

# WWW을 이용한 입후보자 프리젠테이션 서비스를 위한 멀티미디어 정보 검색 시스템

유재수, 최한석  
목포대학교 전산통계학과

## A Multimedia Information Retrieval System for Candidate Presentation Services Based on WWW

Jae Soo Yoo, Han-Suk Choi  
Department of Computer Science & Statistics, Mokpo National University

### <요 약>

본 논문에서는 WWW을 이용하여 입후보자 프리젠테이션 서비스를 제공하는 멀티미디어 정보 검색 시스템을 설계하고 개발한다. 개발된 입후보자 프리젠테이션 서비스 정보 검색 시스템은 유권자들에게 정확하고 공정한 정보를 제공하여 입후보자 선택을 위한 지침을 제시하고 입후보자들간의 자질 및 정강, 정책 등을 비교 검색할 수 있다. 또한 기존의 WWW에서 화일 단위로 정보를 다룸으로써 발생하는 문제를 해결하기 위해, 입후보자에 관한 정형화된 데이터를 사용하여 데이터베이스를 구축하고 구축된 데이터베이스와 WWW 시스템과의 통합을 도모한다.

### 1. 서론

최근들어 정보화의 중요성에 대한 인식이 증가되고 초고속 정보 통신망의 구축이 이루어지고 있음에 따라 정보는 기관과 조직에서 뿐만 아니라 개인에게도 매우 중요한 자산으로 인식되고 있다. 이러한 정보는 그 종류가 다양하고 양적으로 방대하다. 이와같은 방대한 정보를 효율적으로 처리하기 위해 하이퍼미디어 시스템(hypermedia system)이 등장하였다[4].

WWW(World Wide Web)이란 인터넷 상에 방대하게 산재하고 있는 정보들을 효율적으로 접근하기 위한 대규모의 분산 하이퍼미디어 시스템이라고 할 수 있다[3,4]. WWW는 클라이언트/서버 모델을 기반으로 하고 있으며 각기 전세계의 서로 다른 장소에 존재할 수 있는 각종 문서들이 키워드를 통해 서로 연결되어 있다. 또한 텍스트, 그래픽, 이미지, 음성 등 다양한 데이터 타입을 지원하며, 인터넷 상에서 이미 제공하고 있는 각종 서비스 및 정보를 그

대로 사용할 수 있도록 각 프로토콜, 데이터 포맷, 언어들에 대한 통합 인터페이스를 제공한다.[3]

본 논문에서는 이와같은 WWW의 특성을 이용하여 입후보자 프리젠테이션 서비스를 제공하는 멀티미디어 정보 검색 시스템을 설계하고 개발한다. 개발된 입후보자 프리젠테이션 서비스 정보 검색 시스템은 유권자들에게 입후보자 선택의 정확하고 공정한 정보를 선거 운동 기간중 언제든지 제공할 수 있고, 입후보자들간의 자질 및 정강, 정책 등을 비교 검색할 수 있다. 기존의 WWW에서 사용하는 정보의 저장 단위는 주로 화일 단위로 정보를 다루고 있기 때문에 많은 문제점이 있다. 따라서 다양한 형태의 입후보자 정보를 관리하기 위해 데이터베이스와 WWW 시스템과의 통합을 추진한다.

본 논문은 다음과 같이 구성된다. 2장에서는 WWW의 개념과 HTML(Hypertext Markup Language)에 관하여 기술한다. 3장에서는 WWW상에서 입후보자 프리젠테이션을 위한

정보 검색 시스템의 설계과정, 구현 환경 및 구현 접근 방법을 제시한다. 4장에서는 결론 및 향후 연구 방향에 관하여 기술한다.

## II. WWW(World Wide Web)

### 2.1 WWW의 개념

WWW 프로젝트는 1989년 CERN이라고 하는 고차원 에너지 물리학 연구소에서 전세계에 흩어져 있는 연구자들의 연구결과 및 아이디어를 공유할 수 있는 방법을 모색하기 위해 시작했다. 처음 계획할 무렵에는 WWW를 “하이퍼텍스트 프로젝트”라고 명명하였다. 하이퍼텍스트는 텍스트 또는 다른 정보들의 연결된 조각들로 구성된다. 각 정보 유니트들을 노드라 하고, 이 노드들은 링크(link)라고 하는 포인트에 의해 연결된다[2].

WWW 정보 검색 시스템은 사용자가 웹 브라우저(web browser)를 사용함으로써 일반 문자 정보뿐만 아니라 그래픽, 이미지, 오디오, 비디오 등과 같은 멀티미디어 정보를 사용자가 쉽고 일관된 방식으로 검색할 수 있도록 해주기 때문에 최근 정보 검색 시스템 관리자들의 관심의 대상이 되고있다. 특히 이러한 시스템에서 사용되는 웹브라우저와 표준화된 통신 프로토콜(HTTP: Hypertext Transfer Protocol)은 인간이 정보를 접근해 가는 방식과 유사한 하이퍼텍스트 개념을 도입함으로써 정보 검색 효율을 높여주었다.

그림 2.1은 WWW의 동작과정을 보여준다. 브라우저가 URL(Uniform Resource Locator)로 웹서버에게 파일을 요구하면 웹서버는 브라우저가 요구하는 HTML 문서를 찾아서 보내 주게되고 이때 그 문서에 끼어있을 수 있는 멀티미디어를 다루기 위해 메일을 확장한 MIME(Multipurpose Internet Mail Extension)을 이용하여 전달한다. HTML 문서에서는 서버가 다른 응용프로그램으로 정보를 요청하기 위한 여러가지 품들을 사용자에게 제공한다. 그리고 사용자가 정보를 요청하게 되면 서버는 입력을 받아서 그 정보를 CGI(Common Gateway Interface) 프로그램에게 넘겨준다.

CGI 프로그램은 입력을 받아서 자신이 가진 능력을 이용해서 정보를 제공하고 HTML 파일로 변환한 후에 다시 서버를 거쳐서 브라우저

에게 보낸다. 브라우저는 HTML을 해석해서 사용자가 요청한 정보를 디스플레이한다.

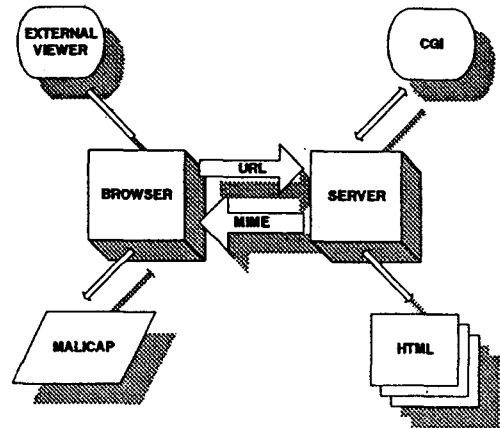


그림 2.1 WWW의 동작과정

### 2.2 HTML(HyperText Markup Language)

모자이크가 출현하기 전에는 인터넷은 일종의 다중 플랫폼 환경으로서 문서의 교환이 다소 힘들었다. 이러한 문서 교환의 문제를 해결하기 위한 방책으로 SGML(Standard Generalized Markup Language)이라는 특정 언어를 개발했다. SGML에서는 문서의 구성 요소에 초점을 맞추게 되므로, 정보 수령자는 원작자의 독점적인 선택에 따르지 않아도 된다[5].

SGML 문서를 사용하면, 윈도우즈의 크기를 조정하여 화면을 최대한 활용할 수 있고, 문서를 인쇄하여 레이아웃을 그대로 유지할 수 있는데, HTML은 전세계의 하이퍼텍스트를 전송하기 위한 간단한 형식의 독점적이지 않은 서식으로 SGML에서 파생한 것이다. HTML은 SGML처럼 데이터를 변환하고 포맷하는데 사용하는 일반적인 수단이다.

HTML은 아주 새로우며 언어 자체를 쉽게 익힐 수 있다. 웹 문서는 보통 HTML로 작성하고 “html”이나 “htm” 같은 확장자를 붙인다 [6,7]. 이러한 HTML 문서는 포맷 코드를 담고 있는 표준 아스키(ASCII) 파일에 불과한 것으로, 포맷 코드에는 텍스트 모양, 문서 제목, 문단, 항목 그리고 하이퍼링크와 같은 레이아웃에 대한 정보가 들어 있다.

HTML은 마크업 언어라고 부르는데, 텍스트를 나타내는 방법을 규정하는 구성 요소들

집합적으로 마크업(markup)이라고 한다. 마크업이란, 코드를 사용하여 웹 브라우저에 텍스트를 어떻게 나타낼 것인지 알려 주는 정보를 의미한다. 문서란 텍스트로 구성되어 있는데, 이 텍스트의 실마리를 바로 마크업에서 얻게 된다. 마크업은 보통 태그라고 칭하는 코드로서 효과를 지정하는데 사용하는 시작 코드와 마크업의 끝을 알리기 위해 앞부분에 슬래시를 붙인 종료 태그로 이루어져 있다. 시작 태그와 종료 태그를 구성요소라고 말하기도 하며, 그 사이에 텍스트나 자료를 기재할 때에는 구성 요소 전체를 묶어서 컨테이너(container)라고도 한다. 텍스트가 들어 있지 않고 종료 태그도 필요하지 않은 구성 요소들은 공간 구성 요소(empty element)라고도 부른다.

마크업의 기술 내용을 “문서 종류 정의(Document Type Definition)”, 또는 DTD라고 한다. HTML은 매우 제한적인 포맷 언어의 일종이다. HTML에 있는 마크업 구성 요소의 종류에는 머리글, 문단, 다양한 문자 강조 형식, 인라인 이미지, 하이퍼텍스트 링크, 목록, 미리 포맷되어 있는 텍스트 그리고 간단한 검색 도구 등이 있다.

### III. WWW상에서 입후보자 프리젠테이션을 위한 정보 검색 시스템

#### 3.1 입후보자 정보의 특성

제14대 총선(1992년)과 1995년 4대 지방선거 시 한국 갤럽 여론 조사 결과에서 유권자들이 선택한 입후보자 투표 결과를 보면 인물, 경력, 신념 등이 포함된 자질이 34.9%, 안정기여 13.2%, 정당 12.6%, 공약 11.1%, 여야 균형 및 변화기대 6.8%, 연고 및 안면도 4.5% 순으로 나타났다. 즉, 유권자들이 후보자들을 선택하는 기준은 입후보자들의 인품 및 경력의 비중이 매우 높다는 의미이다.

유권자들이 현재 선거법에 의거하여 입후보자들의 인품, 경력 및 공약 내용 등을 쉽게 접할 수 있도록 하기 위하여 주요 선거과정에서 모든 유권자들에게 필수적으로 제공되어야 할 입후보자 프리젠테이션 서비스의 정보 규격을 “공직 선거 및 선거 부정 방지법”에 근거하여 정하였다. 제작될 입후보자 프리젠테이션의 표현 기법 및 주요 내용은 표 3.1과 같다.

멀티미디어 표현 기법	표현 내용
Text	성명, 기호, 소속, 정당성, 경력
그래픽	사진, 정장, 정책
동영상 및 오디오	정견 발표
애니메이션	선거구호 및 기타홍보에 필요한 사항

<표 3.1> 입후보자 프리젠테이션의 표현기법 및 내용

#### 3.2 시스템 구성도 및 구현 환경

그림 3.1은 본 연구에서 개발된 전체 시스템의 구성도를 보여준다. 그림 3.1에서 보여주는 바와 같이 입후보자 프리젠테이션 정보중 텍스트 데이터에 해당하는 성명, 기호, 소속, 정당명, 경력 등은 데이터베이스로 구축하여 Informix DBMS를 통하여 관리하고, 그래픽, 동영상, 애니메이션 등과 같은 멀티미디어 표현 기법으로 표현된 정보들은 HTML 파일로 작성하여 파일 시스템으로 관리하였다. 기존의 WWW 정보 검색 서비스와는 다르게, WWW 서버와 데이터베이스 게이트웨이를 이용하여 데이터베이스화한 입후보자 정보를 WWW 서버가 빠르게 접근하여 이용할 수 있도록 하였다.

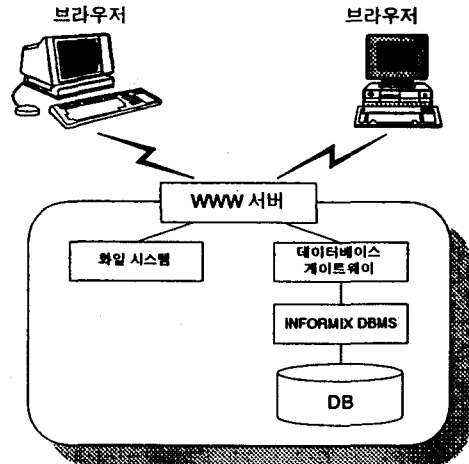


그림 3.1 시스템 구성도

또한 사용자는 같은 시스템에 있거나, 외부에서 인터넷을 통하여 접근하거나 관계없이

많이 사용되고 있는 웹 브라우저인 Mosaic, Netscape, HotJava 등을 이용하여 WWW서버에서 제공하는 서비스를 받을 수 있도록 설계되었다. 물론 사용자가 사용하는 플랫폼은 PC, Workstation, Macintosh에 관계없이 사용 가능하다.

본 연구의 구현은 SUN 4 Sparc 워크스테이션상에서 Informix DBMS를 이용하여 입후보자의 텍스트 정보를 관리하였고, 검색 인터페이스는 HTML 3.0을 이용하여 설계되었다. 또한 WWW에서 데이터베이스를 접근하고 검색하는데 데이터베이스 게이트웨이 부분은 WWW에서 제공하는 cgi-bin 프로그램과 ESQ/C를 사용하여 구현하였다.

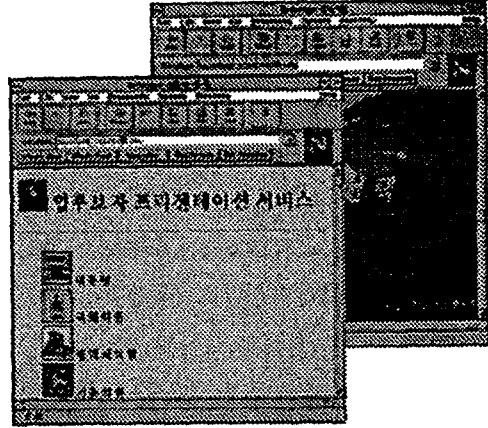


그림 3.2 입후보자 검색 시스템의 웹 홈페이지

### 3.3 입후보자 프리젠테이션 정보 검색 시스템

개발된 WWW상에서 입후보자 프리젠테이션 서비스를 위한 정보 검색 시스템은 대통령, 국회의원, 4대 지방선거에 관한 정보 서비스를 제공한다. 유권자들이 정확히 입후보자들의 특성 및 장단점을 파악할 수 있도록 가능한 많은 정보를 수집하였고 쉽고 편안하게 해당 정보를 접근할 수 있도록 구조화하였다. 또한 해당 선거별로 투표장에서 어떠한 방법으로 투표를 해야하는지를 알려주는 안내페이지를 별도로 설계하였다. 본 논문에서는 지면관계상 설계하여 구현한 다양한 형태의 입후보자 정보를 제공하지 못하고 홈페이지이외의 몇개의 페이지에 대해서만 보여주고 설명한다.

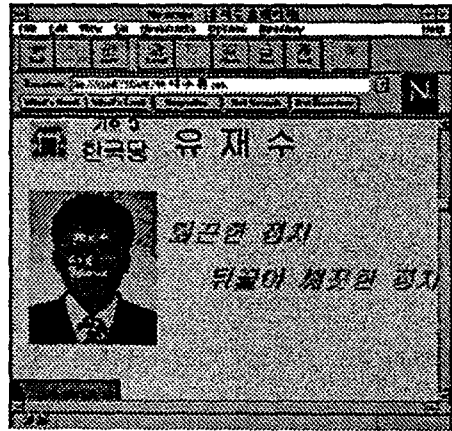


그림 3.3 대통령 후보 유재수에 관한 경력 정보

그림 3.2는 개발된 입후보자 프리젠테이션 서비스를 위한 WWW 정보 검색 시스템의 홈페이지를 보여준다. 그림 3.2에서 보여주는 바와같이 개발된 검색 시스템은 데이터베이스에 저장된 입후보자의 텍스트 데이터를 빠르게 검색하기 위하여 키워드 검색을 지원한다.

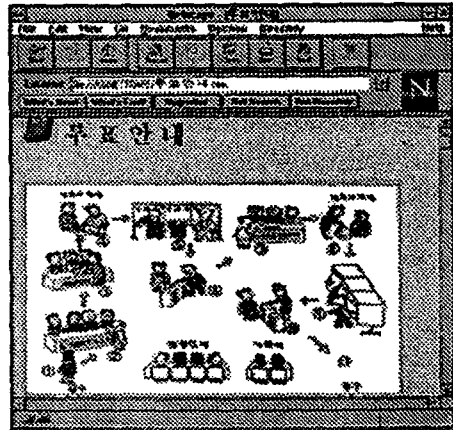


그림 3.4 4대지방선거에 관한 투표 요령

예를들어, 대통령 후보 유재수에 관한 경력 정보를 보고 싶을 경우, 출마형태란에 대통령을 입력하고 입후보자란에 유재수를 입력하면 대통령 후보 유재수에 관한 경력 정보를 포함하는 그림 3.3과 같은 페이지를 화면에 디스플레이한다. 또한 홈페이지에서 해당 선거별 투표장에서 투표하는 요령을 습득하기 위해 그림 3.4과 같은 페이지로 갈 수 있다.

그림 3.5은 입후보자들의 활동에 관한 비디오 정보를 보여준다.

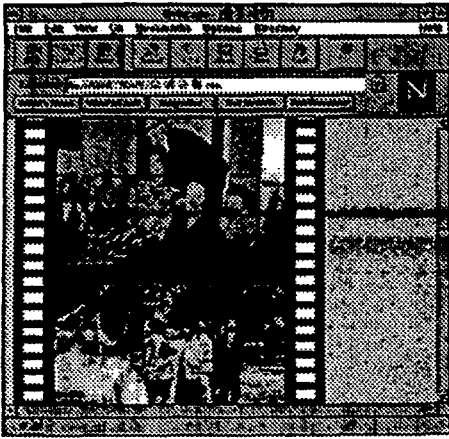


그림 3.5 입후보자들의 활동 사항

### 3.4 데이터베이스 게이트웨이의 구현

그림 3.6에서처럼 데이터베이스 게이트웨이는 WWW 서버측에서는 CGI 프로그램에 해당하며, 데이터베이스 측면에서는 클라이언트 프로그램이 되는 것이다.

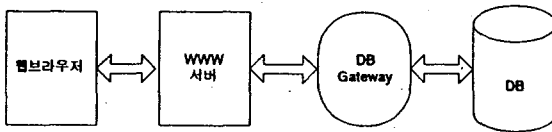


그림 3.6 데이터베이스 게이트웨이의 역할

데이터베이스 게이트웨이 부분은 주로 CGI 프로그래밍으로 이루어졌으며 cgi-bin 프로그램에서 사용자의 질의를 처리하기 위해 POST와 GET과 같은 두가지 방식이 있는데, 표준 입력 출력 방식으로 텍스트 형태의 질의를 받기 위해서는 GET보다는 POST 방식을 이용하는 것이 바람직하다. 따라서 사용자의 질의 내용은 cgi-bin 프로그램의 POST 방식에 의해 (name, value) 쌍으로 넘겨지며, 이후 cgi-bin 프로그램이 처리한 값은 다시 ESQL/C 프로그램을 이용하여 데이터베이스를 접근하고 원하는 값을 검색하게 된다.

데이터베이스 검색이 끝나고 해당하는 검색 데이터들이 데이터베이스로부터 넘어오면 이를 사용자의 브라우저에서 볼 수 있도록 처리해준다. 이런 일련의 과정들을 데이터베이스 게이트웨이 부분에서 처리해주고 있다.

### IV. 결론

본 논문에서는 유권자들에게 입후보자 프리

젼테이션 서비스를 제공하기 위해 WWW 상에서 멀티미디어 입후보자 정보 검색 시스템을 설계하고 개발하였다. 이를 위해 입후보자 프리젼테이션 서비스 내용 및 전체 시스템 요구 사항을 분석하였다. 입후보자 정보중 성명, 기호, 소속 정당등과 같은 텍스트 정보는 데이터베이스로 구축하여 관리하여 키워드를 통한 빠른 검색이 가능하도록 하였고 사진, 정견 발표, 선거 구호등과 같은 정보는 HTML 화일로 작성하여 관리하였다. 개발된 입후보자 프리젼테이션 서비스를 위한 멀티미디어 정보 검색 시스템은 유권자들에게 보다 정확하고 공정한 정보를 제공하여 입후보자 선택을 위한 선택지침을 제시한다.

향후 연구되어야 할 과제로는 다양한 형태의 질의 검색과 다양한 멀티미디어 데이터 지원 가능하도록 확장하는 것이다. 또한 유권자들에게 보다 쉽고 편하게 입후보자 정보를 검색할 수 있도록 Netscape나 HotJava와 같은 웹브라우저를 개발할 것이다.

### 참고문헌

- [1] I-Common Gateway Interface, [http://akebono.stanford.edu/yahoo/Computers/World\\_Wide\\_Web/CGL\\_Common\\_Gateway\\_Interface/](http://akebono.stanford.edu/yahoo/Computers/World_Wide_Web/CGL_Common_Gateway_Interface/)
- [2] Berners-Lee, Connolly, "Hypertext Markup Language: A Representation of Textual Information and Metainformation for Retrieval Interchange, Internet Working Draft CERN.
- [3] Kevin Huges, Entering the World-Wide Web: A guide to cyberspace, EIT, May 1994.
- [4] Jakob Nielson, Hypertext and Hypermedia, Academic Press, 1990.
- [5] Bryan, M., SGML : An Author's Guide, Addison Wesley, 1991.
- [6] Ian S. Graham, HTML Source Book, Wiley, 1995
- [7] Raggett, Dave, "HTML 3.0 Document Type Definition," Geneva, 1995.
- [8] Daniel Connolly, Atrium, "HyperText Markup Language Internet Draft IIR Working Group," CERN, June 1993.