

# TV-anytime 표준을 지원하는 멀티미디어 검색 및 지능화 서비스 시스템의 설계 및 구현

이종설, 이윤주, 박우출, 이석필, 조위덕

전자부품연구원 인터넷미디어연구센터

{jslee, yjlee0618, wcpark, lspbio, chowd}@keti.re.kr

## Design and Implementation of a System for multimedia retrieval and Intelligent service supporting the Specification of TV-Anytime

Korea Electronic Technology Institute

Jong-Sul Lee<sup>0</sup>, Yun-Ju lee, Woo-Chool Park, Seok-Pil Lee, We-Duke Cho  
Internet media Research Center, KETI

### 요 약

본 논문에서는 인터넷 및 방송망 등과 같은 분산된 다수의 환경을 통해 배포되는 대용량의 멀티미디어 콘텐츠를 검색 및 저장하기 위한 TV-Anytime 멀티미디어 콘텐츠 맞춤형 시스템을 설계 및 구현한다. 본 논문에서는 멀티미디어 콘텐츠의 저장 및 검색, 획득을 위해 TV-Anytime 표준과 Mpeg-7 표준을 사용하였다. 구현한 저장관리 시스템은 메타데이터 객체 관리자, 메타데이터 인덱스 관리자, 개인화엔진, 메타데이터 검색 API, Location Handler 모듈 등으로 구성된다. 본 논문에서 구현한 저장관리 시스템은 메타데이터의 처리를 위해 DOM/SAX 인터페이스를 지원하며 메타데이터에 대한 내용검색과 애트리뷰트에 기반한 검색을 제공하며, 사용자 프로파일을 이용하여 개인화된 지능형 맞춤 서비스를 제공한다.

#### 1. 서론

방송망과 네트워크 기술 그리고 컴퓨터 기술의 발달에 따라 방대한 양의 분산된 멀티미디어 데이터를 처리할 수 있는 색인 및 검색 도구의 요구가 커지고 있으며, 개인용 대용량 저장 매체를 이용하여 자기가 선호하는 콘텐츠를 원하는 시간에 시청하기 위한 개인화 된 맞춤형 서비스의 필요성이 대두되었다. 대용량의 멀티미디어 콘텐츠를 검색을 위해서 현재 웹 상에서는 텍스트 기반 정보검색 기술과 P2P 애플리케이션 등 많은 기술들이 개발되어 사용되어 지고 있다. 그러나, 이와 같은 응용기술들을 방송과 연계하기에는 많은 문제점이 발생한다. 이에 본 논문에서는 방송과 인터넷이 결합된 통합환경에서 개인화 된 콘텐츠의 검색/저장/재생을 위한 표준인 TV-Anytime 표준에 기반한 대용량 멀티미디어 콘텐츠를 위한 맞춤형 시스템을 설계 구현한다. 이를 위해 멀티미디어 콘텐츠를 위한 메타데이터 저장 관리기를 설계 및 구현하고, 대량의 멀티미디어 콘텐츠 중에서 특정 배우, 감독 등의 여러 가지 검색 조건으로 콘텐츠를 검색 및 선택하기 위한 인터페이스를 설계 구현한다. 이 인터페이스를 이용하여 사용자는 하나의 멀티미디어 콘텐츠 내에서 특정 비디오세그먼트 (Video Segment)에 대한 검색이 가능하다. 또한, 개인화 서비스 엔진(Personalized service engine)을 이용하여 개인에 대한 맞춤형 서비스의 지원이 가능하다.

본 논문의 구성은 다음과 같다. 2장에서는 멀티미디어 콘텐츠를 위한 메타데이터와 TV-Anytime에서의 CRID와 로케이터에 대하여 살펴본다. 3장에서는 본 연구에서 구현한 시스템인 X-REY (eXpert REsolution sYstem)에 대

해서 알아본다.

#### 2. 관련 내용

##### 2.1 메타데이터란 무엇인가?

메타데이터란 비디오, 오디오, 텍스트 등 실제로 저장하고자 하는 데이터와 직접적으로 혹은 간접적으로 연관된 정보를 제공하는 데이터를 나타낸다. 메타데이터에 대한 연구는 최근 다양한 분야에서 많은 연구가 진행되어지고 있다. 현재 멀티미디어 콘텐츠를 위한 메타데이터를 정의하고 표준을 만드는 국제표준화단체는 대표적으로 MPEG-7과 TV Anytime Forum이 있다. MPEG-7은 주로 멀티미디어 콘텐츠의 저장, 전송 그리고 검색을 위한 기술(Description)방법에 대한 표준을 제정하고 있으며 2001년 11월에 최종 국제 표준초안인 FDIS를 제정하였다. 반면에 TV Anytime Forum은 원래 저장매체가 있는 사용자환경에서 디지털 방송 서비스를 제공하기 위한, 즉 Anytime Service를 위한 표준을 제정하고 있으며 2002년 8월에 ver.1.3 메타데이터 규격을 완성하였다. Anytime Service란 사용자가 원하는 콘텐츠를 자동적으로 Search, Capture, Store하고 사용자가 원하는 시간에 볼 수 있는 서비스를 말한다. 그 밖에 미국의 CableLabs에서는 "VoD 메타데이터"라는 디지털 케이블망을 이용한 VoD서비스를 위한 메타데이터규격을 만들어 놓고 있다.

본 연구에서 사용한 TV-Anytime 메타데이터의 종류는 크게 4부분으로 나눌 수 있다. 첫째로 콘텐츠 기술 메타데이터(Content description metadata)로 콘텐츠의 기술정보를 나타내는 메타데이터이다. 여기에는 Description, Audio/Video Information, Program Information, Group

Information, Media Review DS에 대해서 정의한다. 콘텐츠의 기술정보로는 프로그램 제목, 장르, 줄거리 개요, 등장인물 정보 등이 해당된다. 둘째는 프로그램에 관련된 프로그램 방송시간, 채널정보, 다운로드 위치, 프로그램 개체의 이용 규정 등 로케이션 정보와 서비스 정보를 나타내는 메타데이터로서 개체 기술 메타데이터(Instance description metadata)이다. 셋째는 한 콘텐츠 안에서 하이라이트나 키프레임을 나타낼 수 있는 세그멘테이션 메타데이터(Segmentation metadata)이며, 마지막으로 사용자의 취향이나 습성을 나타내는 사용자 메타데이터(Consumer metadata)로 나눌 수 있다. 사용자 메타데이터는 크게 사용자의 콘텐츠 사용 내역에 대한 정보인 사용정보(Usage history)와 사용자가 직접 기입한 사용자의 선호 정보(User preference)로 나누어진다.

## 2.2 Content Reference Identifier 와 로케이터

### 2.2.1 CRID

TV-Anytime 시스템에서 사용자는 방송망 또는 리턴채널을 통해 전송된 ECG 정보 또는 트레일러를 통해 멀티미디어 콘텐츠에 대한 정보를 획득하거나, 콘텐츠 검색 인터페이스를 통해 콘텐츠를 검색하여 콘텐츠 선택할 수 있다. TV-Anytime에서는 이와 같은 콘텐츠의 획득 절차를 위해 Content Reference Identifier(CRID)를 정의하였다. CRID는 사용자가 선택한 콘텐츠가 정확히 어떤 것인지 가리키는 일종의 식별자이다. 선택된 CRID는 위치 분석(Location Resolution)을 통해 콘텐츠의 저장된 위치, 또는 방송 시간 등을 나타내는 로케이터를 얻을 수 있고, 이 로케이터를 이용해 사용자는 선택한 콘텐츠를 획득할 수 있다. 하나의 CRID는 위치 분석과정에서 여러 가지의 CRID들로 나뉘어 질 수 있는데, 예를 들어 사용자가 선택한 콘텐츠인 사극 '제국의 아침'이 속편들을 포함한 타이틀인 경우 위치 분석 과정에서 최초 CRID는 각 속편들 하나 하나의 CRID들로 나뉘어진다. 이와 같은 CRID를 표현하는 형식은 다음과 같다.

CRID://<authority>/<data>

authority는 콘텐츠에 대해 저작권을 가진 콘텐츠 제작자의 도메인 이름으로 이루어지며 동일 도메인 이름을 가진 여러개의 권한이 있을 경우 추가 될 수 있다. authority는 아래와 같은 형식으로 표현 될 수 있다.

<DNS name><name\_extension>

따라서 CRID는 아래 형식으로 표현된다.

CRID://<DNS name><name\_extension>/<data>

### 2.2.2 로케이터

콘텐츠의 위치는 여러 가지 다양하게 있을 수 있다. 예를 들어 방송국의 콘텐츠 저장 서버일수도 있고, 어느 서비스 업체의 데이터 저장 서버일수도 있고, 인터넷 웹서버의 파일형태로 있을 수도 있다. 따라서 사용자의 콘텐츠 검색을 통해 선택된 콘텐츠는 언제 어디로부터 어떻게 사용자에게 전송될지를 알아야 하는데 이런 정보를 로케이터라 한다. 로케이터는 사용자가 콘텐츠를 선택 한 후 획득한 CRID를 이용해 Location Resolution과정에서 검색해 낸다. 이렇게 검색된 로케이터 정보(해당 시간, 해당 위치, 해당 전송방법)에 따라 사용자에게 콘텐츠가 전송된다. 로케이터를 표현하는 방법은 아래와 같다.

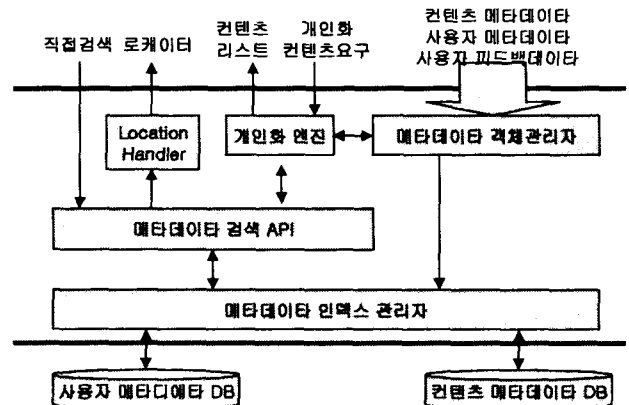
<transport mechanism>:<transport system specific>

위 로케이터 표현 양식중에 <transport mechanism>은 공중파 또는 인터넷 망과 같은 콘텐츠를 전송하는 방법을 기술한 부분이다. <transport system specific>은 콘텐츠 저장위치(location), 콘텐츠 가용 시간의 종류(type of availability)에 대한 정보를 포함하고 있다.

## 3. X-REY (eXpert REsolution sYstem) 시스템의 구현

본 연구에서 구현한 시스템을 이용하여 사용자는 직접 콘텐츠에 대한 검색이 가능하며, 또한 시스템 등록 시 작성한 사용자 메타데이터를 바탕으로 사용자에게 적합한 콘텐츠 리스트를 자동으로 받는 맞춤형 서비스를 받을 수도 있다. 사용자는 전달받은 콘텐츠의 CRID를 통해 로케이터를 획득할 수 있으며, 획득한 로케이터가 방송용 로케이터 일 경우에는 해당 프로그램에 대해서 녹화가 가능하며, URL을 통한 On-Demand 콘텐츠일 경우에는 해당 콘텐츠에 대한 다운로드를 통해 획득이 가능하다.

### 3.1 시스템 구조



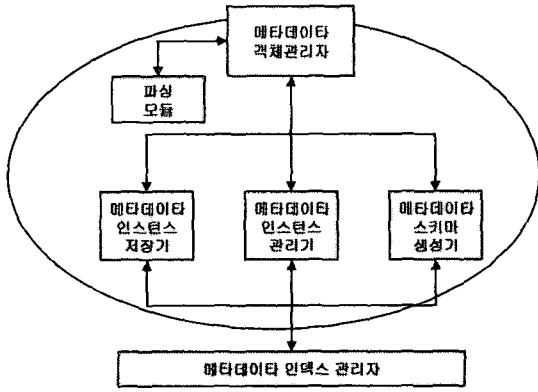
(그림 2) 시스템 구조도

본 연구에서 구현된 X-REY 서버의 구성은 (그림 2)와 같다. 메타데이터 서버는 1)메타데이터 객체 관리자 2) 메타데이터 인덱스 관리자 3) 개인화 엔진 4) 메타데이터 검색 API 5) Location Handler 모듈로 구성된다. 콘텐츠 제작자는 배포할 콘텐츠에 대해 콘텐츠 기술 메타데이터와 세그먼트 메타데이터, 그리고 콘텐츠의 배포정보인 개체 정보 메타데이터를 생성하여 메타데이터 서버에 등록한다. 메타데이터 서버에 등록된 메타데이터는 메타데이터 객체 관리자를 통해 데이터베이스에 저장되고 각 부분별 구분 분석이 실행되어지며 메타데이터 인덱스 관리자를 통해 파싱된 부분에 대한 인덱스가 생성된다. 사용자는 검색 인터페이스를 통해 콘텐츠에 대한 검색을 수행하고 CRID를 획득한다. 사용자가 콘텐츠를 획득을 최종 결정하게 되면 Location Handler를 통해 로케이터를 획득하게 된다. 또한 사용자가 작성한 사용자 프로파일은 사용자 메타데이터로 작성되어 개인화 엔진에 제공되어진다. 개인화 엔진에서는 정보 필터링 모듈을 통해 저장된 개인 선호 정보와 콘텐츠 메타데이터를 바탕으로 각 사용자에게 대해 개인화된 콘텐츠 리스트를 결정하게 된다. 개인화된 콘텐츠 리스트는

사용자에게 제공되어 사용자의 반응을 기다리고 사용자의 해당 콘텐츠에 대한 피드백을 통해 사용자의 선호도를 학습하여 변경한다.

### 3.2 메타데이터 객체관리자

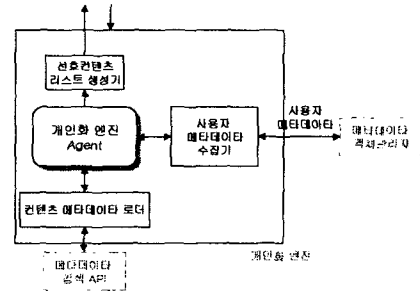
메타데이터 객체관리자는 콘텐츠 기술 메타데이터, 세그먼트 기술 메타데이터, 개체 기술 메타데이터 등을 저장 및 관리한다. 이를 위해 저장관리자는 메타데이터 스키마에 따라 데이터베이스 스키마를 생성한다. 또한, 메타데이터를 저장하고 삭제 변경하는 등 관리하는 일을 담당한다. 본 연구에서 메타데이터 저장관리자는 메타데이터를 파일로 유지하는 동시에 데이터베이스에 파싱해서 관리하는 혼합 저장방식을 적용한다.



(그림 3) 메타데이터 객체 관리자

메타데이터 객체관리자는 스키마 생성기, 인스턴스 관리기, 인스턴스 저장기, 파싱모듈 등으로 구성되며, 각 모듈들에 대한 통합 인터페이스를 제공한다. 저장될 메타데이터는 메타데이터 객체 관리자를 통해 파싱 모듈로 전해진다. 파싱 모듈에서는 TV-Anytime 메타데이터 표준에 따라 메타데이터의 Descriptor 구조를 추출하기 위한 API로 W3C에서 제안하고 있는 JAXP1.2를 기본으로 사용한다. 파싱 모듈에서는 DOM/SAX 인터페이스를 이용하여 Descriptor 단위의 추출을 수행한다. 추출된 Descriptor를 바탕으로 메타데이터 스키마 생성기에서는 해당 Descriptor에 대한 스키마의 생성 여부를 확인하여 메타데이터 저장관리기의 기본이 되는 스키마를 생성하며 관리한다. 메타데이터 인스턴스 저장기에서는 추출된 DS와 해당 내용을 각 테이블에 저장한다. 이때 콘텐츠는 CRID와 로케이터를 통해서 식별되어진다. 메타데이터 인스턴스 관리기는 저장된 메타데이터에 대해 변경 및 삭제, Fetch 등의 작업을 수행한다. Fetch는 3rd Party에서 메타데이터의 일부분 또는 전체를 필요로 할 경우 제공하는 기능이다. 메타데이터의 일부분 즉 DS단위의 Fetch는 DS단위의 저장된 DS일 경우는 해당 부분을 제공해주며 그렇지 않은 경우는 전체 문서에서 해당 부분을 추출해서 전달한다. 효과적인 메타데이터의 관리를 위해 문서 저장기 내에서는 메타데이터에 대하여 콘텐츠 식별자(Content ID)를 부여하여 관리한다.

### 3.3 검색 인터페이스 및 사용자 선호 정보를 이용한 개인화 엔진



(그림 4) TV-Anytime에 기반한 개인화 엔진의 구조도

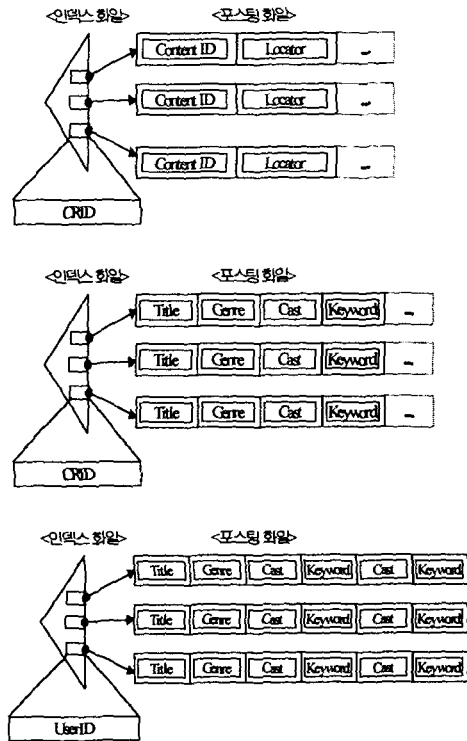
(그림 4)는 사용자 메타데이터를 이용한 개인화 엔진(Personalized Engine)의 전체적인 구조를 보여주고 있다. 개인화 엔진에서 사용되어지는 사용자에 대한 정보는 첫째로 사용자 인터페이스에서 제공하는 일정한 양식에 따라서 항목별로 사용자의 선호 정보를 직접 입력하는 명시적 선언(Explicit Declaration)에 의한 방식과 특정 프로그램 시청 중에 리모콘에서 제공하는 일정한 버튼을 이용하여 해당 프로그램에 대한 선호 표시를 하는 묵시적 선언(Implicit Declaration)에 의한 입력 방식을 바탕으로 생성되는 사용자 선호 메타데이터(User preference metadata)와 둘째로 사용자가 셋톱박스의 작동에 대한 기록을 나타내는 콘텐츠 사용 정보(Usage history) 등이 제공되어진다. (그림 2)에서와 같이 사용자의 개인화 정보들은 메타데이터화 되어 메타데이터 객체 관리자를 통해 저장되어진다. 개인화 엔진 Agent는 시스템의 부하가 적은 적당한 시간에 사용자 메타데이터 수집기를 통해 사용자 정보에 대한 수집 및 정렬을 수행하고 해당 사용자에 대한 선호도와 콘텐츠 메타데이터와의 필터링을 통해 사용자에 적합한 개인화 된 콘텐츠 리스트 정보를 생성한다. 현재 개인화 엔진을 통해 수집 및 비교되는 항목정보로는 Title, Channel, Genre, Cast, Keyword, Time DS이다. 다음은 개인화 엔진 Agent에서 사용자의 선호 정보와 콘텐츠 메타데이터 정보를 이용하여 사용자에 적합한 콘텐츠를 자동으로 추천하는 과정을 단계별로 설명한 것이다.

- Step 1 : 개인 선호 정보 인덱스로부터 항목별로 사용자 선호 정보 추출
- Step 2 : 콘텐츠 인덱스로부터 항목별(Title, Channel, Genre, Cast, Keyword, Time)로 사용자 선호 정보를 포함하는 프로그램들의 식별자를 검색하여 항목별 추천 프로그램 리스트 작성
- Step 3 : 항목별 추천 프로그램 리스트의 합집합을 구하고 추천 후보 프로그램 리스트 작성
- Step 4 : 추천 후보 집합에 있는 각각의 ContentID에 대하여 항목별 추천 프로그램 리스트의 존재 여부를 판단하여 6차원의 선호도 벡터 구성한다.
- Step 5 : 각각의 추천 후보 프로그램의 선호도 벡터의 가중치의 합을 구함
- Step 6 : 추천 후보 프로그램들을 각각의 가중치 합을 기반으로 정렬
- Step 7 : 상위 N개의 프로그램들만 사용자에게 추천

### 3.4 메타데이터 인덱스관리자

파싱 모듈을 통해 구조 분석된 메타데이터는 메타데이터 객체 관리자를 통해 저장된 이후 인덱스 관리자에게 전달된다. 인덱스 관리자에서는 메타데이터의 검색속도와 관리의 편리성을 위해 그림과 같이 역파일 링크드 리스트 형태로 저장된다. 첫째로 그림 과 같이 location resolution을

위해 CRID 와 ContentID, 로케이터에 대한 역파일 인덱스를 생성 관리하며, 둘째로 콘텐츠 메타데이터에 대한 검색을 위해 Title, Genre, Cast, Keyword등에 대한 DS 값과 CRID에 대해서 역파일 인덱스를 생성 관리한다. 마지막으로 개인화 정보를 위해 UserID 와 개인 선호 정보 항목에 대해서 역파일 인덱스를 생성 관리한다.



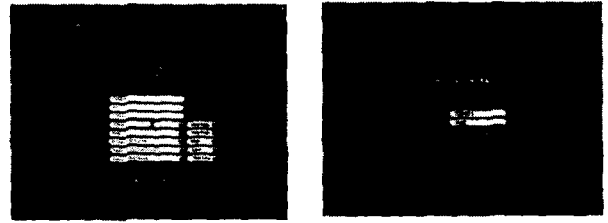
(그림 5) 메타데이터 인덱스 관리자의 역파일 링크드 리스트 구조

### 3.5 Location Handler

개인화 엔진을 통해 사용자에게 제공된 콘텐츠 리스트의 CRID와 직접 검색을 통해 사용자가 획득한 CRID는 Location Handler 모듈을 통해 해당 콘텐츠에 대한 로케이터를 추출한다. Location Handler 에서는 사용자에게 선택되어진 CRID를 바탕으로 해당 콘텐츠에 대한 로케이터를 추출한다. 그림과 같이 로케이터를 추출할 CRID는 Location Handler의 입력으로 주어지며, Location Resolution Handler에서는 메타데이터 인덱스 관리자에서 구성한 역파일 정보를 바탕으로 해당 콘텐츠의 로케이터를 추출한다. 추출되어진 로케이터는 사용자에게 전송되어진다. 본 연구에서 구현한 Location Handler는 자체 시스템의 메타데이터 인덱스 관리자를 통해서 해당 콘텐츠의 로케이터를 추출하지만, 시스템 확장을 통해 다른 위치의 Location resolution 시스템을 통해서도 로케이터의 추출이 가능하다.

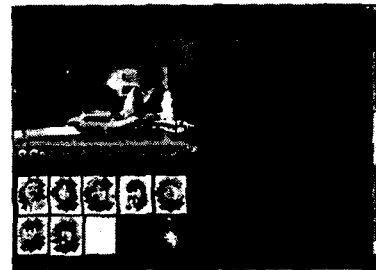
### 3.6 사용자 인터페이스

사용자는 맞춤형 서비스를 이용하기 위해 최초 사용시 또는 사용중 개인 정보의 변경시 사용자 정보를 (그림 6)과 같이 입력하여야 한다. 입력된 정보를 바탕으로 TV-Anytime 메타데이터의 사용자 기술 메타데이터가 생성되며 사용자 메타데이터는 X-REY 시스템으로 전달되어 개인화엔진에 전달된다.



(그림 6) 사용자 인터페이스 화면

로그인한 사용자는 X-REY 시스템에서 전송한 콘텐츠의 리스트를 확인할 수 있다. 사용자에게 대한 개인별 콘텐츠 리스트는 (그림 7)과 같다. 메인 화면은 사용자의 콘텐츠 선호도에 맞게 추출, 저장된 콘텐츠를 보여주는 콘텐츠 리스트(Electronic Content Guide)와 동영상 플레이어, 이미지 뷰, 텍스트 뷰로 구성된다. 사용자가 콘텐츠 리스트에 나타난 콘텐츠 중에서 원하는 콘텐츠 바에 버튼을 클릭하면, 콘텐츠에 관련된 이미지와 텍스트 정보가 디스플레이되며 예고편 파일이 있을 경우 해당 동영상상을 재생하게 된다. 사용자가 해당 콘텐츠에 대한 획득을 결정하게 되면 Location Handler를 통해 해당 콘텐츠의 로케이터를 획득하여 최종적으로 콘텐츠를 획득하게 된다. 또한, 사용자가 획득한 콘텐츠에 대한 재생 중의 동작에 대해서는 사용 히스토리를 생성하여 콘텐츠에 대한 개인화 엔진의 피드백 정보로 활용되어진다.



(그림 7) 지능형 개인화 서비스 ECG

### 4. 결론

본 논문에서 구현한 X-Rey 시스템은 TV-Anytime 메타데이터를 이용하여 동영상 콘텐츠에 대한 검색 및 해당 콘텐츠에 대한 로케이터 추출이 가능한 시스템이다. 동영상 콘텐츠에 대한 검색은 현재 Text에 기반하여 각 Descriptor에 대한 키워드 검색만이 가능하다. 또한 사용자에 대한 선호 정보를 바탕으로 개인화 서비스가 가능한 시스템이다. 현재 콘텐츠의 내용 기반 검색에 대한 표준을 정의하고 있는 Mpeg-7 과 TV-Anytime 간의 업무 협력 관계를 고려할 때 본 시스템 또한 내용 기반 검색 엔진에 대한 연구가 추가적으로 필요하다.

### 참조 문헌

- [1] 이석필, "TV Anytime Forum 현황", 한국방송공학회지 제7권 제1호 2002.3
- [2] 강경옥 김진웅, "메타데이터를 이용한 지능형 방송 기술", 대한전자공학회지 제29권 제 7호 2002.7
- [3] 박우철, 이종설, 이윤주, 이석필, 조위덕, "콘텐츠 메타데이터와 사용자 선호도에 의한 지능화 EPG 서비스 시스템", 한국통신학회 하계 학술대회 2002. 7.
- [4] MPEG-7 Multimedia Description Schemes XM(v5.0), La Baule, France, Oct. 2000.
- [5] The TV-Anytime Forum <http://www.tv-anytime.org>