

# 사용자 적응형 에이전트 기반의 전자도서관 시스템

소영준, 김인태, 김범수, 박영택  
승실대학교 정보과학대학 컴퓨터학과

## Digital Library System based on User Adaptive Agents

Young-Jun So, In-Tae Kim, Bum-Su Kim, Young-Tack Park

{(yjs, powerkit, bskim, park)}@multi.soongsil.ac.kr

Dept. of Computer Science, Soongsil Univ.

### 요약

본 논문에서 구현하고자 하는 에이전트 기반 전자도서관 시스템은 웹을 통해 도서정보를 검색하는 사용자에 대한 도서 정보 관심도를 모니터 에이전트에 의해 추출하게 하고 해당 정보를 대상으로 학습 과정을 거쳐 사용자별 프로파일을 구축할 수 있게 한다. 이 과정에서 모니터 에이전트에서 추출되는 사용자 history와 학습 예제 정보는 사용자가 행하는 다양한 작업의 결과 차별화된 중요도가 적용된다. 학습 에이전트를 통해 구축된 사용자 관심 정보는 각 사용자에게 세부 DB 추천 모듈과 관심 정보를 푸쉬해주는 에이전트에 적용된다. 본 시스템에서는 전자 도서관에 가입하는 사용자들의 공통된 그룹정보를 관리하여 이를 통해 동일 그룹내의 사용자들 사이의 정보 이용이 가능하게 하였다. 본 논문에서는 개인과 그룹을 대상으로 하는 모니터 에이전트와 학습에이전트를 설정하여 도서관을 이용하는 사용자들에게 정확하고 신속한 정보를 제공하는 것을 목적으로 한다.

### 1 서론

전자 도서관을 이용하는 사용자들이 늘어나고 도서관에서 제공되는 정보와 서비스가 점점 다양해지고 증가함에 따라 기존의 검색 시스템을 이용하여 자신이 원하는 정보를 정확하게 찾기란 쉬운 일이 아니다. 기존의 도서관 시스템은 모든 사용자에게 공통적인 서비스를 제공함으로써 특정 분야에 관심을 둔 사용자는 자신의 관심정보를 매번 동일한 여러 과정을 거쳐야 얻을 수 있는 불편이 있다. 이와 같은 문제점들은 전자도서관을 사용하는 데 있어서 검색 효율을 떨어뜨리게 할뿐만 아니라 새로운 관심 정보가 등록되었다 하더라도 이를 필요한 때에 받아 볼 수 없게 한다. 본 논문에서는 이러한 사용자에게의 무차별적인 서비스가 아닌 각 사용자마다의 개별적인 관심정보 제공의 적응형 서비스를 위해 지능형 에이전트 기반의 전자도서관 시스템을 제안한다.

지능형 에이전트는 사용자가 관심을 가지는 정보의 주제를 파악하여 관련된 정보를 수집, 제공하는 목적을 갖고있다[Marko]. 그러므로 본 논문에서는 SGML 전자도서관에서 사용자가 수행한 작업들 중에서 모니터 에이전트를 통해 자동으로 관심 정보를 추출한다. 추출된 사용자의 관심 정보는 특정 관심영역을 주제로 하여 수집되며 학습 에이전트에 의해 프로파일의 형태로 저장된다. 해당 프로파일은 도서관의 검색 속성이자 SGML문서의 중요 속성에 따른 다중 프로파일로 구성된다.

각각의 프로파일은 각 사용자가 관심 있어 한 정보를 구성하고 있는 중요 키워드를 기반으로 구축되며 사용자가 전자도서관을 이용하면서 행하는 검색 작업과 개별적인 관심문서 푸쉬 작업등에 이용된다.

본 연구에서는 각각의 사용자가 에이전트를 통해 자신의 검색정보는 물론 그룹 정보를 토대로 관심분야에 대해 도움을 받을 수 있고, 학습된 결과를 적용한 푸쉬 서비스를 받을 수 있는 에이전트 시스템을 제안한다. 2장에서는 관련 전자 도서관 에이전트에 관해 개괄적으로 다루고, 3장에서는 본 연구에서 적용하는 전자도서관 에이전트의 사용자별 또는 그룹별 관심도 추출기능에 대하여 설명하며, 4장에서는 학습 에이전트를 통한 다중 프로파일 구축과 응용 에이전트의 기능에 대해 설명하고, 5장에서는 응용 에이전트에 대한 실험 결과와 향후 연구방향에 대하여 기술한다.

### 2. 관련 연구

웹 상의 전자도서관이 갖는 사용자상의 편리함과 운영상의 이점으로 인해 전자도서관의 구축 건수가 늘어가고 있지만 단순한 검색엔진으로는 방대한 자료들 중에서 사용자가 원하는 정보를 찾는 데 어려움이 많아졌다. 이에 전자도서관의 기능을 에이전트와 연계하여 제공함으로써 자동으로 사용자의 관심 정보를 추출, 이를 대신 검색하여 제공해주는 지능형 에이전트 기반의 전자도서관에 대한 연구가 이루어지고 있는

주제이나, 이러한 전자도서관 에이전트의 목적은 추출된 사용자 관심 자료를 정확하고 신속하게 검색하여 사용자에게 전달하는데 있다. 이러한 전자도서관 에이전트에서 기본적으로 주목하고 있는 부분은 각 사용자의 관심도 추출, 추출된 관심도를 기반으로 하는 관심 정보 검색기능과 검색시 관련 세부 DB 추천 기능 등의 응용분야이다. 위 작업을 통하여 전자도서관 에이전트는 사용자에게 보다 정확한 관심 정보를 제공한다. 현재 세팅된 에이전트를 적용한 전자도서관들의 공통적인 사항은 개별적인 사용자의 프로파일을 구축하여 각각의 사용자에게 맞는 적용화된 서비스를 제공한다는 것이다.

2.1 UMDL(University Michigan Digital Library)

UMDL은 미국의 미시간대학의 전자도서관 프로젝트로, 도서관의 구축을 위해 세 가지 소프트웨어 에이전트를 두고 있다. 이 중에 사용자 인터페이스 에이전트는 각각의 사용자가 요구하는 정보의 종류를 파악하여 해당하는 정보를 제공하기 위한 목적을 갖는 에이전트이다 [Michigan]. 사용자가 검색한 히스토리 정보를 각 사용자마다 저장하여 두었다가 다음 로그인 하였을 경우 해당 정보 또는 수정된 정보를 재 사용할 수 있게 하였다. 그러나 사용자의 관심도를 구체화하는 프로파일의 생성작업은 이루어지지 않는다.

2.2 Stanford University

미국의 스탠포드 대학의 전자도서관 프로젝트는 네트워크 상에 흩어져있는 여러 정보들을 대상으로 통합적인 서비스를 제공하는데 있어서 이용자에게 필요한 도서관 서비스와 적합한 정보를 검색할 수 있는 기술을 개발한다. 예로서 GLOSS(Glossary-of-Servers Server) 서비스가 있는데 이것은 사용자의 특정 질의를 만족시키는 정보저장소의 위치를 지시해 주는 metadata를 유지하는 것이다[Stanford]. 이를 통해 사용자는 자신이 찾고자 하는 정보를 검색할 여러 정보DB중에 시스템이 추천해주는 특정DB를 선택할 수 있게 된다. 하지만 단순한 검색 시에 주어지는 질의어로 인한 사용자에게 맞는 DB추천 기능일 뿐 사용자의 종합적인 관심정보를 학습하여 이를 제공하는 기능은 추가되지 않았다.

3. 지능형 에이전트 기반 전자도서관 시스템

본 논문에서 구현한 지능형 에이전트 기반 전자도서관 시스템의 전체적인 구성은 크게 세 부분으로 구성되어 있다. 첫째, 전자도서관을 검색하는 각각의 사용자와 그들이 속한 그룹의 관심 정보를 추출해내는 모니터 에이전트로 주로 사용자가 행한 검색 키워드와 선택한 검색 결과 그리고 검색 히스토리 등의 정보를 저장하여 관심 정보를 추출하는데 이용한다. 둘째, 추출된 정보를 기반으로 다중 프로파일을 구축하는 학습 에이전트로 단순한 검색 키워드 프로파일이나 아닌 검색된 정보의 저자, 제목 등의 여러 속성을 대상으로 여러 개의 프로파일을 구축한다. 셋째, 해당 프로파일을 적용하여 검색작업에서 세부 DB 추천 및 질의어 확장 등의 기능을 수행하는 응용 에이전트로 사용자가 특정 질의어를 입력하였을 경우 해당 정보가 저장되어 있는 세부 DB를 추천하거나 제목, 저자, 요약 등의 정보를 확장하여 검색하는 기능을 제공한다.

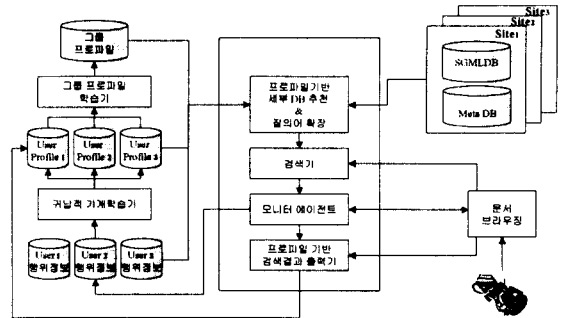


그림 1 에이전트 기반 전자도서관 시스템 구성도

3.1 모니터 에이전트

모니터 에이전트는 학습을 통한 사용자 프로파일 생성을 위한 학습 자원을 만들어내는 기능을 수행하는 부분으로 그 기능과 구조는 다음과 같다. 첫째, 현재 사용자가 전자도서관 시스템에서 수행하는 모든 작업들을 저장한다. 이는 검색 작업 시 해당 정보 저장과 검색 결과 중에서 사용자가 선택한 정보 등으로 구분되어 관리된다. 사용자가 검색을 위해 입력한 질의어들은 본 시스템 상에서는 저자, 제목, 요약의 정보 별로 구분되어 관리된다. 선택한 검색 결과에서도 해당 속성을 추출하기 위해 본 시스템에서는 자체적으로 문서의 메타데이터를 설정하여 사용하였다. 이로써 각각의 속성에 따른 프로파일을 생성하는 기초자료로 사용된다. 또한 사용자의 검색 히스토리 정보는 갱신된 정보 확인을 위해 이전의 질의어와 같은 형식으로 입력하려 할 경우 반복된 질의 작업이 필요 없이 히스토리 정보중 해당 정보를 선택함으로써 수행하도록 해준다. 둘째, 사용자의 사용자 단위가 아닌 특정 사용자들로 구성된 그룹 정보관리이다. 이는 사용자가 속해있는 그룹의 정보들 기반으로 사용자에게 관련된 정보를 제공하기 위한 것으로 그룹의 분야가 사용자의 관심 영역을 나타낸다. 사용자는 자신의 관심영역을 먼저 생성하거나 이미 생성된 분야에 참여함으로써 어느 정도 구체적인 관심도를 지정해 놓고 해당 그룹의 미리 구축되어 있는 공통적인 관심 정보를 제공받을 수 있다. 이는 사용자가 자신의 학습된 프로파일이 존재하지 않았을 경우 관심 정보를 서비스 받을 수 있다는 이점이 있다.

3.2 학습 에이전트

학습 에이전트는 모니터 에이전트가 보내오는 사용자 관심 정보에 대해서 사용자가 생성 또는 참여한 관심 그룹의 주체로 되어있는 정보들의 중요 키워드에 대해 C4.5의 귀납적 기계학습 시스템을 통해서 사용자 프로파일을 생성한다. 모니터 에이전트에서 추출된 사용자 행위 정보는 먼저 각 속성별로 구분되어 별도의 프로파일을 생성하고 생성된 사용자 프로파일은 사용자가 속하는 그룹의 프로파일을 생성하는데 이용된다. 프로파일은 검색 시 입력된 속성을 중심으로 다중 프로파일을 생성하며 이는 각 속성에 입력된 질의어를 확장하는데 적용된다. 구축된 프로파일은 또한 시스템의 세부 DB 추천 기능에도 이용된다. 본 논문에서 각 프로파일을 구성하는 키워드 학습 방법으로 TF, DF방식

과 키워드 백터를 추출한 뒤 C4.5학습 시스템을 이용하여 결정 트리를 구한 후 생성된 룰을 이용하는 방식을 적용하였다[Quinlan]. 다음의 그림 2. 이러한 다중 프로파일의 생성되는 과정을 나타낸 것이다.

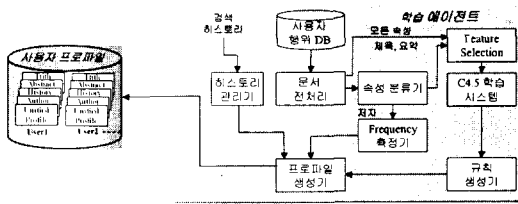


그림 2. 학습 에이전트 구조

### 3.3 그룹 프로파일을 이용한 사용자 관심정보

본 논문에서는 사용자가 자신의 관심 영역을 모든 사용자에게 공유하여 관리하는 그룹을 생성하는 기능을 제공한다. 해당 그룹은 생성한 사용자에게 의해 관리되며 이에 참여한 사용자들은 자신이 해당 그룹을 통해 검색한 행위 정보들을 제공하는 것과 동시에 그룹의 관심 정보를 제공받게 된다. 이러한 그룹 프로파일로 인해 사용자가 얻게 되는 이점은 다음의 두 가지로 요약될 수 있다. 첫 째, 사용자의 프로파일이 생성되지 않은 가입 초의 시기에도 관심영역에 대한 그룹에 참여함으로써 관심정보에 대한 서비스를 제공받게 된다. 둘째, 많은 사용자가 참여하고 있는 동일한 그룹을 나타내는 중요 키워드를 단 한사람의 검색 행위에서 보다 더욱 자세하게 추출할 수 있어 관심영역의 정확도를 높일 수 있다.

### 3.4 프로파일을 적용한 응용 에이전트

본 논문에서 적용되는 학습된 사용자 프로파일의 이용 기능은 크게 세 가지로 구분할 수 있다. 첫 째, 여러 DB들로 구성된 전자도서관에서 검색 행위를 나타내는 DB선택 작업을 수행하여 주는 세부 DB 추천 기능, 둘째, 사용자가 제목, 저자, 요약 등의 속성에 입력한 질의어를 프로파일과 비교 확장하여 주는 기능, 셋 째, 검색 결과로 출력되는 정보들과 사용자 프로파일의 유사도를 측정하여 가장 높은 수치를 갖는 결과를 순서대로 사용자에게 출력해주는 기능이 그것이다.

## 4. 전자도서관의 사용자 인터페이스

본 논문에서 제안한 전자도서관 에이전트의 기능을 이용하기 위해 사용자는 먼저 시스템에 등록을 하고 검색작업을 수행한다. 이는 사용자의 관심도에 맞게 검색 정보를 제공하기 위해서 사용자 정보 등록과 동시에 사용자 관심 DB와 본 논문에서 정의한 SGML DTD에 해당하는 사용자 프로파일 DB 및 검색 History DB가 생성이 된다. 사용자의 검색 한 건과에 의해서 사용자 관심 DB에 해당하는 Keyword와 검색 결과에서 찾아진 문서 및 전체 문서를 다운로드 시 각각 구분된 가중치를 갖는 DB에 저장을 하게된다. 이 데이터를 가지고 위의 설명한 단계들을 통해 사용자 프로파일과 그룹 프로파일을 생성하는 순서로

진행되었다. 그리고 Keyword History DB는 사용자로 하여금 이전에 검색한 정보를 가지고 있어 재 검색을 할 때 검색 입력부분에 사용자가 입력하지 않고 자동으로 입력하여 사용자의 편리성을 제공을 하게 된다. 다음 그림은 전자도서관 시스템의 인터페이스 화면이다.

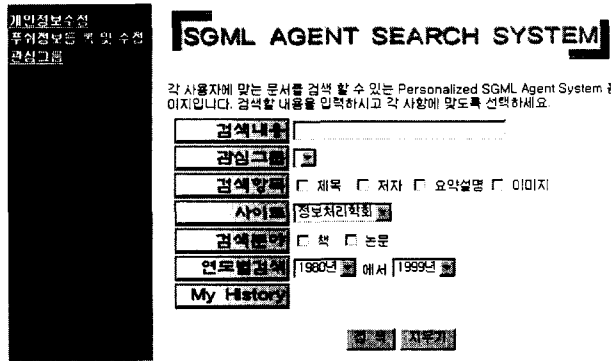


그림 3 전자도서관 에이전트의 인터페이스 화면

## 4. 결론 및 향후 연구과제

본 논문은 에이전트기능을 적용한 전자도서관 시스템을 소개하였다. 사용자는 모니터 에이전트를 통하여 전자도서관 정보에 대한 관심도를 추출되게 하고 학습 시스템에 의하여 관심분야에 관한 구체적인 키워드 정보를 갖게 된다. 구축된 다중 프로파일은 사용자가 전자도서관내의 특정 정보를 찾기 위해서 검색엔진을 사용하는 행위에 자문을 줄 수 있는 기능으로 확장될 것이며, 정보가 저장되어 있는 최최의 DB를 추천 받을 수 있게 하였다. 이로써 기존의 전자도서관 사용자가 정보 검색 시 갖는 불편을 해결할 수 있을 것이며 사용자에게 맞는 정보 서비스를 받게 될 것이다. 본 시스템에 적용되는 에이전트 시스템은 전자도서관 뿐만 아니라 웹 상에서 제공되는 모든 서비스 시스템에 응용이 가능할 것이다.

## 5. 참고문헌

[Marko] *Marko Balabanovic and Yoav Shoham*, "Learning Information Retrieval Agents: Experiments with Automated Web Browsing", AAAI Spring Symposium on Information Gathering, Stanford, CA, March 1995.

[Stanford] *L. Gravano, H. Garcia-Molina, and A. Tomasic*, "The Effectiveness of GIOSS for the Text-Database Discovery Problem," Proc. SIGMod Conf., ACM Press, New York, 1994.

[Michigan] *B. Schatz et al.*, "Interactive Term Suggestion for Users of Digital Libraries: Using Subject Thesauri and Co-Occurrence Lists for Information Retrieval," Proc. First ACM Int'l Conf. Digital Libraries, ACM Press, New York, 1996, pp. 126-133.

[Quinlan] *J. R. Quinlan*, "C4.5 Programmes for Machine Learning", San Mateo, CA: Morgan, Kaufman, 1993.