공한다.
 - Query입력 : 테스트할 쿼리문을 입력한다. (아래 * 주의사항 * 참조)
 - 결과확인 버튼을 이용하여 입력 쿼리에 관한 결과화면을 출력한다.
 - 페이지 네비게이션 기능을 이용하여 1~N 번의 필드 전체를 확인 가능하다.
 - 잘못된 쿼리 입력시 오류메시지를 띄운다.



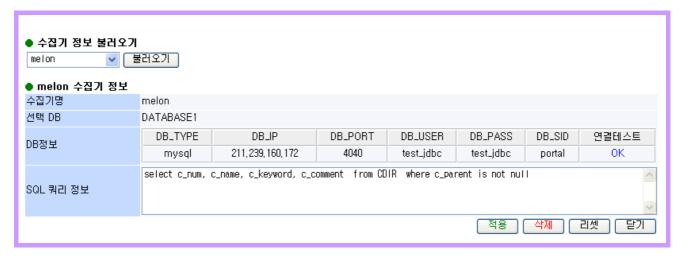
[그림 3.12] ExodGate 화면(쿼리 테스트 오류 메시지 출력)

* 주의 사항 *

- 너무 많은 데이터의 출력을 막기 위해 SELECT 쿼리에 한하여 내부적으로 결과 개수 제한 키워드를
 사용한다. (ORACLE: BETWEEN, MYSQL: LIMIT, MSSQL: TOP)
- 데이터베이스의 보안을 위하여 아래 키워드의 사용을 제한한다. (CREATE, DROP, DELETE, INSERT, UPDATE, ALTER)

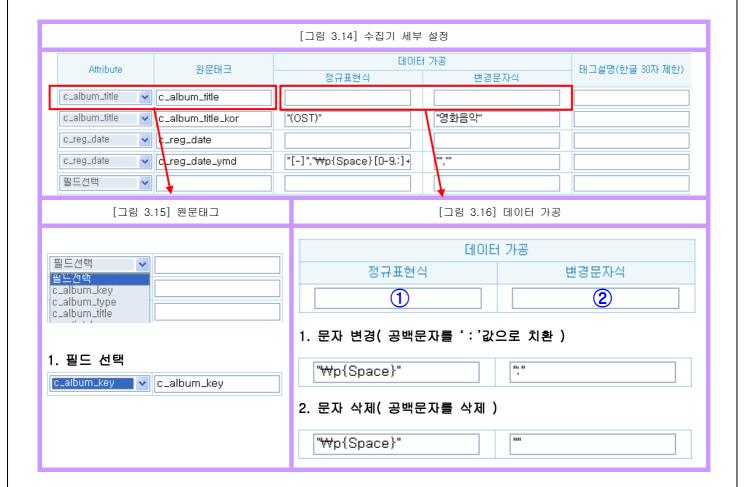
3.2.3. 수집기 관리.

- 등록된 수집기의 정보 확인과 관리를 할 수 있는 기능으로 세부적인 수집 설정을 입력할수 있어 실제 생성될 수집파일의 원문 태그설정, 수집데이터의 필드별 가공등의 핵심적인 작업을 수행할 수 있다.
- 수집 SQL 쿼리에 대한 수정과 해당 수집기의 삭제를 할수 있다.
- 해당 수집기의 인증 정보를 확인하여 실제 색인에 반영할수 있는지에 대한 유무를 판단한다. (수집데이터 미리보기, 인증현황에서의 결과 값을 확인)



[그림 3.13] 수집기 관리 화면1

- 수집기 관리 화면 1
 - 해당 수집기의 데이터베이스 정보와 SQL 수집쿼리 정보 확인이 가능하다.
 - 수집쿼리를 변경하여 적용 버튼을 눌러서 저장할 수 있으며 삭제 버튼을 이용하여 수집기를 삭제할 수 있다.



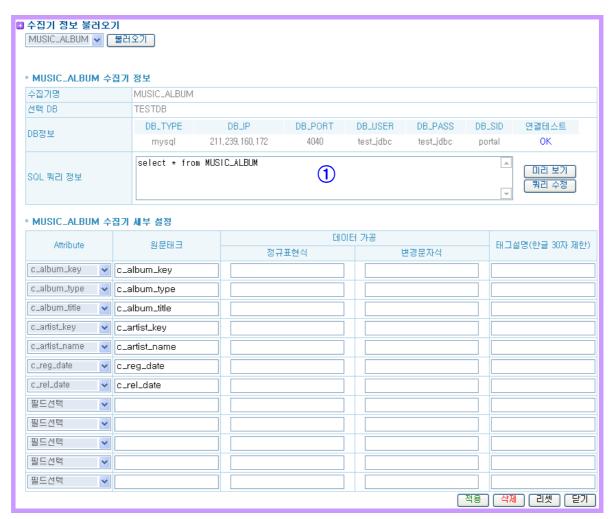
● 수집기 세부 설정 화면

- 수집할 데이터에 대한 원문태그설정, 데이터 가공, 태그설명을 설정할 수 있다.
- 원문태그: Attribute 필드는 수집쿼리의 결과에서 받아오는 수집 필드명을 화면상에 출력 한 것이다. 필드를 선택하게 되면 상단의 그림과 같이 원문태그 필드로 Attribute값이 전달 되는데 사용자가 원 문태그를 임의로 입력해도 되지만 실제 파일에는 ^원문태그: 의 형식으로 저장된다.
- 데이터 가공: 정규표현식을 이용하여 해당 필드에 대한 데이터 가공이 가능하다. 총 100개까지의 데이터 가공식을 입력할 수 있으며, 하나의 식의 구분자로는 "와,가 사용된다. 만일 두박스에 입력된 구분자의 수가 틀리게 될경우에는 적은 쪽의 갯수 만큼만 변경되기 때문에 반드시 구분자를 확인하여야 한다. 두개의 에디트 박스 중에서 ①박스에는 정규표현식(JAVA)을 입력하고, ②박스에는 변경을 원하는 문자를 입력한다.

◆ 문자 변경 : ①박스에 정규표현식을, ②박스에 변경할 값을 입력한다.

◆ 문자 삭제 : ①박스에 정규표현식을, ②박스는 비워둔다.

- 태그설명 : 한글 30자 이내에서 원문태그 설정에 대한 설명을 등록할 수 있다.
- 등록 : 등록 버튼을 누르면 수집기에 세부설정 사항이 등록되고 재설정이나 수집데이터 미리보기 기 능이 가능한 페이지로 이동한다.



[그림 3.17] 수집기 관리 화면 2

● 수집기 관리 화면 2

- SQL쿼리 정보 : 설정한 수집 쿼리에 대한 확인과 해당 쿼리의 데이터 미리보기, 쿼리를 수정 할 수 있는 쿼리수정이 가능하다. ①(쿼리 정보 입력 박스)에서 쿼리를 수정한후 미리보기 버튼을 클릭하게 되면 수정된 쿼리에 관해서 미리보기가 가능하므로 수정한 쿼리를 저장하기 전에 반드시 미리보기를 이용하여 먼저 데이터를 확인하는 작업이 필요하다.
- 수집기 세부설정 : 수집기 관리화면 1과는 두가지의 차이가 있다. 첫번째는 원문태그변경을 위한 Attribute 필드의 수가 늘어 났다는 것이다. 이는 동일한 필드에서 다양한 형태의 데이터를 가공하여 각가 다른 원문태그로 저장하기 위한 용도에서 꼭 필요한 부분이다. 그리고 두번째는 수집기의 삭제와 수정이 가능하다.
- 수집기 인증정보: 등록된 수집기가 정상적으로 데이터를 수집할 수 있는지의 여부와 수집데이터를 미리볼수 있는 서비스이다. 수집데이터 미리보기 기능은 이제까지의 미리보기 기능과는 또 다른 것으로 실제 수집할 SCD의 내용을 보는 것과 마찮가지의 기능이라 할 수 있겠다. 파일로 수집을 하기 전에 설정에서 등록된 원문태그나 수집데이터의 가공값을 반영한 결과를 확인 할 수 있다.



[그림 3.18] 수집 데이터 미리보기 화면

4. INDEX

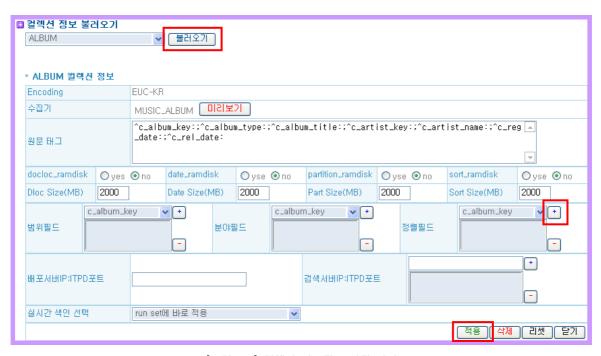
- 수집기를 선택하거나 정의된 SCD 구조에 맞게하여 컬렉션을 생성 (구성설정) (수집기를 사용하지 않을 경우 사용자가 exod/+INDEXER/crl/컬렉션.crl.sh 스크립트를 사용자가 사용 하는 수집기를 사용하도록 수정해서 사용해야 한다.)
 - 날짜필드, 분야필드, 정렬필드 설정
 - 색인 및 검색서버 설정
 - 실시간 색인 방법 설정
- 필요색인을 설정 (구성설정)
 - 색인방식, 색인옵션, 중요도
 - 기타 색인옵션(인접도, 절단검색, 대소분자 구분)
- 배치/실시간 색인 실행 및 색인 결과 확인 (색인 스케쥴러, 색인 결과조회)
 - 배치/실시간 색인 실행 등록 및 결과 확인
- 공유메모리 (공유메모리)
 - 현재 서비스되고 있는 SET 확인
 - 공유메모리 올리기, 내리기
- 실시간 색인 삽입/삭제 (실시간 색인)
 - 실시간으로 현재 서비스되고 있는 SET 이나 Online SET에 문서 삽입/삭제
- 색인 내용 확인 (색인 체크)
 - 색인에 등록된 Term과 Term이 등록된 색인을 컬렉션 별로 확인

4.1. Collection



[그림 4.1] 컬렉션 등록 화면

- 수집기를 선택하고 컬렉션 이름을 입력하고 [등록] 버튼을 누른다.
- 수집기를 사용하지 않을 경우 ☑ 버튼을 누르고 컬렉션 이름을 입력하고 사용자가 SCD에 사용할 원 문태그를 입력한다.
- 등록된 컬렉션을 선택후 ^{불러오기} 버튼을 눌러서 기본 정보를 설정한다.



[그림 4.2] 컬렉션 기본정보 설정 화면

- docloc_ramdisk, date_ramdisk, partition_ramdisk, sort_ramdisk 기능 설정 및 색인볼륨사이즈 설정 (XTEN3.5 이상)
 - InMemory 기능을 사용할 경우 yes로 설정하고 각 볼륨의 사이즈를 적어준다.
 - 사이즈 설정시 볼륨의 크기와 온라인 삽입시 늘어날 크기를 계산해서 적어준다.

- 범위필드, 분야필드, 정렬필드 설정
 - 범위필드 : 문서의 번호나 날짜등 숫자로된 필드가 있을 때 설정해 줄 수 있다. 이 필드를 설정해 놓으면 색인시 DATE 볼륨을 생성하여 검색결과에서 범위와 정렬방법을 설정할 수 있다.
 - 분야필드: category를 나눌 때 사용한다. 필드가 분야를 구분할 수 있는 값으로 설정되어 있을 때 사용할 수 있다. 이 필드를 설정해 놓으면 색인시 PART 볼륨을 생성하여 검색결과에서 category 와 정렬방법을 설정할 수 있다.
 - 정렬필드 : 제목이나 본문내용을 가지고 있는 필드로 설정해 줄 수 있다. 이 필드를 설정해 놓으면 색인시 SORT 볼륨을 생성하여 검색결과에서 그 필드에 대한 가나다정렬을 설정할 수 있다.

- 색인서버 및 검색서버 설정

- 색인서버 : 실제로 색인을 실행하는 server다. 이 server에서 색인을 돌린 후 검색 server에 배포를 하게 된다. 색인 server의 IP를 적은 후 :와 itpd port번호를 적어 준다.

- 실시간 색인 설정

- run set에 바로 적용 : 실제 service되고 있는 색인에 추가되는 문서를 삽입/삭제하는 방식이다.
- online set을 만들어서 적용 : 실제 service되고 있는 색인에는 영향을 미치지 않고 setonline을 만들어서 사용하는 방법이다. 검색결과가 service되고 있는 set과 setonline이 merge 된다.
- 수집기 호출 스크립트 생성 : ^{적용} 을 누르면 위에서 설정된 내용을 포함한 파일이 /+INDEXER/inf/컬렉션.inf 로 생성되고, 수집과 색인 및 실시간 색인에 필요한 shell script가 생성된다.

● 하나의 컬렉션에 최대 64개의 색인 등록이 가능



[그림 4.3] 색인 등록 화면

- 색인명 : 색인명을 입력한다. 하나의 컬렉션에 대해서 64개의 색인을 생성할 수 있다.
- 필드명 : 색인할 필드를 select box에서 선택해준다.
- 형태소분석방식 : 선택한 필드에 대한 다양한 형태소분석방식을 선택가능하다.
- 형태소분석옵션 : 숫자와 문자등을 분리 색인가능하다.
 (설정시 반드시 xten manual을 참고해야 한다. 형태소분석 방식에 따라서 적용되는 옵션이 틀리기

■ +digit : 숫자를 색인어로 추출한다.

때문이다.)

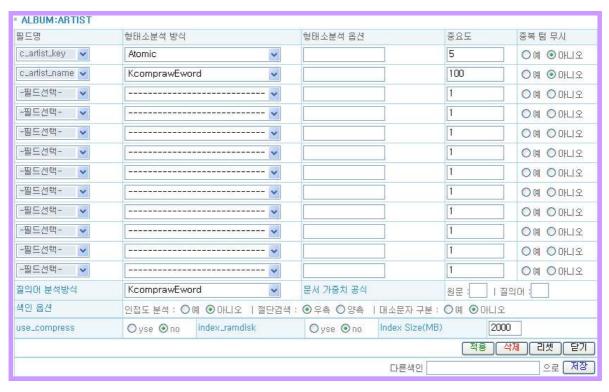
■ +josa: 조사를 색인어로 추출한다.

■ +word : 어절을 색인어로 추출한다.

- 중요도 : 검색시 이 필드에 가중치를 부여할 수 있다.
- 중복텀 무시 : 검색어가 이 필드에 여러 번 나왔을 경우 중요도(가중치)를 한번만 적용할 것인지 나오는 횟수대로 적용할 것인지를 설정하는 부분이다.
- 질의어 분석방식 : 형태소분석 방식과 질의어 분석방식은 동일하길 권장한다. 검색어가 분석되는 결과와 색인한 결과가 다르면 원하지않은 검색결과가 나올수 있기 때문이다. 입력한 검색어를 분석하는 방식을 설정하는 부분이다.
- 가중치공식 : 가중치를 계산하는 방법을 설정하는 부분이다.
- 인접도 분석 : 여러 질의어를 검색어로 검색할 경우 질의어들의 거리를 계산해서 가중치를 주는지를 설정하는 부분이다.

- 절단검색 : 기본값은 우측 절단이다.
 - 우측 : 검색시 우측 절단을 수행할 수 있도록 한다.
 - 양측 : 검색시 좌측, 우측, 양측 절단을 모두 수행할 수 있도록 한다.
 - 일반적인 정보서비스에서는 좌측 절단 또는 양측 절단이 필요하지 않는 경우가 대부분이다. 또한 "양측" 옵션을 사용하면 좌측절단을 지원하기 위해 더 많은 정보를 색인에 저장해야 하므로 색인을 저장해야 하는 디스크 오버해드가 발생한다. 따라서 특별히 좌측 절단 또는 양측 절단 연산이 필요하지 않다고 판단되면, "양측" 옵션을 사용하지 말 것을 권장한다.
- 대소문자 구분 : 대소문자를 구분해서 검색할지를 설정하는 부분이다.
- user_compress(XTEN3.5 이상): 색인 압축 기능을 사용할지 여부를 설정한다. (색인 사이즈가 클 경우 권장)
- index_ramdisk 기능 설정 및 색인볼륨사이즈 설정 (Index) (XTEN3.5 이상)
 - InMemory 기능을 사용할 경우 yes로 설정하고 각 볼륨의 사이즈를 적어준다.
 - 사이즈 설정시 볼륨의 크기와 온라인 삽입시 늘어날 크기를 계산해서 적어준다.

★ 자세한 내용은 XTEN 색인관리자 Manual을 참고



[그림 4.4] 색인 등록 확인 화면

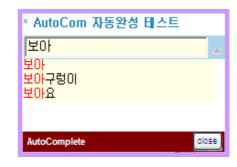
4.2. Collection(자동완성/동의어)

- 사전을 선택하고 컬렉션 이름을 입력하고 등록 버튼을 누른다.
- 등록된 컬렉션을 선택후 ^{불러오기} 버튼을 눌러서 기본 정보를 설정한다.
- 나머지 설정방법은 4.1의 방법과 동일하고 자동완성/동의어 컬렉션은 색인에 대한 기본적인 설정들이 되어있다.
- **자동완성 기능**은 색인생성 후

http://192.160.1.1:7578(exod서버주소:포트)/autocom?colname=AutoCom(컬렉션명)&keyword=보아(쿼리)

위의 URL을 자바스크립트 등에서 호출하게 되면





[그림 4.5] autocom

[그림 4.6] EXOD 테스트화면

보아|보아구렁이|보아요

보아|보아구렁이|보아요

를 받을 수 있다. 이기능을 이용하여 자동완성 기능을 자바스크립트 등에서 파싱해서, EXOD의 테스트 페이지와 같이 AJAX를 이용해서 사용하면 된다.

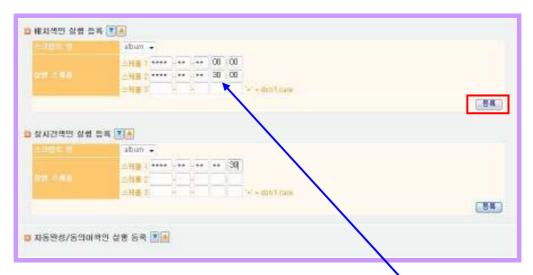
● **동의어 기능**은 색인생성 후 **검색설정**에서 **질의어 확장 컬렉션 메뉴에서 동의어 컬렉션으로 설정**해주면 검색시 동의어 색인에서 키워드를 확장해서 검색하게 된다. (단, 질의어 확장 컬렉션 사용시 동의어 컬 렉션에 검색을 한번 한 후에 질의어를 확장해서 검색을 하기 때문에 오버해드가 발생한다.)



[그림 4.7] 동의어 컬렉션 등록화면

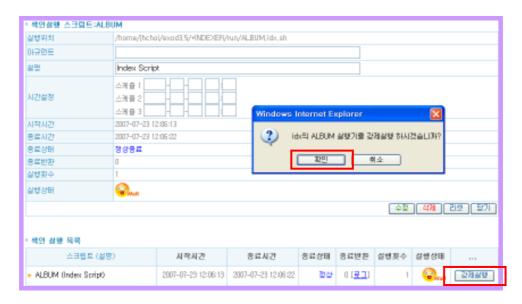
4.3. IndexScheduler

● 색인설정을 한 컬렉션에 대해 배치/실시간 색인을 설정



[그림 4.8] 스케쥴러 등록 화면

- 실행할 색인을 등록한다. 색인 실행단위는 컬렉션이다.
- 스크립트명 : 색인할 컬렉션을 선택한다. 컬렉션을 선택하면 컬렉션에 포함된 모든 색인을 실행한다.
- 실행 스케쥴: 한 컬렉션에 3개의 시간설정이 가능하다. 스케쥴 순서대로 **년-월-일-시-분:초** 순으로 입력한다. 일주일 단위로 색인을 돌리고 싶면 ****-**-MO 식으로 입력한다. 일주일 단위로 요일마 다 MO, TU, WE, TH, FR, SA, SU을 지원한다. 위의 예는 매시간마다 30분과 00분에 색인을 실행한다. 즉 30분마다 새로운 색인을 한다.
- 등록 을 클릭하면 설정된 실행정보의 확인 및 수정이 가능하다.



[그림 4.9] 스케쥴러 강제실행 화면

- 강제실행 : 실행 스케쥴에서 설정한 시간외에 ^{강제실행} 버튼을 눌러서 강제로 색인을 실행할 수 있다. 버튼을 누르면 확인창이 나오고 ^{확인}을 누르면 색인을 실행한다.

| ■ 색인 실행 목록 | | | | | | | | | |
|----------------------------|---------------------|---------------------|------|---------------|------|------|------|--|--|
| 스크립트 (설명) | 시작시간 | 종료시간 | 종료상태 | 종료반환 | 실행횟수 | 실행상태 | | | |
| ► ALBUM (Index Script) | 2007-07-23 18:02:53 | | | [<u>로그</u>] | 2 | Run | 강제종료 | | |
| - 자동완성/동의머 실행 목록 | | | | | | | | | |
| 스크립트 (설명) | 시작시간 | 종료시간 | 종료상태 | 종료반환 | 실행횟수 | 실행상태 | | | |
| ▶ aaaa (No-Description) | 2007-07-23 12:09:04 | 2007-07-23 12:09:04 | 정상 | 0 | 1 | Wait | 강제실행 | | |
| ► AutoCom (No-Description) | 2007-07-23 17:42:57 | 2007-07-23 17:42:57 | 정상 | 0 | 1 | Wait | 강제실행 | | |

[그림 4.10] 스케쥴러 실행중 화면

- 시작시간 : 색인이 시작된 시간 - 종료시간 : 색인이 종료된 시간

종료상태 : 색인의 종료상태를 표시한다.실행횟수 : 색인이 실행된 횟수를 표시한다.

- 실행상태 : 색인이 실행되고 있는 상태를 표시한다.

| 실행상태 | 설명 | | | | | | |
|-------|--------------------------------|--|--|--|--|--|--|
| READY | 색인이 실행될 시간이 되었을 때의 상태."준비"로 표시 | | | | | | |
| RUN | 색인이 실행중일 때의 상태."실행"으로 표시 | | | | | | |
| CLOSE | 색인이 실행중이지 않을 때의 상태."대기"로 표시 | | | | | | |

[표 4.1] 스케쥴러 실행상태

- 강제종료: 색인이 실행중일 때 ^{강제종료} 버튼을 눌러서 색인을 강제로 종료할 수 있다. 강제종료가 되었어도 색인을 실행시키는 스크립트의 process는 종료되지만 스크립트에서 파생된 자식 process 들은 종료되지 않는다.
- 색인설정을 한 자동완성/동의어 컬렉션에 대해 자동실행을 설정
 - 설정방밥은 위와 동일하다.

4.4. LogInquire

● 색인 과정별 각 프로세스의 진행상태 확인할 수 있다.



[그림 4.11] 로그조회 화면

- batch script: 색인을 실행하는 script다
- assign index : 색인을 돌릴 set을 정하는 부분이다. exod는 기본적으로 2개의 set을 가지고 있다.
- unload idxmem : memory에 올린 공유메모리를 내린다.
- final shm : 색인을 돌릴 set의 공유메모리를 내린다.
- init dir : 색인 처음 실행되는 경우 set에 색인이 실행될 컬렉션에 대한 폴더를 생성
- crawl data : 수집기로 scd를 생성한다.
- make loc : 색인에 대한 컬렉션.loc 파일을 생성한다.
- create locvol: 색인에 대한 locvol을 생성한다.
- load locvol: locvol을 load한다.
- create date : date volume을 생성한다.
- load date : date volume을 load한다.
- create part : part volume을 생성한다.
- load part : part volume을 load한다.
- create sort : sort volume을 생성한다.
- load sort : sort volume을 load한다.
- create index : index volume을 생성한다.
- build index: index volume을 build한다.
- memvol to disk: 메모리에 생성한 색인을 파일로 만든다.
- init shm: index volume을 load한다.

- clear rindex : setoline의 index volume을 공유메모리에서 내리고, setoline폴더를 삭제하고 실시간 색 인이 실행중인 실시간 프로세스들을 죽인다.
- serv idx : service 되고 있는 set을 지금 배치 색인이 끝난 set으로 바꾼다.

4.5. SharedMemory

● 검색 속도 향상을 위해 색인데이터를 메모리에 로딩하는 역할



[그림 4.12] 공유메모리 관리 화면

- 기본적으로 스케쥴러에 의해서 색인이 실행되면 공유메모리를 올린다.
- 서비스되고 있는 공유메모리를 확인한 뒤 강제로 조정가능

4.6. RealTimeIndex

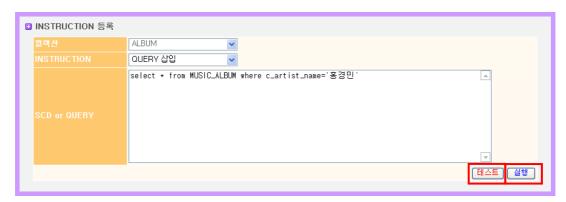
● 실시간으로 색인에 문서를 삽입/삭제 지원



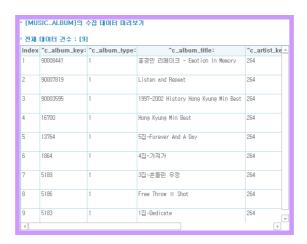
[그림 4.13] 실시간 색인 화면

- 해당 컬렉션을 선택
- INSTRUCTION을 선택 (INSTRUCTION에 대한 자세한 내용은 8. REAL TIME INDEX 참조)
- SCD or QUERY에 INSTRUCTION에 해당 하는 문서의 내용을 입력하고 ^{테스트} 버튼으로 결과를 확 인하고 ^{실행} 버튼으로 INSTRUCTON을 수행

● QUERY 삽입



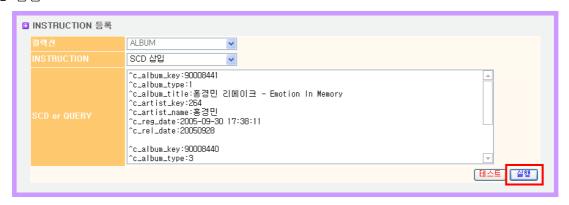
[그림 4.14] QUERY 삽입 화면



[그림 4.15] QUERY 삽입 테스트 화면

- SCD or QUERY에 삽입할 문서를 가져올 수 있는 SQL을 입력하고 데스트 버튼으로 결과를 확인하고 일행 버튼으로 INSTRUCTON을 수행

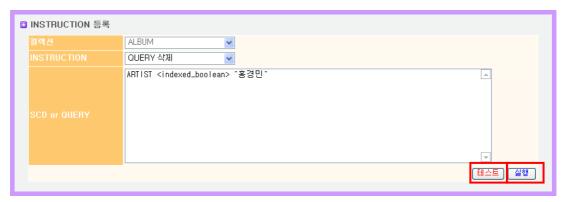
● SCD 삽입



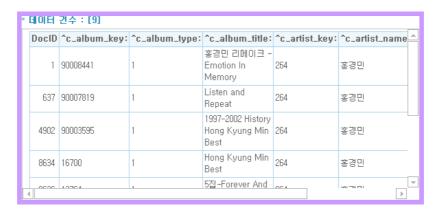
[그림 4.16] SCD 삽입 화면

- SCD or QUERY에 삽입할 문서(SCD)를 입력하고 실행 버튼으로 INSTRUCTON을 수행 (테스트 페이지는 제공하지 않는다.)

● QUERY 삭제



[그림 4.17] QUERY 삭제 화면



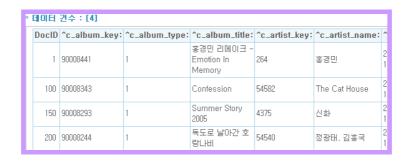
[그림 4.18] QUERY 삭제 테스트 화면

- SCD or QUERY에 삭제할 문서를 검색할 수 있는 질의어를 포함한 쿼리(색인명 + 질의방식 + 질의어)를 입력하고 데스트 버튼으로 결과를 확인하고 집행 버튼으로 INSTRUCTON을 수행

● DOC ID 삭제 RUN SET



[그림 4.19] DOC ID 삭제 RUN SET 화면

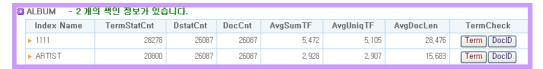


[그림 4.20] DOC ID 삭제 RUN SET 테스트 화면

- SCD or QUERY에 삭제할 문서의 DOC ID를 입력하고 테스트 버튼으로 결과를 확인하고 실행 버튼으로 INSTRUCTON을 수행
- DOC ID 삭제 ONLINE SET
 - 컬렉션 설정에서 실시간 색인 선택에서 online set을 만들어서 적용을 선택하였을 경우 사용한다.
 - 사용법은 DOC ID 삭제 RUN SET과 동일하다.

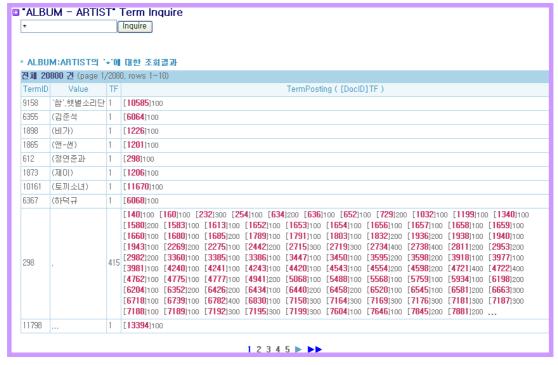
4.7. IndexCheck

● 컬렉션의 색인 별로 Term에 대한 기본정보를 확인할 수 있다.



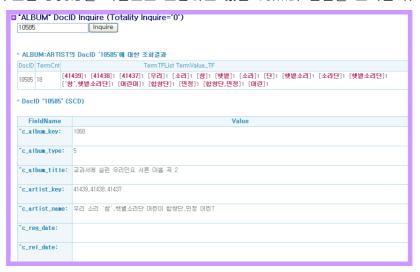
[그림 4.21] 색인 확인 메인 화면

● Term 버튼을 누르면 Term을 기준으로 Term을 포함하고 있는 DocID를 보여준다.



[그림 4.22] 색인 확인 Term 기준

● DocID 버튼을 누르면 DocID를 기준으로 포함하고 있는 Term과 원문을 보여준다.



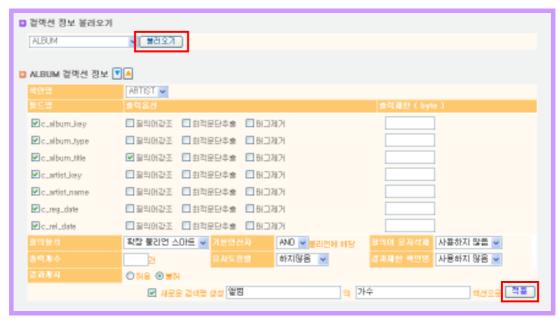
[그림 4.23] 색인 확인 DocID 기준

5. SFARCH

- 정의된 색인 컬렉션을 통해 검색 설정
 - 출력 옵션(질의어 강조, 최적문단)
 - 질의형식, 출력개수
- 검색 결과 확인 : TXT, XML, XLS, HTML 형태를 통한 검색결과
- 검색결과 로그 통계 : 검색결과에 대한 로그를 피크타임/인기검색어로 통계
- 검색결과 로그 통계 스케쥴러 : 검색결과 로그 통계를 주기적으로 수행

5.1. SearchSetup

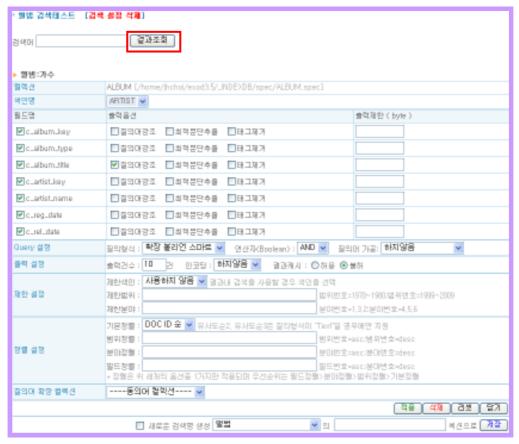
● 구성설정은 컬렉션 구성정보를 토대로 만들어진다. 컬렉션 구성에서 설정해 놓은 필드들을 토대로 검 색을 할 수 있다.



[그림 5.1] 검색 등록 화면

- 컬렉션 선택후 **불러오기** 버튼을 누르면 해당 컬렉션 정보를 볼 수 있다.
- 색인명 : 검색에 사용하고자 하는 색인을 선택한다.
- 필드명 : 검색시 출력하고자 하는 필드를 선택한다.
- 출력옵션
 - 질의어 강조 : 해당 필드에 질의어가 있으면 굵은 글씨로 출력한다.
 - 최적문단추출 : 검색결과 출력시 질의어가 찾아진 부분을 출력한다.
 - 태그제거 : 검색결과에서 html 태그를 제거한다.
- 출력제한 : 해당필드가 검색결과로 나올경우 몇 byte까지 출력할지를 설정할 수 있다.
- 질의형식
 - 질의어를 분석하는 형식이다.
 - Boolean, Boolean Pnorm, Boolean Smart, Text방식을 제공한다.

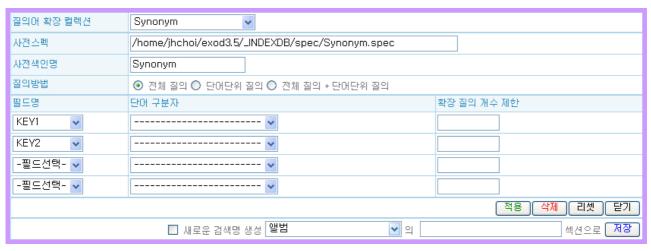
- 기본연산자 : 질의어에 띄어쓰기를 사용해서 입력했을 때 질의어를 AND 또는 OR로 처리할지를 설정
- 질의어 문자삭제 : 질의어에서 기호문자, 공백문자, 기호와 공백문자를 삭제하고 싶을 때 설정한다.
- 출력개수 : 검색시 출력되는 결과의 개수를 설정한다.
- 유사도정렬 : 유사도에 따라 검색결과를 정렬한다.
 - DOC ID 순: 유사도를 반영하지 않고 **DOC ID 순**으로 검색결과를 정렬한다.
 - 유사도순 : 검색결과를 유사도 순으로 정렬한다.
- 결과제한 색인명 : 검색결과를 제한할 색인을 선택한다.
- 결과캐시 : 검색결과를 캐시에 올려놓을지를 설정한다.
- 검색명을 select box에서 선택하거나, 새로운 검색명을 생성하려면 check box를 체크하고 검색명을 입력 후 섹션명을 입력하고 적용 버튼을 누르면 설정된 내용들이 저장된다.
- [그림 5.1]에서 ^{적용} 버튼을 누르면 좀 더 세부적인 검색결과제한, 검색결과정렬, 질의어 확장에 대한 설정을 할 수 있다.



[그림 5.2] 검색 섹션 등록 화면

- 결과조회 버튼을 통해서 검색결과를 확인 할 수 있다
- 제한 설정
 - 제한색인 : 제한을 적용할 색인명을 사용한다. (결과내 검색에 사용된다.)
 - 제한범위: 색인설정에서 범위필드를 설정해 놓았으면 설정해 놓은 순서대로 **첫번째 필드의 범위는** 2006년 1월 1일부터 2007년 1월 1일까지 이고 두번째 필드의 범위를 1900년부터 2007년까지로 제한 하려면 1=20060101~20070101;2=1900~2007 식으로 설정할 수 있다.
 - 제한분야 : 색인설정에서 분야필드를 설정해 놓았으면 설정해 놓은 순서대로 첫번째 분야가 1,2,3이고 두번째 분야가 4,5,6인 것으로 제한하려면 1=1,3,2;2=4,5,6 식으로 설정할수 있다.

- 정렬 설정 : 검색결과를 정렬하는 방식을 설정할 수 있다.
 - 우선순위는 **기본정렬 < 범위정렬 < 분야정렬 < 필드정렬**이다. (우선순위에 따라 1가지만 적용)
 - 기본정렬: 기본으로 DOC ID순으로 정렬하고, 유사도에 따라서 유사도순1, 유사도순2, 유사도순3으로 설정할 수 있다.
 - 범위정렬: 색인설정에서 범위필드를 설정해 놓았으면 설정해 놓은 순서대로 첫번째 필드를 오름차순으로 두번째 필드를 내림차순으로 정렬하려면 1=asc;2=desc식으로 설정할 수 있다.
 - 분야정렬: 색인설정에서 분야필드를 설정해 놓았으면 설정해 놓은 순서대로 첫번째 필드를 오름차순으로 두번째 필드를 내림차순으로 정렬하려면 1=asc;2=desc식으로 설정할 수 있다.
 - 필드정렬: 색인설정에서 정렬필드를 설정해 놓았으면 설정해 놓은 순서대로 첫번째 필드를 오름차순으로 두번째 필드를 내림차순으로 정렬하려면 1=asc;2=desc식으로 설정할 수 있다.
 - 범위정렬과 분야정렬은 숫자를 기반으로한 정렬이고. 필드정렬은 문자를 기반으로한 정렬이다.

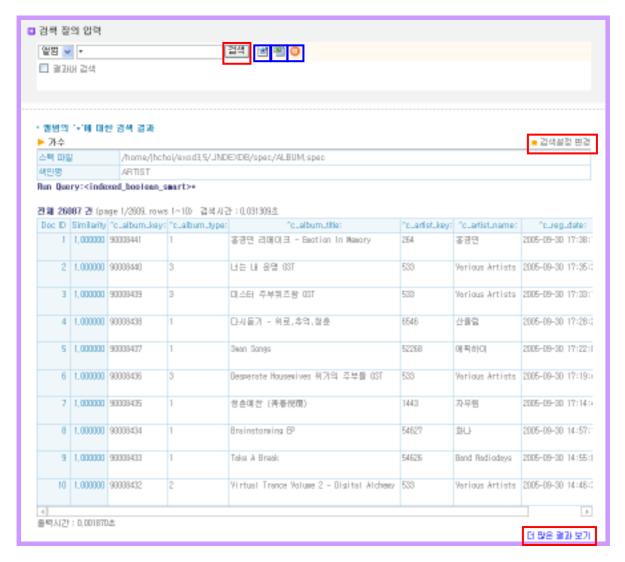


[그림 5.3] 질의어 확장 화면

- 질의어 확장: 설정된 색인 이외에 다른 색인에서 원래 질의어와 연관된 질의어들을 확장해서 검색하고 싶을 경우 사용한다. 동의어와 유사어등을 검색결과에 반영하기 위해서 사용한다. 질의어 확장기능을 쓰면 질의어를 확장하기 위한 검색을 한번 더 수행한다. **사전스펙, 사전색인명, 질의방법**은 질의어를 검색할 색인에 대한 정보이고 **필드부분**은 질의어를 확장할 필드에 대한 설정을 하는 부분이다.
 - 질의어 확장 컬렉션 : 컬렉션 설정에서 동의어 컬렉션으로 등록한 컬렉션을 선택하고 ^{적용} 버튼을 누르면 세부설정을 할 수 있다.
 - 사전스펙 : 질의어를 확장할 컬렉션의 USER SPEC 경로를 써준다.
 - 사전색인명 : 검색대상 색인 명을 입력한다.
 - 질의방법
 - 단어단위 질의 : 질의어에서 각각 단어별로 질의어를 확장한다.
 - 전체 질의 : 질의어를 하나로 보고 질의어 전체에 대해서 확장한다.
 - 필드명 : 질의어를 **확장할 필드명**을 입력한다.
 - 단어구분자 : 필드에 확장될 질의어의 구분자를 설정한다. **공백으로 구분, 콤마로 구분, 기호로 구분, 공백+기호로 구분**이 있다.
 - 확장 질의 개수 제한 : 질의어 확장시 **확장된 질의어**가 많으면 검색시간이 오래걸리기 때문에 **개수** 를 **제한** 할 수 있다.

5.2. SearchInquire

● TXT 출력



[그림 5.4] TXT 출력 화면

- EXOD에서 검색어를 입력하고 검색 버튼을 눌렀을 때 검색설정에 따라 Doc ID(SCD문서번호)와 Similarity(유사도)를 포함해서 결과를 보여준다.
- 검색결과를 변경하고 싶으면 버튼을 눌러서 다시 설정할 수 있다.
- 검색결과를 세부적으로 조회하고 싶으면 <mark>더 많은 결과 보기</mark> 를 눌러서 해당 섹션의 검색결과를 확인할 수 있다.

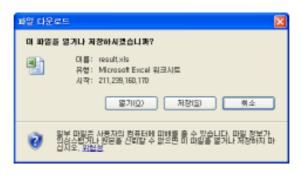
● XML 출력

```
weta_storage_list date='2007/01/05 16:44:31">
- <section name="가수">
<usrspec>/obuser/exodtft/exod.xxx/_INDEXDB/spec/ALBUM.spec</usrspec>
    <idxname>ARTIST</idxname
     <qryflag>boolean</qryflag>
    <srtflag>0</srtflag>
    <totcnt>26087</totcnt>
    <maxcnt>20</maxcnt>
<outcnt>20</outcnt>
     <pagenum>1</pagenum>
    cpagenum>1
celapse>0.009587sec</elapse>
<doc id="1" sim="1.000000">
- <att name="^c_album_key:">
<![CDATA[ 90008441 ]]>
       </att>
     - <att name="^c_album_type:">
          <![CDATA[ 1 ]]>
       <att name="^c_album_title:">
       <![CDATA[ 홍경민 리메이크 - Emotion In Memory ]]>
</att>
       <att name="^c_artist_key:">
       <![CDATA[ 264 ]]>
</att>
     - <att name="^c_artist_name:">
          <![CDATA[홍경민 ]]:
       </att>
     - <att name="^c_reg_date:">
  <![CDATA[ 2005-09-30 17:38:11 ]]>
       </att>
       <att name="^c rel date:">
         <![CDATA[ 20050928
       </att>
     </doc>
```

[그림 5.5] xml 출력 화면

- 검색결과를 xml형식으로 얻고 싶으면 질의어를 입력하고 [□] 버튼을 누르면 검색결과를 xml형식으로 새로운 창에서 보여준다.
- base64 encoding은 기본으로 Y로 설정되어 있다. 위의 화면은 base65 encoding을 base64=N로 설정했을때의 화면이다.

● XLS 출력



[그림 5.6] xls저장 확인 화면

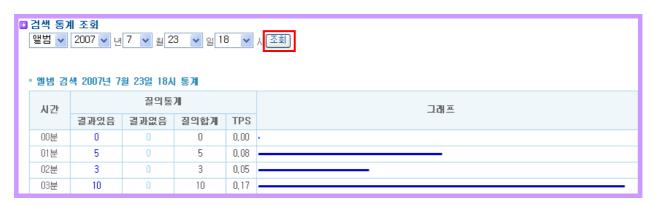
- 검색결과를 xls파일로 저장하고 싶으면 질의어를 입력하고 ^웹 버튼을 누르면 검색결과를 xls형식으로 저장할 수 있다.

● 템플릿 출력

- 검색결과를 사용자가 작성한 템플릿을 이용해서 보고 싶으면 질의어를 입력하고 [□] 버튼을 누르면 검색결과를 템플릿을 적용해서 볼 수 있다.
- 자세한 내용은 6. 템플릿에 나와있다.

5.3. UsageInquire

● 질의어의 검색 통계에 대해 년/월/일/시/분 까지 조회가 가능하다.



[그림 5.7] 검색 로그조회 화면

- 조회하고 싶은 검색명을 선택후 년/월/일/시/분을 선택 후 조회 버튼을 누르면 날짜별, 시간별 통계를 확인 할 수 있다. 좀 더 세부적인 통계를 원할 경우에는 해당 날짜/시간을 클릭하면 분단위까지 통계조회가 가능하다. 검색통계정보는 exod/+SRCHINF/qry/컬렉션/ 폴더에 날짜별로 저장된다.

5.4. 템플릿

5.4.1. 템플릿 불러오기

- 템플릿을 입력하기 위한 폼을 불러오거나 미리 만들어진 템플릿을 통해서 검색 테스트를 할수 있다. 템플릿 정보는 검색 설정이 되어있는 컬렉션과 연계하기 때문에 템플릿을 만들기 위해서는 검색 설정을 등록해야 한다.
- 검색어 박스에 원하는 질의어를 넣고 결과조회를 누르면 만들어진 템플릿을 통해서 원하는 결과가 출력된다. 미리 등록된 템플릿 페이지가 없을 시에는 아무결과도 조회되지 않는다.



[그림 6.1] 템플릿 불러오기 화면

5.4.2. 템플릿 변수 리스트

● 템플릿은 사용자에 의해서 다양한 플렛폼으로 만들어지지만 검색결과에대한 값들은 사용자들이 정의할 수 없다. 그래서 EXOD에서는 템플릿 변수를 직접 제공한다. 변수는 네가지로 나누어진다. 전체 폼에서 언제든지 사용가능한 공통변수, 하나의 섹션에서만 사용가능한 섹션변수, 검색결과 폼에서만 사용가능한 BODY변수, 마지막으로 다양한 검색 결과를 제공해주는 폼변수가 있다. Exod에서 제공하는 템플릿 변수는 반드시 정해진 형식과 정해진 페이지에서만 사용할수 있으므로 변수명을 사용자가 설정할수 없다.



[그림 6.2] 템플릿 변수리스트 화면

● 템플릿 변수리스트

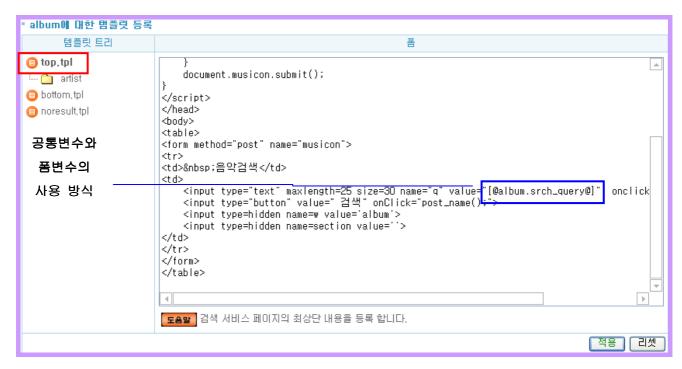
- 공통변수 : 모든 템플릿 페이지에서 공통적으로 사용할 수 있는 변수로써 사용자에게 입력 받는 검 색 질의어가 포함된다.
- 섹션변수: Exod의 검색 설정에서는 하나의 컬렉션안에 여러가지의 섹션으로 설정이 가능하다. [그림 3.15]에서는 cdir이라는 섹션이 등록 되어있고, cdir섹션의 모든 템플릿 페이지에서는 검색건수, 검색 시간, 출력 시간의 세가지 변수들을 사용할수 있다.
- **body변수**: 섹션의 body.tpl 페이지에서만 사용할 수 있는 변수가 body변수이다. 검색 결과 출력시 내부적으로 지원하는 문서번호와 검색 유사도를 기본으로 지원하며 나머지 변수들은 각 필드의 결과 값을 출력하게 된다.
- **품변수** : CGI, JSP 등의 인터페이스 프로그램을 이용하여 html 동적 페이지를 만들 때 사용자의 input을 output으로 변경해주는 폼안의 변수들을 폼변수라고 하는데 Exod도 이런 폼변수들을 지향하고 있기 때문에 스크립트안의 폼변수들을 이용하여 다양한 페이지 구성이 가능하다.

| 변수명 | 설명 |
|---------|--------|
| w | 컬랙션명 |
| section | 섹션명 |
| q | 검색 질의어 |

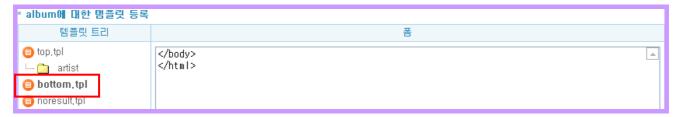
[표 6.1] 필수 입력 변수

5.4.3. 템플릿 등록

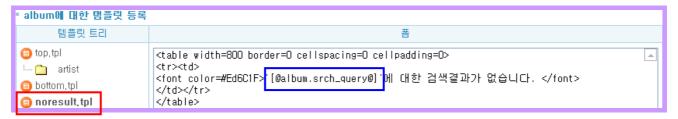
- 컬렉션과 섹션의 내용을 파악하고 변수리스트를 확인하였다면 각각의 템플릿 페이지에 검색 서비스 화면을 만들어줄 템플릿 소스를 등록 하여야 한다. 템플릿 페이지는 컬렉션 템플릿과 섹션 템플릿으로 나뉘어진다.
- 컬렉션 템플릿에는 top, bottom과 전체 결과가 없을 때 출력되어지는 noresult페이지가 있다.
- 섹션 템플릿에서는 head, body, tail과 더많은 결과 출력을 의미하는 more, 페이지 링크 네비게이션을 지원하는 link, 현재 섹션에 결과가 없을 때 화면에 출력하는 noresult 페이지가 있다.
- 템플릿 페이지를 여러 페이지로 구성한 것은 각각의 모듈화 작업을 원활히 하기 위해서 이기 때문이며, 변수는 반드시 변수 리스트에 있는 것을 사용하여야 검색 서비스 화면에서도 정확한 결과를 얻을수 있다.



[그림 6.3] 템플릿 등록(top)

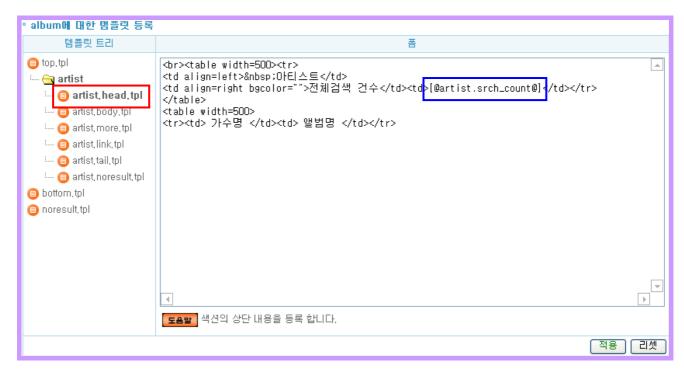


[그림 6.4] 템플릿 등록(bottom)



[그림 6.5] 템플릿 등록(noresult)

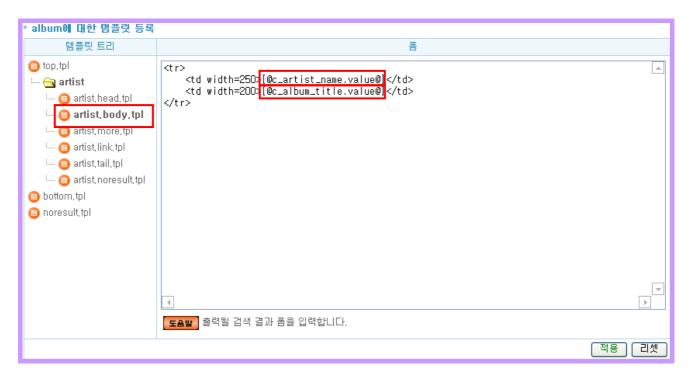
- 템플릿 등록 (컬렉션 top, bottom, noresult)
 - top : 문서의 가장 상위 부분을 작성한다. [그림 6.3]에서는 Html의 head 부분과 자바스크립트 , 그리고 검색 바등의 소스가 들어가게 되는데 공통변수와 폼변수를 이용하여 작성할 수 있다.
 - bottom : 문서의 가장 하단 부분을 작성한다. 상위에서 어떻게 입력 되었는지에 따라 작성 소스가 틀려지며 상단부 입력 소스에 따라서 입력하지 않아도 무방하다. 공통변수와 폼변수를 이용하여 작성할 수 있다.
 - noresult : 컬렉션에 관한 출력결과가 없을때의 메시지를 작성하는 페이지이다. 출력결과가 1건이라도 있게되면 화면에 출력되지 않게 설계 되어있으며 입력 방식은 top, bottom과 동일하다.



[그림 6.6] 템플릿 등록(섹션 head)

● 템플릿 등록 (섹션 head)

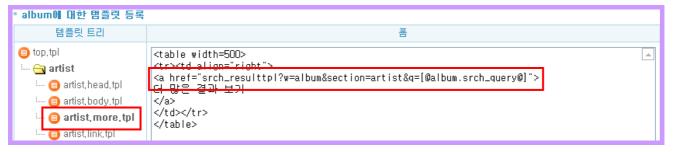
- 각 섹션템플릿의 가장 상위 부분을 작성한다. 컬렉션 별로 해당 섹션이 여러개 존재할 수 있으므로 구분을 쉽게 하기 위해서는 태그를 이용하여 섹션별로 묶어서 작업하는 것을 권장한다. 섹션템플릿에서는 섹션변수가 사용 할 수 있으며, 검색시간, 출력시간, 검색건수를 출력할 수 있다. [그림 6.6]에서도 섹션변수, 공통변수등이 사용되어 졌고, "더 많은 결과보기"가 상단에 포함된 화면이다.



[그림 6.7] 템플릿 등록(섹션 body)

● 템플릿 등록 (섹션 body)

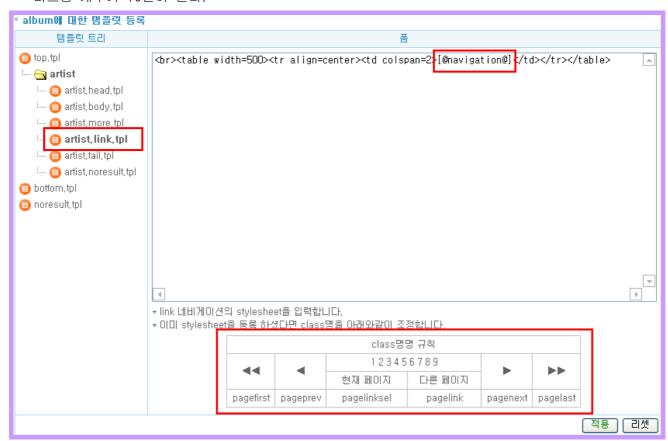
- 각 섹션템플릿의 검색결과 리스트 출력을 담당하는 것이 body부분이다. 출력개수에 의해서 각 문서의 결과를 반복출력한다. 검색결과 확인의 가장 주가 되는 부분이므로 여러가지 변수의 이용법과 출력과정을 미리 확인하는 것이 좋다.
- 출력개수: 우선 한페이지에 출력되는 출력개수를 미리 파악해야 하는데, 그 이유는 검색결과 body 부분이 미리 정해진 출력개수 값 만큼 화면에 리스팅하기 때문이다. 이 출력개수 값은 Exod의 [검색설정->출력개수 입력품] 에서 설정이 가능한데 만일 값이 20으로 되어있다면 출력결과에서는 한 화면에 20건의 문서가 출력된다. 또한 출력개수값은 폼변수를 이용하여 수정 가능하다. 폼변수명은 outmax 이며 입력값은 원하는 출력개수를 주면 된다.
- body변수: 출력개수를 확인했다면 body폼을 작성해야할 차례이다. 페이지 상단에 있는 변수 리스트를 참고 하면 body변수에 관한 정보를 확인할 수 있다. 기본적으로 제공하는 문서ID, 유사도가 변수로 등록되어 있으며 나머지 변수들은 수집되어진 SCD파일에 있는 필드명을 이용하여 각각의 변수명이 등록되어 있다. [그림 6.7] 에서도 등록된 변수를 이용하여 입력된 body 페이지의 화면을 확인할 수 있다.



[그림 6.8] 템플릿 등록(섹션 more)

● 템플릿 등록 (섹션 more)

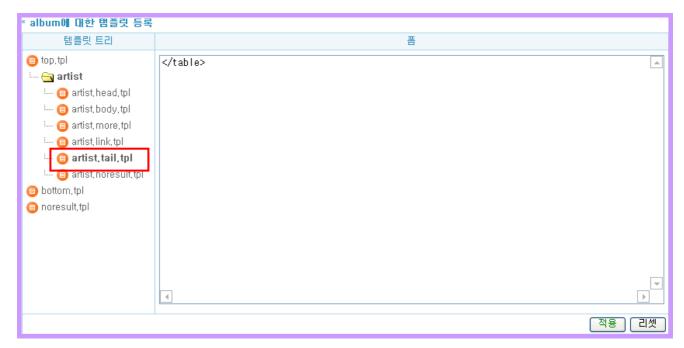
- 각 섹션의 더 많은 결과 보기 부분을 입력하는 폼이다.
- [그림 6.8]의 예제를 보면 공통변수, 폼변수를 이용하여 더 많은 결과 보기 링크를 만드는 방법을 확인할 수 있는데, outmax 값을 10으로 정하였기 때문에 앨범 섹션의 결과 페이지에서는 한 화면에 리스팅 개수가 10건이 된다.



[그림 6.9] 템플릿 등록(섹션 link)

● 템플릿 등록 (섹션 link)

- 각 섹션의 더 많은 결과 보기 후에 페이지 이동을 시켜주는 것이 link 페이지이다. Exod에서는 링크 네비게이션 부분이 미리 만들어져 있다. [그림 6.9] 화면의 도움말에서도 확인 가능 하듯이 페이지 처음으로 가기(◀◀), 이전 페이지로 가기(◀), 다음 페이지로 가기(▶), 맨 끝 페이지로 가기(▶)와 현재 페이지와 다른 페이지로 구분된 숫자링크 또한 미리 정해져 있다. 이처럼 미리 짜여져있는 네비게이션에서 사용자들이 입력할 것은 stylesheet 이다. 각 태그의 클래스명을 확인하여 stylesheet를 입력하면 색감을 변경한다거나 크기를 변형 시키는 등의 작업이 가능하다.



[그림 6.10] 템플릿 등록(섹션 tail)

● 템플릿 등록 (섹션 tail)

- 각 섹션의 가장 하단 부분의 폼을 작성한다. tali 폼에서는 공통변수, 섹션변수가 사용 가능하고 이 부분역시 상단의 페이지에서 어떻게 입력했는지에 따라 소스를 입력하여도 되고 비워두어도 무관하다. [그림 6.10]에서는 이미지를 이용하여 화면에 라인처리를 하고 테이블을 닫는 부분이 들어가 있다.



[그림 6.11] 템플릿 등록(섹션 noresult)

● 템플릿 등록 (섹션 noresult)

- 각 섹션에 관한 출력결과가 없을때의 메시지를 작성하는 페이지이다. 컬렉션에 있는 noresult와 동일하게 출력결과가 1건이라도 있게되면 화면에 출력되지 않게 설계 되어있으며 입력 방식도 비슷하다. 그렇지만 잘 구분해야 할 것은 전체 출력결과가 없을 때 메시지가 중복되어 출력 될수 있으므로 각기 다른 메시지를 만들던가, 한쪽 부분은 아예 입력하지 않는 것이 좀더 짜임새 있는 구성이 될 것이다.

5.4.4. 템플릿 검색

● 템플릿 등록 폼을 이용하여 템플릿을 등록 하였다면 템플릿 메이커 상단의 검색바를 이용하여 페이지 를 확인해 볼수 있다.



5.5. LogStatistics

5.5.1. 통계 등록

- 검색어 로그가 있는 검색 설정 목록을 선택(중복선택 가능)하여 등록 가능하다.
- 통계주기(시간, 일, 주)와 통계타입, 필드등의 옵션 선택이 가능하다.

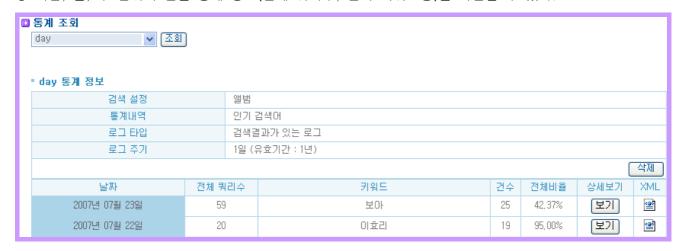


[그림 5.8] 새로운 통계 등록 화면

- 검색로그 선택시 멀티 셀랙트 박스를 이용하여 다중 선택이 가능하다.
- 인기검색어 통계가 아닌 해당 검색어의 피크타임 수치로도 통계를 생성할 수 있다.

5.5.2. 통계 조회

- 통계 등록에서 설정했던 정보의 확인이 가능하고, 통계 설정을 삭제 할 수 있다.
- 시간, 일, 주 단위의 건별 통계 정보(전체 쿼리수, 인기 키워드등)를 확인할 수 있다.

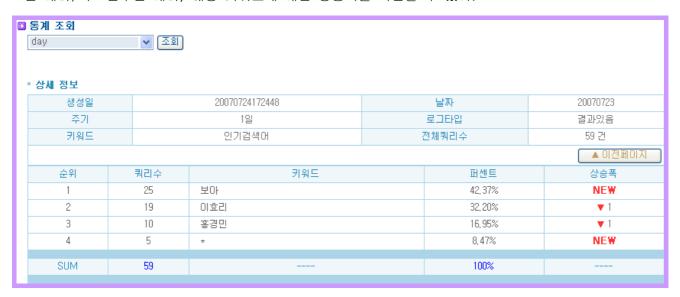


[그림 5.9] 통계 조회 화면

- 상세보기의 보기 버튼을 누르면 상셰 통계를 확인할 수있다.
- 해당 날짜의 XML 이미지를 클릭하면 XML 형식의 통계 정보를 확인할 수 있다.

5.5.3. 상세 통계

- 상세한 로그 통계를 확인할 수 있다.
- 통계 파일의 HEAD 부분에 정의 되어있는 파일 생성일자, 통계 날짜, 주기, 타입등이 확인가능 하다.
- 각 키워드의 순위와 해당 쿼리수, 키워드 내용, 전체에서 차지하는 퍼센티지와 전일대비(시간: 한시간 전 대비, 주: 한주전 대비) 해당 키워드에 대한 상승폭을 확인할 수 있다.



[그림 5.9] 통계 조회 화면

- 상승폭 필드 설명 ▲ : 상승, ▼ : 하강, == : 동일, NEW : 새로추가된 키워드
- 20건씩 나누어진 키워드들의 쿼리수와 퍼센티지의 SUM을 확인 할수 있다.

5.5.4. 통계 XML

- 해당 일자의 키워드 통계 정보를 XML 페이지로 확인할 수 있다.
- 검색결과 XML 페이지에서와 마찬가지로 form 변수를 사용할수 있다.
 - base64 : base64 encoding은 Y, N 값으로 설정할 수 있다.
 - outmax : 한페이지에 출력될 전체 건수를 지정할 수 있다.

```
<?xml version="1.0" encoding="euc-kr" ?:
<meta_storage_list date="2007/05/07 14:15:02">
       <statis name="day":
              <statisfile>/home/kardozo/SKBMT/exod/+SRCHINF/statis/day/d_0503.kwd</statisfile>
               <crtdate>20070504155031
               <logdate>20070503
               <cycle>d</cycle>
               <type>SR</type>
               <field>2</field>
              <totcnt>465</totcnt>
             cranking id="1" count="364" percnt="78.28" hdigit="==">
- <att name="keyword">
                             <![CDATA[ 사람 ]]>
                      </att>
               </ranking
       - <anking id="2" count="28" percnt="6.02" hdigit="NEW">
- <att name="keyword">
<![CDATA[ 서비스 ]]>
                     </att>
               </ranking:
               <ranking id="3" count="21" percnt="4.52" hdigit="NEW">
             - <att name="keyword":
                              <![CDATA[ 장나라 ]]>
                     </att>
               </ranking:
        - (anking id="4" count="14" percnt="3.01" hdigit="8">
- (anking id="4" count="14" percnt="3.01" hdigit="8")
- (anking id="4" count="14" percnt="3.01" hdigit="8" percnt="14" percnt="3.01" hdigit="8" percnt="14" percnt
                      </att>
                </ranking:
```

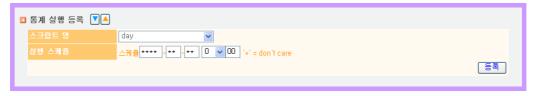
5.6. LogScheduler

- 통계등록을 통해 설정한 컬렉션에 대해 자동실행을 설정
- 시간, 일, 주 3가지 타입의 등록 창으로 설정 되어있다.



[그림 5.11] 시간단위 통계 등록 화면

- 한시간에 한번씩 실행되어져야 하기 때문에 1~60사이의 숫자로 등록할 수 있다.



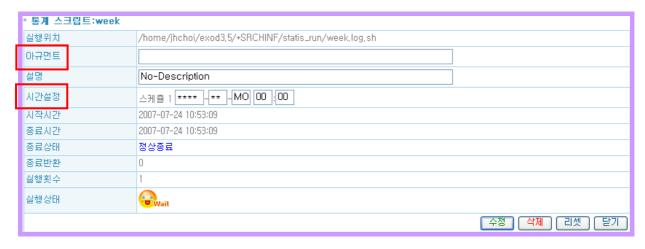
[그림 5.12] 일단위 통계 등록 화면

- 하루에 한번씩 실행되어져야 하기 때문에 1~23 사이의 숫자로 등록할 수 있다.



[그림 5.13] 주단위 통계 등록 화면

- 일주일에 한번씩 실행되어져야 하기 때문에 MO~SU 사이의 단어로 등록할 수 있다.



[그림 5.14] 스케쥴러 업데이트 화면

- 아규먼트를 이용하여 등록하면 해당하는 날짜의 통계를 강제로 생성할수 있다. 예를 들어 위 그림에 아규먼트를 '20070506'으로 넣고 수정후 실행하게 되면 2007년 5월 6일의 통 계 정보를 일단위로 생성할 수 있다.
- 등록창에서는 변경할 수 없었던 시간설정들도 업데이트 창에서는 가능하게 된다.

| " 통계 실행 목록 | | | | | | | | | |
|-------------------------|---------------------|---------------------|------|------|------|------|------|--|--|
| 스크립트 (설명) | 시작시간 | 종료시간 | 종료상태 | 종료반환 | 실행횟수 | 실행상태 | | | |
| ▶ day (No-Description) | 2007-07-24 10:51:25 | 2007-07-24 10:51:25 | 정상 | 0 | 2 | Wait | 강제실행 | | |
| ▶ week (No-Description) | 2007-07-24 10:53:09 | 2007-07-24 10:53:09 | 정상 | 0 | 1 | Wait | 강제실행 | | |

[그림 5.15] 스케쥴러 실행 목록 화면

- 비정상 종료시에 종료반환 코드가 화면에 출력 된다.
- 상태값들은 색인스케쥴러의 설정과 모두 동일하다.

5.7. Search Option

- 기본적으로 구성설정에서 설정되는 값들을 Url 값을 변경함으로써 확인 할 수 있다.
- w (where) 검색설정이름 ex) **w=SONG**
- section 검색 설정 내 섹션이름 ex) w=SONG§ion=artist
- q (query) 질의 ex) w=SONG§ion=artist&**q=홍경민**
- pg (page number) 페이지설정 ex) w=SONG§ion=artist&q=홍경민&**pg=1**
- outmax 한페이지에 출력되는 검색결과 수 ex) w=SONG§ion=artist&q=홍경민&pg=1&outmax=10
- base64 (xml에서만 사용 가능) Base64 encoding 여부: Y(encoding) ex) w=SONG§ion=artist&q=홍경민&pg=1&outmax=10&base64=Y
- sq (subquery)
 - q=*&sq={질의어}
 - query로 뽑은 docs를 q로 검색한 것에 scope하는 방식
- csq (complex subquery)
 - q=*&csq={idxname:검색} | {idxname2:엔진} | {idxname3:검색 ! 엔진}
 - complex query로 뽑은 docs를 q로 검색한 것에 scope하는 방식

★ 정렬우선순위 sort < dsort < psort < fsort

- sort
 - 정렬옵션: 0 (정렬없음), 1,2,3(유사도1,2,3) ex) w=SONG§ion=artist&q=홍경민&pg=1&outmax=10&base64=Y&**sort=0**

★ Url에서는 dsort=10,101 검색 구성설정에서는 1=asc,10=desc

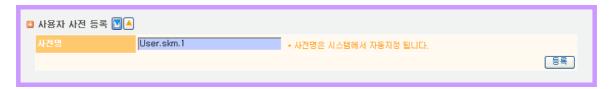
- dsort (DATE sort)
 - 범위필드: 10,11 ~ 640,641
 ex) w=SONG§ion=artist&q=홍경민&pg=1&outmax=10&base64=Y&sort=0&**dsort=11,20**
- psort (PART sort)
 - 분야필드: 10,11 ~ 640,641 ex) w=SONG§ion=artist&q=홍경민&pg=1&outmax=10&base64=Y&sort=0&**psort=11**
- fsort (SORT sort)
 - 정렬필드: 10,11 ~ 640,641 ex) w=SONG§ion=artist&q=홍경민&pg=1&outmax=10&base64=Y&sort=0&**fsort=20**
- d1~d64 (DATE scope)
 - 수치 범위 관련 옵션(64개 까지 설정가능) d1=20060101~20070101&d2=1111~2222
- p1-p64 (PART scope)
 - 파티션 설정 옵션(64개 까지 설정가능) p1=1,4325,32423&p2=34,124

6. 사 전

사전 편집 및 반영 절차는 다음과 같다.

: 사전을 내용을 편집하고, 편집된 내용으로 사전 바이너리 파일을 만든 후, 사전을 빌드후에 sspd DAEMON을 restart 시켜야 사전이 반영된다.

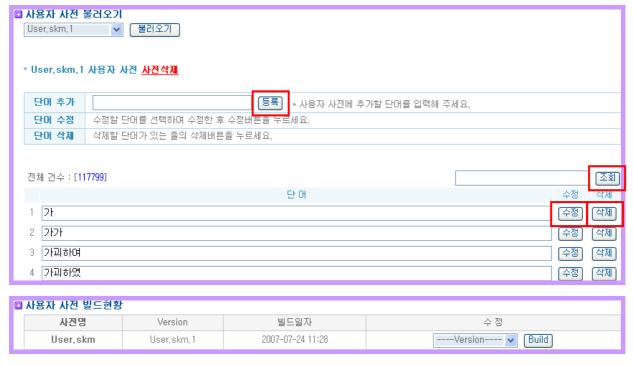
6.1. User Dictionary



[그림 7.1] 사용자 사전 등록 화면

6.1.1. 사용자 사전 등록

- 등록 버튼을 눌러 등록한다.
- 사전명은 EXOD에서 번호순으로 자동생성된다.



[그림 7.2] 사용자 사전 편집 및 빌드 화면

6.1.2. 사용자 사전 편집

- 사전 삭제 : <u>사전삭제</u> 를 눌러서 사전을 삭제할 수 있다. 현재 빌드된 사전은 삭제 할 수 없다.
- 단어 조회 : 단어를 조회하고 싶으면 단어를 입력하고 조회 버튼을 누른다.
- 단어 등록 : 새로운 단어를 등록하고 싶으면 단어 추가 창에서 단어를 입력하고 등록 버튼을 누른다.
- 단어 수정 : 수정하고 싶은 단어의 단어창을 수정 후에 수정 버튼을 누른다.
- 단어 삭제 : 삭제 하고 싶은 단어의 삭제 버튼을 누른다.

6.1.3. 사용자 사전 빌드

- 빌드현황 : Version에서 현재 빌드되어 있는 사전 Version을 확인할 수 있다.
- 빌드: 수정창에서 빌드한 Version을 선택 후 Build 버튼을 누르면 사전이 빌드된다. 하지만 실제로 사전에 적용시키기 위해서는 sspd DAEMON을 restart 시켜야 사전이 반영된다.

6.2. PreAnalysis Dictionary.

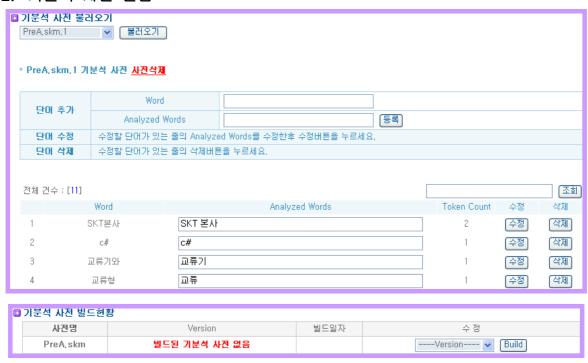


[그림 7.3] 기분석 사전 등록 화면

6.2.1. 기분석 사전 등록

- 등록 버튼을 눌러 등록한다.
- 사전명은 EXOD에서 번호순으로 자동생성된다.

6.2.2. 기분석 사전 편집



[그림 7.4] 기분석 사전 편집 및 빌드 화면

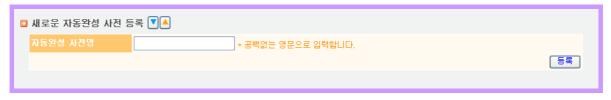
- 사전 삭제 : <mark>사전삭제</mark> 를 눌러서 사전을 삭제할 수 있다. 현재 빌드된 사전은 삭제 할 수 없다.
- 단어 조회 : 단어를 조회하고 싶으면 단어를 입력하고 조회 버튼을 누른다.
- 단어 등록 : 새로운 단어를 등록하고 싶으면 단어 추가 창에서 기분석 사전 등록에서와 같은 방법으로 단어들을 입력하고 등록 버튼을 누른다.
- 단어 수정 : 수정하고 싶은 단어의 Analyzed Words 창을 수정후에 수정 버튼을 누른다.
- 단어 삭제 : 삭제 하고 싶은 단어의 삭제 버튼을 누른다.

6.2.3. 기분석 사전 빌드

- 빌드현황 : Version에서 현재 빌드되어 있는 사전 Version을 확인할 수 있다.
- 빌드 : 수정창에서 빌드한 Version을 선택 후 Build 버튼을 누르면 사전이 빌드된다. 하지만 실제로 사전에 적용시키기 위해서는 sspd DAEMON을 restart 시켜야 사전이 반영된다.

6.3. AutoComplete Dictionary

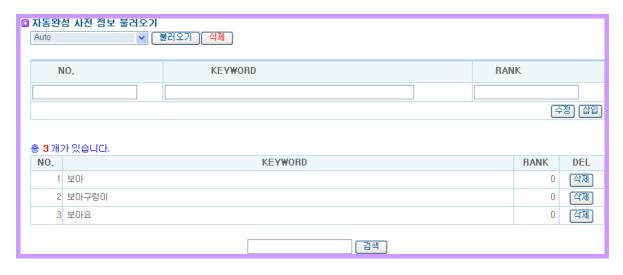
6.3.1. 자동완성 사전 등록



[그림 7.5] 자동완성 사전 등록 화면

● 사전명을 입력 후 등록 버튼을 눌러 등록한다.

6.3.2. 자동완성 사전 관리



[그림 7.6] 자동완성 사전 관리 화면

- 사전 삭제 : <mark>삭제</mark> 버튼을 눌러서 사전을 삭제할 수 있다.
- 단어 조회 : 단어를 조회하고 싶으면 단어를 입력하고 ^{검색} 버튼을 누른다.
- 단어 삽입 : KEYWORD와 RANK(default:0)을 입력하고 ^{삽입} 버튼을 누른다.
- 단어 수정 : 수정하고 싶은 단어의 NO를 입력하고 KEYWORD나 RANK중 수정하고 싶은 곳의 값을 입력후 ^{수정} 버튼을 누른다.
- 단어 삭제 : 삭제 하고 싶은 단어의 ^{삭제} 버튼을 누른다.

6.4. Synonym Dictionary

6.4.1. 동의어 사전 등록

| 🖸 새로운 동의머 사전 등록 | 록 ▼▲ | |
|-----------------|--------------------|----|
| 동의어 사전명 | * 공백없는 영문으로 입력합니다. | |
| | | 등록 |
| | | |

[그림 7.7] 동의어 사전 등록 화면

● 사전명을 입력 후 등록 버튼을 눌러 등록한다.

6.4.2. 동의어 사전 관리

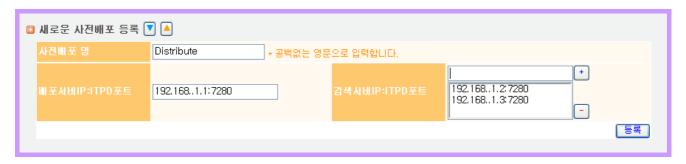


[그림 7.8] 동의어 사전 관리 화면

- 사전 삭제 : <mark>삭제</mark> 버튼을 눌러서 사전을 삭제할 수 있다.
- 단어 조회 : 단어를 조회하고 싶으면 단어를 입력하고 ^{검색} 버튼을 누른다.
- 단어 삽입 : KEYWORD와 RANK(default:0)을 입력하고 ^{삽입} 버튼을 누른다.
- 단어 수정 : 수정하고 싶은 단어의 NO를 입력하고 KEYWORD나 RANK중 수정하고 싶은 곳의 값을 입력후 ^{수정} 버튼을 누른다.
- 단어 삭제 : 삭제 하고 싶은 단어의 ^{삭제} 버튼을 누른다.

6.5. Dict Distributer

6.5.1. 사전 배포 등록



[그림 7.9] 사전 배포 등록 화면

- 사전배포명을 입력한다.
- 배포서버와 검색서버를 입력 후 등록 버튼을 눌러 등록한다.

6.5.2. 사전 배포 실행



[그림 7.10] 사전 배포 화면

- 사전 배포 : [배포하기]를 눌러서 사전을 배포할 수 있다.
- 배포한 사전(User, Pre)을 반영시킬려면 반영할 사전의 sspd DAEMON을 restart 시켜야 한다.

7. 부가기능

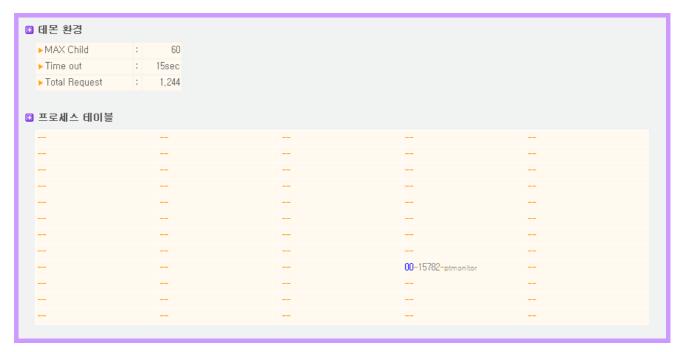
7.1. System Monitor

| ☑ SYSTEM MONITOR <mark>새로고침</mark> | | | | | | | | | |
|------------------------------------|--------|--|-------------|---------------|----------------|--------|-----------------------|-------------------|-----|
| KEF | RNEL | Linux 2,6,8-2-686-smp #1 SMP Tue Aug 16 12:08:30 UTC 2005 i686 | | | | | | | |
| UP1 | ГІМЕ | 11:34:02 up 36 days, 1:47, 3 users, load average: 0,24, 0,07, 0,06 | | | | | | | |
| | | Username | PTS | Mon Day hh:mm | | m | I P | | |
| USE | zpe | hjlee | pts/4 | Jul | 24 | 08:48 | 08:48 (121, 140, 155, | | |
| USI | -nə | Ijb0356 | pts/5 | Jul | 24 | 10:01 | (| (61, 33, 250, 10) | |
| | | hjlee | pts/3 | Jul | 24 | 08:46 | (1 | 21, 140, 155, 22) | |
| ноя | STNAME | searchasp | | | | | | | |
| CPU | | 0,00% us 0,00 | 0% sy 100,0 | 0% id 0, | 00% wa | а | | | |
| | | | total | used | | free | free etc | | |
| ME | MEMORY | Mem | 2075, 86N | 1697, 49M | | 378 | ,37M | 324,60M buffe | ers |
| | | Swap | 2714,941 | 1 121 | 9,90M 1495,04M | | ,04M | 586,12M cacl | hed |
| | | Filesystem | 1K-blocks | Used | Availa | able C | apacity | Mounted on | |
| | | /dev/sda1 | 255912 | 122980 | 11 | 9279 | 51% | / | |
| | | tmpfs | 1037928 | (| 103 | 7928 | 0% | /dev/shm | |
| | | /dev/sdb1 | 141122196 1 | 08813652 | 2 3230 | 8544 | 78% | /data | |
| DIS | | /dev/sda9 | 130099712 1 | 16814064 | 1 1328 | 5648 | 90% | /home | |
| | | /dev/sda8 | 369000 | 8335 | 34 | 1005 | 3% | /tmp | |
| | | /dev/sda5 | 4807056 | 2358452 | 2 220 | 4420 | 52% | /usr | |
| | | /dev/sda6 | 2885780 | 2102912 | 2 63 | 6280 | 77% | /var | |
| | | tmpfs | 10240 | 692 | 2 | 9548 | 7% | /dev | |
| | | | | | | | | | |

[그림 7.1] system Monitor

● EXOD가 설치된 서버의 상태를 보여준다. (KERNEL, UPTIME, USERS, HOSTNAME, MEMORY, DISK)

7.2. Process Monitor



[그림 7.2] Process Monitor

- 프로세스 상태확인
 - 최대 60개의 프로세스를 사용할 수 있다.

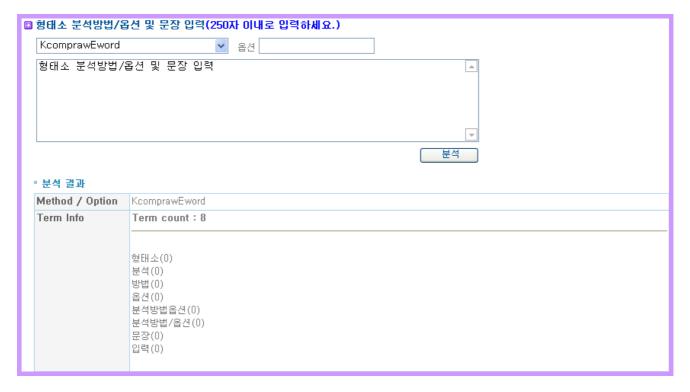
7.3. Others Scheduler



[그림 7.3] Others Scheduler

● 색인 스케쥴러, 로그 스케쥴러외에 사용자가 필요한 스크립트(/+INDEXER/run/*.etc.sh)를 주기적으로 등록해서 실행할 수 있다.

7.4. Morph Analyzer



[그림 7.4] Morph Analyzer

- 빈 텍스트창에 분석을 원하는 문장 혹은 단어를 입력하고, Analyze 버튼을 누르면 분석 결과를 확인할 수 있다.
- 수정한 사전에 따른, 각 색인 방법별 형태소 분석 결과를 확인할 수 있다.
- "Compare with past result"를 선택하면 기존 사전의 결과도 함께 비교할 수 있도록 지원하고 있다.

7.5. Daemon Controler

| • | DAEMON LIST | | | | | | | | |
|---|-------------|------------|-------|-------|----------------|--|--|--|--|
| | DAEMON | PORT (PID) | AC | CTION | Service Status | | | | |
| | ITPD | 7280 (-1) | Start | Stop | Stopped | | | | |
| | SYND | (-1) | Start | Stop | Stopped | | | | |

[그림 7.5] Daemon Controler

● Daemon Control 및 Daemon에서 사용하는 PORT를 설정

7.6. IMAGE VIEWER

● exod/imgs폴더에 사용하고 싶은 IMAGE를 넣고, URL에서

http://EXODserverIP:port번호/imgview?f=IMAGE파일명 을 입력하는 식으로 IMAGE VIEWER를 사용할 수 있다.

Ex) exod/imgs폴더안에 xten.gif 파일을 넣고,

http://EXODserverIP:port번호/imgview?f=xten.gif 을 URL에 입력하면 아래와 같은 결과를 확인할 수 있다.



[그림 7.6] Image Viewer

7.7. 영한 자동 변환

● URL에서

http://EXODserverIP:port번호/eng2han?q=QEURY를 입력하는 식으로 영한 자동 변환 기능을 사용할수 있다.

Ex) URL에

http://EXODserver IP:port번호/eng2han?q=ahskflwk 을 입력하면 아래와 같은 결과를 확인할 수 있다.



[그림 7.7] 영한 자동 변환

● eng2han Library는 EXOD 설치 EXOD 설치 폴더 밑에 lib/e2hlib 에 들어가 있다.

8. REAL TIME INDEX

● 실시간 색인의 ReadMe file의 내용은 아래와 같다. 실시간 색인하는 방법을 간단히 설명하고 있다. exod/+INDEXER/rtinstruction 폴더는 컬렉션 별로 폴더를 가지고 있다. 실시간 삽입/삭제가 필요한 경우 해당 컬렉션의 폴더안에 아래에 나와있는 방법으로 삽입/삭제를 하면 된다.

```
exod/+INDEXER/rtinstruction/collectionname/
* <<!NSERT>>
* file_name =
* insertscd* || insertary* 로 파일생성
 file_content =
 insertscd* : scd파일을 직접입력
   ex) ^III:title
       ^CON:content
       ^DAT:2006.11.28
 insertary* : 직접 query를 입력
   ex) select * from MUSIC where name=从이
* <<DELETE>>
 file name =
 deletedocid* || deletegry* 로 파일생성
 file_content =
 deletedocid* : Doc_ID
   ex) 1234 1 2 3
 %%% 한번에 여러 문서를 지울수 있다.
 _deletegry* : idx_name <indexed_boolean> keyword
  ex) INDEX_1 <indexed_boolean> 사이
 %%% online set을 만들어서 사용하는 경우
    deletedocidrun*
    deletedocidon*
    deletearv*
```

[그림 8.1] 실시간 색인 ReadMe 파일

8.1. run set에 바로 적용

실제 service되고 있는 set에 xten_InsertDoc xten_DeleteDoc을 이용해서 실시간으로 문서를 삽입/삭제를 하는 방법이다. 실시간 삽입/삭제시 중간에 멈추게 되면 색인이 깨질수도 있다. 반드시 삽입/삭제가 완료될 때까지 기다려야 한다. 현재 EXOD에서는 한 색인주기에 10만개까지의 문서를 삽입할 수 있게 설정되어 있다.

8.1.1. Insert

● SCD 삽입

```
exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM> Is insertscd exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM> cat insertscd ^c_album_key:1 ^c_album_type:1 ^c_album_title:홍경민 리메이크 - Emotion In Memory ^c_artist_key:264 ^c_artist_name:홍경민 ^c_reg_date:2005-09-30 17:38:11 ^c_rel_date:20050928 exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM>
```

[그림 8.2] 실시간 SCD 삽입 화면

- insertscd로 시작하는 파일명으로 해당 컬렉션 폴더안에 scd내용을 넣어서 파일을 만들어주면 된다. scd 안에 여러 개의 문서가 있어도 삽입이 가능하다.
 - ex) "홍경민 리메이크 Emotion In Memory" 라는 문서를 삽입할 경우
- QRY 삽입

```
exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM> Is
insertgry
exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM> cat insertgry
select + from MUSIC_ALBUM where ^c_album_key=90008441
exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM>
```

[그림 8.3] 실시간 Query 삽입 화면

- insertqry로 시작하는 파일명으로 해당 컬렉션 폴더안에 db에서 삽입할 문서를 구별해서 가져올 수 있는 query를 넣어서 파일을 만들어주면 된다. query를 이용하여 수집기가 scd를 만들어서 문서를 삽입하는 방법이다.
 - ex) c_album_key가 90008441인 문서를 삽입할 경우

8.1.2. **Delete**

● DocID 삭제

```
exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM> Is
deletedocidrun
exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM> cat deletedocidrun
1
2
100
1000000
exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM>
```

[그림 8.4] 실시간 DOC ID 삭제 화면

- deletedocidrun으로 시작하는 파일명으로 해당 컬렉션 폴더안에 지울 문서의 DocID를 넣어서 파일을 만들어주면 된다.
 - ex) DocID가 1, 2, 100, 1000000인 문서를 지울경우
- QRY 삭제

```
exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM> Is
deletegry
exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM> cat deletegry
ARTIST <indexed_boolean> "홍경민"
exodtft@srchteam:~/jihoon/exod/+INDEXER/rtinstruction/ALBUM>
```

[그림 8.5] 실시간 Query 삭제 화면

- deleteqry로 시작하는 파일명으로 해당 컬렉션 폴더안에 색인명과 검색방법 검색어를 넣어서 파일을 만들어주면 된다.
 - ex) ARTIST 색인에 <indexed_boolean>로 "홍경민"이라는 검색어로 찾아지는 문서를 지울 경우
 - ARTIST <indexed_boolean> "홍경민"

8.2. online set을 만들어서 적용

실제로 service되는 set이 아닌 setonline을 만들어서 실시간 색인을 하는 방법이다. 이방법을 사용하면 원본 색인에는 영향을 끼치지 않는다는 장점이 있다. 단 검색시 service되는 set의 결과와 setonline의 결과를 merge시키는 시간이 추가된다.

8.2.1. **Insert**

● "run set에 바로 적용"의 방법과 같다.

8.2.2. **Delete**

● DocID 삭제

- deletedocidrun과 deletedocidon으로 시작하는 파일명으로 해당 컬렉션 폴더안에 지울 문서의 DocID를 넣어서 파일을 만들어주면 된다. deletedocidrun 과 deletedocidon으로 나눠져 있는 이유는 내가 지울문서가 service되고 있는 set의 문서인지 setonline의 문서인지 알아야 하기 때문이다. 파일은 만드는 방법은 "run set에 바로 적용"의 방법과 같다.

● QRY 삭제

- "run set에 바로 적용"의 방법과 같다. 단 주의 할 점은 service되고 있는 set의 문서와 setonline의 문서를 모두 검색해서 지운다. 지우기 전에 다시 한번 검색어가 내가 지울문서만 지우게 될 지 생각해보는 것을 추천한다.

9. Add On

● EXOD외에 타 웹 인터페이스 관리기 및 사이트 링크. 등록하고 싶은 사이트가 있으면 /conf/addon.conf 파일에 메뉴명=사이트주소 폼으로 추가한다. ex) 구글=http://www.google.co.kr



[그림 9.1] Add On 화면