

하이퍼텍스트 정보검색에 관한 연구동향

이 영 자*

< 목 차 >

- | | |
|--------------------------------------|--|
| I. 서 언 | 4. 정보검색 보조도구들의 개발에 대한 연구 |
| II. 전통적인 정보검색모델의 하이퍼텍스트에의 통합에 관한 연구 | III. 온라인목록 검색을 위한 하이퍼텍스트 기법의 적용에 관한 연구 |
| 1. 일반적 논의 | IV. 인공지능의 하이퍼 미디어에의 적용에 대한 연구 |
| 2. 전통적인 정보검색기법들의 하이퍼텍스트에의 통합 | V. 결 언 |
| 3. 이용자-시스템간의 상호작용성 향상을 위한 하이퍼텍스트의 사용 | |

I. 서 언

정보검색이란 컴퓨터시스템과 이용자간의 상호작용과정에 그 본질을 두고 있다고 말할 수 있다. 이 상호작용성은 정보요구를 충족시키기 위하여 이용자로 하여금 한 번의 탐색기간 중에 한 탐색전략에서 다른 탐색전략으로 여러번 옮길 수 있도록 허용하며, 또한 원하는 결과에 도달할 때까지 다양한 피드백 기법들을 사용할 수 있게 한다. 시스템과 이용자간의 상호작용은 이용자의 정보요구의 목적 즉, 수행해야 할 과제의 성격에 따라, 선택된 탐색기법에 따라 다양한 양상으로 이루어질 수 있다. 즉, 알려진 한 항목들을 탐색할 경우에 Boolean 탐색기법은 최선이 될 수 있고, 같은 토픽에 관한 많은 항목들을 찾아야 할 경우에는 브라우징 기법이, 그리고 이미 제시된 결과를 계속 추적하려고 할때는 링크들의 사용이 권장된다.

* 경북대학교 문헌정보학과 교수

전통적인 정보검색기법들은 주로 질의어의 입력에 근거하며 탐색결과들은 질의어에 적합한 유용한 여러 항목들을 찾아볼 수 있게 하는 탐색기법(searching)이라 할 수 있다. 한편, 이 탐색된 여러항목 가운데 가장 적합한 항목을 중심으로 그와 관련된 혹은 연결된 항목들을 선택할 수 있는 기법은 브라우징(browsing)기법이라고 불린다.

브라우징기법을 사용하는데는 브라우징 해야할 공간이 너무 방대할 때 이용자측에서 정보탐색의 방향을 상실할 수 있는 단점이 야기된다. 따라서 이 문제점을 극복하기 위한 다양한 전략들도 연구, 개발되어오고 있으며 질의어에 의한 탐색기법을 브라우징기법과 통합하는 기법이 두드러진 한 연구영역을 이루어오고 있다. 하이퍼텍스트라는 새로운 기술이 초기에는 주로 상호작용성이 강한 교육도구들이나 기술편람, 사전, 백과사전 등의 전자출판 등에 많이 적용되어 오다가 최근에 이르러 정보검색시스템에 대한 사용자인터페이스로 개발되거나 전통적인 정보검색시스템과의 통합에 적용되어오고있다.

본 연구의 목적은 정보검색분야에 있어서의 하이퍼텍스트(혹은 하이퍼미디어)의 통합 내지 적용에 대한 최근의 연구 동향을 검토함으로써 하이퍼텍스트 정보검색시스템의 연구 개발의 방법들과 현황들을 파악하고자 하는데 있다. 30 여편의 문헌검토로써 망라적이고 철저한 동향과악이 이루어질 수는 없을 것이지만 검토된 문헌들을 통하여 몇가지 연구영역들의 범주를 구별하려고 시도한 것이며 이들은 서로서로 배타적인 것이 될 수 없고 어떤 문헌은 하나 이상의 범주에 속할 수도 있을 것이다.

그래서

첫째, 질의어에 의한 고전적 정보검색기법과 하이퍼텍스트의 브라우징기법의 통합에 관한 연구.

둘째, OPACs에 대한 인터페이스로서의 하이퍼텍스트시스템에 관한 연구,

셋째, 인공지능검색기법과 하이퍼텍스트기법의 통합에 관한 연구. 로 나누어 각각에 대한 연구동향이 검토될 것이다.

II. 전통적인 정보검색모델의 하이퍼텍스트에의 통합에 관한 연구

1. 일반적 논의

정보검색에 하이퍼텍스트기법을 통합하기 위한 연구는 대규모 도서관의 자료집성에서, 알고있지 않은 전문영역의 자료를 찾으려고 할 때 사서의 도움이나 목록시스템의 도움을 받아(혹은 받지않거나) 질의어로 출발하는 것이 훨씬 더 용이하다는 점과 규모가 훨씬 작은 도서관자료에서 이용자가 자기에게 익숙한 특정분야의 자료를 찾을때 서가를 열람하는것(브라우징하는것)이 더 쉽고 효율적이라는 점을 결합함으로써 새로운 검색전략을 개발하려는 의도에 그 초점을 두고있다. 한마디로 요약하면 질의어탐색을 거쳐서 이용자가 브라우징할 영역을 찾아낼 수 있음으로써 대규모 자료집성의 탐색이 보다 적절하고 효율적일 수 있으리라는 가정을 입증하려는 시도들이라고 볼 수 있다.

전통적인 질의어-기반의 탐색은 Boolean연산자에 근거하여 질의어와 레코드의 완전일치에 의하여 탐색결과를 얻는 경우와 질의어를 구성하는 용어들의 전부 혹은 그중의 몇개를 포함하는 레코드들을 출력하는 최적일치 혹은 부분일치기법을 포함한다. 이 부분일치기법은 용어에 가중치를 부여하는 벡터공간기법과 확률에 근거하는 기법이 있을 수 있다. 이 각 각의 전통적인 검색기법들을 하이퍼텍스트기법에 적용 내지 통합하려는 시도가 연구의 한 영역을 이루어오고 있다. 한편으로는 하이퍼텍스트-기반의 디소러스의 개발과 분류체계를 하이퍼텍스트에 적용하려는 연구도 아울러 이루어져오고 있다.

2. 전통적인 정보검색기법들의 하이퍼텍스트에의 통합

전통적인 정보검색모델을 하이퍼텍스트시스템에 통합함으로써 보다 효율적인 정

보검색을 수행할 수 있도록 하기 위한 연구들을 연대순으로 논의해 보면 다음과 같다.

1987년에 보베이(J. D. Bovey)와 브라운(P. S. Brown)은 기존의 질의어-기반의 불리언검색기법에 대한 프론트 엔드(front-end) 시스템으로서의 하이퍼텍스트 시스템을 통합하는 문제에 대한 연구를 발표하였다.¹⁾

이 연구는 정보검색시스템에서 문헌들을 상호작용적으로 디스플레이하는 설비의 적합성을 논한다. 상호작용적으로 문헌들을 디스플레이할 수 있는 한가지 가능성은 데이터베이스내의 레코드들의 정보가 버튼을 이용하는 계층적방법으로 표현될 수 있도록 GUIDE포맷에 의하여 기록할 수 있다는 점에 있으며 GUIDE를 사용함으로써 기존의 데이터베이스를 수정하지 않고 그로부터 추출된 레코드들을 디스플레이할 수 있다. 이것을 실험하기 위하여 보베이가 설계하고 구축한 기존의 정보검색시스템인 PRISM에 대한 프론트-엔드 시스템으로서 하이퍼텍스트 개발도구인 GUIDE를 이용한 PRIDE라는 원형검색시스템을 PERQ-2 그래픽 워크스테이션에서 가동하였다. PRISM의 bq (Boolean query) 명령어, pr (print) 등의 기본적인 명령어들을 PRIDE에서 사용하지만 원형시스템으로 구축되었기 때문에 모든 명령어들을 다 포함시키지는 않았다. PRIDE는 초기의 하이퍼텍스트시스템으로 탐구되었기 때문에 향상의 여지를 많이 가지고 있다. 그 중에서도 전통적인 질의어 명령에 근거하고 있는 PRIDE가 레코드들의 적합성을 평가할 수 있는 질의어방식을 통합하는 것을 지적할 수 있다.

1988년에 프리세(Mark E. Frisse)는 그래프상에서 최적의 브라우징의 출발점을 선정하기 위하여 가중치기법을 적용하는 질의어 처리에 대한 연구를 발표하였다.²⁾

-
- 1) J. D. Bovey and P. S. Brown, "Interactive Document Display and Its Use in Information Retrieval," *Journal of Documentation*, Vol. 43, No. 2(June 1987), pp. 125-137.
 - 2) Mark E. Frisse, "Searching for Information in a Hypertext Medical Handbook," *Comm. of the ACM*, Vol. 31, No. 7(July, 1988), pp. 880-886.

이 논문에서 저지는 하이퍼텍스트시스템을 위한 색인과 정보검색 방법을 설계함에 있어서 하이퍼텍스트는, 노드의 상대적인 자율성을 강조할 때는 이용자의 질의어와 데이터베이스내의 노드를 대조하여 적합한 노드를 검색하는 정보검색과 동등한 것이 되고, 만약 노드들간의 의미적 연결구조를 강조한다면 방향이 인도되어 있는 그래프화된 네트워크구조를 브라우징할 수 있게 되는데 전자를 패턴일치를 강조하는 소문헌접근방식(small document approach)으로, 후자는 브라우징을 강조하는 그래프형단방식(graph traversal approach)로 명명하고 있다. 저자는 이 논문에서 하이퍼텍스트를 위한 질의어 원칙들을 논했는데 그래프의 브라우징을 위하여 한 개 혹은 그 이상의 최적의 출발점을 이용자에게 제공함으로써 이용자에게 반응해야하며 그 출발점은 노드들의 내재적요소(intrinsic component)와 외재적요소(extrinsic component)의 기중치값의 계산에 의하여 확인될 수 있다고 하였다. 그리고 저지는 사용자피드백을 가능하게 하는 하이퍼텍스트검색을 위하여 신뢰망(belief network)와 결합성(connectionist) 방법의 잠재성을 논하였다.

1989년에 프리세와 쿠진즈(Steve B. Cousins)는 프리세가³⁾ 언급한 신뢰망과 결합성 방법의 잠재성을 연구한 결과를 논문으로 발표하였다.⁴⁾

저자들은 기록된 의학정보의 효과적인 이용을 위한 <Dynamic Medical Handbook>프로젝트에서 사용한 정보단위, 링크, 문헌공간, 정보공간 및 네비게이션 등의 용어를 명확히 정의하였다. 그리고 효과적인 정보접근의 두가지 방법으로서 첫째, 부분접근방법(local access method)은 버튼-기반의 브라우징에 의한 하이퍼텍스트 네비게이션 방법으로, 둘째, 전체접근방법(global access method)은 문헌공간 외부의 데이터구조인 색인공간에 의한 네비게이션 방법으로 정의하였다. 저자들은 문헌공간과 색인공간 둘 다가 평면적인 전통적 검색시스템에서는 한

3) Frisse, 1988.

4) Mark E. Frisse and Steve B. Cousins, "Information Retrieval from Hypertext: Update on the Dynamic Medical Handbook Project," *Hypertext '89-Proceedings*, (Nov. 1989), pp. 199-212.

문헌이나 한 색인용어의 유용성에 대한 정보를 제공해주지 않는, 조건적 독립을 지양하여 평면적이 아닌 계층문헌공간과 계층색인공간을 이용하여 구조화된 의학편람을 효과적으로 탐색하기 위한 신뢰망을 사용하는 추론확률기법을 연구하였다.

같은 해에 크로프트(W. Bruce Croft)와 터틀(Howard Turtle)은 베이지언식 추론망(Baysian inference network)을 구축하여 전통적인 질의어에 의한 부분일치검색을 할 수 있는 하이퍼텍스트 환경을 이룩하는 연구를 발표하였다.⁵⁾

이 논문은 하이퍼텍스트 데이터베이스를 베이지언 추론망으로 구축할 수 있으며 이것이 어떻게 질의어에 답하기 위하여 이용되는지를 설명한다. 데이터베이스를 구성하는 노드들의 내용은 특정의 개념과 그에 관련되는 개념들을 연관시킴으로써 표현되는데 이 연관은 수직업으로 용어를 선택하거나 통계적 언어처리, 자연언어처리 기법에 의거하는 등 다양한 방법을 사용하여 확립될 수 있다. 그리고 데이터베이스의 노드들간의 링크도 여러가지 유형을 갖는데 통계적인 최단근접(nearest neighbor)에 의하여 유도되는 노드들간의 링크, 텍스트와 이용간의 서지사항의 수로부터 유도되는 링크, 계층적구조상에서 표현된 노드들간의 링크들을 들 수 있다. 이 연구의 의의는 전통적인 정보검색시스템에서 취급하는 대부분의 질의는 미리 수집하여 구축된 색인망을 통하여 대답을 갖게 되지만 '보다 더 큰, 보다 더 적은'과 같은 비교평가적인 질의어나 다양한 문헌변수들을 포함하는 질의어의 경우 그 대답을 찾기 어렵기 때문에 연역추론과 비연역추론 등이 실제시스템에 통합될 필요가 있음을 제안한 데 있다. 즉 하이퍼텍스트 환경에서 추론망기법을 통합하여 검색모형을 확장하려는 시도에서 이 연구의 의의를 찾을 수 있다. 그리고 저자들은 이 방법에 의한 검색효율을 평가하기 위하여 정도율과 재현율이라는 측정방법을 적용하였다.

크로프트와 터틀은 1993년에 1989년의(각주 5의) 연구를 확장한 논문을 발표하

5) W. Bruce Croft and Howard Turtle, "A Retrieval Model for Incorporating Hypertext Links," *Hypertext '89 Proceedings*. (Nov. 1989), pp. 213-234.

였다.⁶⁾

연구자들은 이 논문에서 하이퍼텍스트 데이터베이스내의 링크들이 하이퍼텍스트 데이터베이스와 전통적인 정보검색시스템의 문헌데이터베이스를 구분하는 기능을 한다고 보며 추론망에 근거한 확률적 검색모델을 이 링크들에 통합할 수 있는 방안이 하이퍼텍스트 정보검색모델개발에 유용할 수 있다고 피력하고 이 새로운 모델을 평가하기 위한 실험을 제시한다. 즉 서지정보 링크(citation link)와 유사성-기반의 링크들에 연결된 문헌망에 대한 파급활성화(spreading activation)는 하이퍼텍스트를 위한 많은 가능성을 가진 전략으로 지적한다. 그리고 파급활성화 전략과 본 논문에서 제시하고 있는 베이지언 추론망에 기반을 둔 탐색전략을 비교 실험했다. 실험결과는 베이지언 추론망에 기반을 둔 하이퍼텍스트 검색은 파급활성화 전략만큼 효과적임을 보여준다. 연구자들은 복잡한 구조를 가지고 있는 멀티미디어 도큐먼트들을 지원하기 위한 더 깊은 연구가 현재 진행되고 있다고 지적한다.

1993년에 루카레라(D. Lucallela)와 잔지(A. Zanzi)는 정보검색의 탐색기능(searching)과 하이퍼텍스트의 브라우징(browsing)기능을 통합하여 멀티미디어 노드들을 확률추론을 이용하여 탐색하며 적합성 순서에 따라 레코드들이 출력되는 하이퍼텍스트로부터의 정보검색에 대한 연구를 발표하였다.⁷⁾

이 논문에서 제시하는 검색은 이용자가 하이퍼텍스트망을 브라우징하는 과정에서 이루어지거나, 시스템이 지식베이스로서의 하이퍼텍스트망을 활용하기 위하여 탐색기능을 수행할 수 있는 추론과정에서 이루어진다. 또한 이 논문은 질의어기반과 브라우징기반의 검색방법을 결합하는 원형적 모형과 구도(architecture)를 제시하는데 모형은 확률추론에 근거를 두고 있으며, 하이퍼미디어장서는 추론망으로서의

6) W. Bruce Croft and Howard Turtle, "Retrieval Strategies for Hypertext," *Information Processing & Management*, Vol. 29, No. 3(1993), pp. 313-324.

7) D. Lucallela and A. Zanzi, "Information Retrieval from Hypertext: An Approach Using Plausible Inference," *Information Processing & Management*. Vol. 29, No. 3(1993), pp. 299-311.

역할을 한다. 이 모형에서 링크들은 두개의 연결된 노드들 사이에 존재하는 관계의 유형에 따라 이름이 붙여지고 그 관계의 강도(strength)를 나타내기 위하여 링크에 가중치가 부여된다. 이것은 하이퍼텍스트의 상태를 수동에서 능동적인 상태로 수정하는 방법이라고 연구자들은 피력한다. 이용자의 관점에서 볼 때 이 시스템의 탐색모듈은 브라우징을 위한 최적의 선택점을 얻거나 혹은 의미적으로 관련된 노드들의 소재를 확인할 수 있도록 하는 여과능력(filtering capability)으로 지각된다. 이들이 실험결과를 제시함으로써 이 모델의 성능에 대한 어떤 통찰을 주고있다.

던랩(M. D. Dunlap)과 리즈버겐(C. J. van Rijserbergen)은 1993년에 질의 어기법과 브라우징기법을 통합한 혼성하이퍼텍스트시스템을 논하고 비문자적(non-textual) 문헌검색에서 이 혼성기법이 적합한지의 여부를 평가하기 위한 실험을 하였다.⁸⁾

이 논문에서 저자들은 대규모의 텍스트 데이터베이스와 멀티미디어 데이터베이스로부터 정보를 검색하기 위하여 질의어능력과 브라우징을 결합하는 혼성방법을 제시한다. 브라우징-기반의 하이퍼미디어 시스템들은 초보 이용자에게 이용의 용이성을 제공하고 어떤 미디어에 대해서도 동등한 접근 능력을 제공한다. 반면에 질의어-기반의 텍스트 검색시스템은 대규모의 데이터베이스에 대하여는 성능이 훌륭하나 멀티미디어 검색에는 성능이 빈약하다. 이 논문에서 제시한 모델은 이미지 등의 내용에 의해서는 접근될 수 없는 노드들에의 접근을 제공하기 위하여 이미지 노드들에 연결된 문자적 노드들이 하나의 클러스터를 형성하게 한다. 즉 클러스터-기반의 검색모델을 제시함으로써 질의어 검색과 브라우징을 결합하는 혼성시스템을 이룩할 수 있는 것이다. 이 제안된 모델의 타당성을 평가하기 위하여 텍스트문헌을 사용하는 두개의 실험을 수행하여 하이퍼미디어시스템을 위한 평가기법의 개발에 대한 어떤 통찰을 제공한다.

8) M. D. Dunlap & C. J. van Rijserbergen, "Hypermedia and Free Text Retrieval," *Information Processing & Management*, Vol. 29, No. 3(1993), pp. 287-297.

지금까지 살펴본 몇편의 논문들은 하이퍼텍스트 정보검색을 위하여 새로운 검색 모델을 개발하려는데 초점이 맞추어지고 있는데,

첫째, 하이퍼텍스트의 상태를 수동에서 능동으로 수정하기 위하여 다양한 링크유형을 구축하고 각 유형의 링크에 태그나 가중치나 레이블을 부여하는 방법,

둘째, 불리언기법의 질의어 결과를 추론확장하는 모델개발의 문제,

셋째, 브라우저의 출발점을 선택하기 위하여 질의어 기법을 통합하는 방법,

넷째, 베이지언 기반의 추론망이나 확률적 추론망을 구축하여 관련있는 노드들의 탐색을 도모하는 방법,

다섯째, 하이퍼텍스트시스템을 평가할 수 있는 기법개발에 대한 문제 등을 연구한 것으로 정리할 수 있다.

3. 이용자-시스템간의 상호작용성 향상을 위한 하이퍼텍스트의 사용

이상에서 논의한 하이퍼텍스트 정보검색을 위한 새로운 모델개발이라는 연구동향과 유사한 내용을 많이 포함하면서 시스템과 이용자간의 상호작용 성능의 향상에 특히 초점을 맞춘 연구들이 또 하나의 연구흐름을 형성하고있다.

앤더센 등(Michael. H. Andersen, Jacob Nielsen, and Henrik Rasmussen)은 1989년에 이용자 인터페이스로서 유사성-기반의(similarity-based) 하이퍼텍스트 브라우저(browser)를 개발하였다.⁹⁾

이 논문에서 발표한 HyperNews는 대단히 많은 수량의 기사들을 생성해내는 메시지 시스템에 대한 이용자 인터페이스 기능을 향상시키기 위하여 하이퍼텍스트 방법을 사용하는 것이 매우 유망하다는 것을 보여주었다. HyperNews는 역시 전

9) Michael H. Andersen, Jacob Nielsen and Henrik Rasmussen, "A Similarity-Based Hypertext Browser for Reading the UNIX Network News," *Hypermedia*. Vol. 1, Part 3(1989), pp. 255-265.

통적인 정보검색 페더다임과 하이퍼텍스트 방법을 결합한 또 다른 혼성시스템의 예이다. 저자들은 이 논문에서 비구조적 포맷(unstructured format)으로 분산되어 생성되어 있는 정보의 흐름에 대하여 하이퍼텍스트 인터페이스를 구축하는 것이 가능하다는 것을 보여준다. 만약 뉴스 기사들이 보다 구조화된 포맷으로 생성되었다면 더 많은 장점들을 제공하였을 것이다. UNIX Usenet Network News와 같은 기술메시지 교환을 위한 국제적 망은 매일 수천개의 기사들을 생산한다. 이 정보과잉 문제를 해결하기 위한 한가지 방법은 인간편집자가 이들을 분류하여 특수한 관심주제에 적합한 것을 추출하는 것이고 또 다른 대안으로는 컴퓨터가 어떤 지식기반의 우선순위 선정방법을 이용하여 메시지를 선정하는 것일 것이다. HyperNews는 이러한 의도에서 개발된 하이퍼텍스트 인터페이스로서 이것의 노드들은 밤중에 이 망으로부터 도착하는 하나의 기사이고 하이퍼텍스트 링크는 서로 서로에 대하여 논평하는 기사들간의 링크이다. 기사들간의 관련링크들의 구축은 정보검색분야에서 사용되고 있는 매트릭스에 의하여 계산되는 전문 유사성 비율에 근거하는데 이용자가 'similarity button'을 선택하면 현재의 기사에 가장 유사한 순서대로의 기사리스트가 디스플레이된다. 즉 정보검색공식에 근거한 유사성비율에 따라 현재의 기사에 관련있는 외부의 기사들에 대한 링크가 구축된다.

리-크(Michael Leek)는 1989년에 정보과적현상을 막기 위하여 하이퍼텍스트 브라우징기능을 적용하는 연구를 발표하였다.¹⁰⁾

텍스트의 작은 단위들을 가지는 하이퍼텍스트 시스템은 온라인 목록이나 서지검색시스템에서와 같이 너무 많은 검색결과를 초래하는 문제를 야기시키고 있다. 특히 하이퍼텍스트는 정보자료를 작은 단위(node)로 자르기 때문에 이 상황을 더욱 악화시킨다. 이 논문은 정보과적 문제를 다루기 위하여 전자적으로 하이퍼텍스트 단위들에 대한 전후 관계와 조직을 디스플레이하는 방법에 의하여 대규모의 정보를

10) Michael Leek. "What to do When There's Too Much Information," *Hypertext '89 Proceedings*. (Nov. 1989). pp. 305-318.

가진 시스템에서 이용자가 원하는대로 정보를 가려낼 수 있음을 시사해준다. 80만 건의 목록에 대하여 실험을 행하여 빈번한 단어들을 가지는 구들의 디스플레이와 계층적 주제카테고리에 의한 항목의 디스플레이가 유용하다는 것이 입증되었다. 그 디스플레이를 통하여 이용자는 적합한 용어나 표제를 선택할 수 있고 그래서 탐색결과를 줄일 수 있다.

크루츠 등(Donald B. Crouch et. al.)은 하이퍼텍스트 정보검색에 있어서 클러스터 계층구조를 사용하였고 이것이 시각적으로 이용자탐색을 도와줄 수 있음을 피력하였다.¹¹⁾

1980년대 중반부터 PC 보급에 기인하여 정보검색시스템의 최종 이용자를 돕기 위하여 사용자-시스템간의 상호작용은 잘 훈련된 전문탐색증개자의 활동을 대신하고 이용의 용이성을 도모하는 각종의 메뉴방식의 인터페이스 그리고 전문가시스템의 개발에서 구현되어 왔다. 그러나 탐색증개자에 의한 탐색결과의 비일관성과 불완전한 자연어처리의 문제를 완화하기 위한 인터페이스 설계를 위하여 저자들은 하이퍼텍스트시스템 사례연구를 제시한다. 전문가시스템의 자연언어 처리를 대신하여 하이퍼텍스트에 노드들간의 의미정보를 제공하는 클러스터 계층구조를 구축하고 이것이 이용자로 하여금 시각적으로 문헌집성을 효율적으로 탐색하도록 도와줄 수 있다고 피력한다. 이 연구의 직접적인 목적은 문헌집성을 쉽고 정확하게 탐색하고 브라우징할 수 있는 검색시스템을 개발하는 것으로 부분조망과 전체조망으로 구성된 상호작용적 브라우저가 클러스터 탐색과정동안에 이용자 통제 효과와 이용자 개입의 효과를 나타내는지를 평가하고 그래픽 데이터표현의 시각적인 분석의 효과를 평가하기 위하여 다양한 실험을 설계하는 것이다. 포괄적이고 신속성이 있는 실험 지원을 할 수 있도록 하이퍼카드를 이용하여 Machintosh II V Computer 상에서

11) Donald B. Crouch, Carolyn J. Crouch and Glenn Andreas, "The Use of Cluster Hierarchies in Hypertext Information Retrieval," *Hypertext '89 Proceedings*. (Nov. 1989), pp. 225-237.

실험을 이행하였고 메킨토시는 SMART 정보검색시스템이 장착되어있는 SUN 시스템에 LAN을 통하여 연결되었다.

1991년에 마르체티(Pier Giorgio Marchetti)와 벨킨(Nicholas J. Belkin)은 상호작용적인 온라인 탐색을 위한 탐색식 형성을 지원하는 지능적시스템 개발에 대한 연구를 발표하였다.¹²⁾

이 연구프로그램의 목표는 대규모 서지정보검색시스템에서 최종이용자와의 상호작용을 지원하기 위하여 지능적시스템을 개발하려는 것이다. 이를 위하여 이 연구는 첫째, 정보검색문제의 인지적 업무분석(cognitive tasks analysis)을 행하고 둘째, 서지데이터베이스의 일반적인 하이퍼텍스트 모델을 구축하고, 셋째, 지능적 인터페이스 설계, 라는 점진적인 이행전략을 세우고 있다. 이 연구에 의하여 생산된 두개의 서로 다른 구체적 산물로서는 첫째, ESA-QUEST시스템에 대한 운용적 설비인 HYPERLINE이라는, 최종 사용자용 디소러스인데 이것은 서지정보를 브라우징하는 기능을 제공하고 둘째, 최종 이용자가 탐색문을 형성하는 것을 지원하는, ESA-QUEST에 대한 상호작용적인 그래픽 사용자 인터페이스이다. 이 논문은 인지적 업무 분석에서 확인된 많은 문제점들이 용어 및 개념확인과 그것의 선택과 관계가 있음으로 이용자로 하여금 이를 용이하게 할 수 있도록 개념의 앞, 뒤를 네비게이션할 수 있고 용어간의 의미연관을 볼 수 있는 HYPERLINE을 실행함으로써 사용자-시스템간의 상호작용성을 높이려고 하였다. 이 논문은 어떻게 정보검색의 문제들-정보검색 과정, 정보검색 데이터베이스의 하이퍼텍스트 모델에 관련된 상호작용 문제들-에 대한 이론적 연구가 최종 이용자의 탐색을 실질적으로 도울수 있는 운영시스템을 생산하였는지를 입증한 것이다.

1993년에 벨킨 등(N. J. Belkin, P. G. Marchetti and C. Cool)은 상계는

12) Pier Giorgio Marchetti & Nicholas J. Belkin, "Interactive Online Search Formulation Support," ed. by Martha E. Williams *12th National ONLINE Meeting: Proceeding-1991*. New York, (May 7-9, 1991). pp. 237-242.

문에 ZOOM과 HYPRRLINE개념을 통합하여 이용자-시스템간의 상호작용을 강력하게 지원할 수 있는 인터페이스 설계에 대한 논문을 발표하였다.¹³⁾

저자들은 정보검색행위를 본질적으로 이용자-시스템간의 상호작용 과정으로 간주하고 사람들이 정보탐색에 종사할 때는 다양한 탐색행태를 나타냄으로 이들 행태를 반영할 수 있는 시스템을 설계하는 것이 중요하다고 지적한다. 그리고 이 행태들에 근거한 정보탐색 전략 모델 및 두 단계의 하이퍼텍스트 모델에 의하여 대규모 서지 정보검색시스템을 위한 인터페이스를 구축할 수 있음을 시사한다. 이 논문은 매우 상세한 인지적 업무분석을 이행하여 정보탐색목표를 정의하고 그 목표달성과 관련되는 필요사항들을 결정하여 1991년의 마르체티와 벨킨의 연구의 확장이라 할 수 있는 BRAQUE(Browsing and Query Formulation)를 개발하였다. 저자들은 인지적업무분석에 근거한 다양한 탐색전략유형들의 이론적 구조에서 실제적인 시스템 이행에 이르기까지에는 다음 문제들 즉 첫째, 개념적 설계를 운영시스템으로 이행하는 것, 둘째, 제안된 인터페이스 설계의 효과를 평가하는 것, 셋째, 탐색전략모델을 정당화하고 확장하는 것이 해결되어야 한다고 피력한다. 현재 그 첫 단계가 시행중에 있고 이것이 완성된 후 다음 두 단계가 계속 연구될 것이라고 결론을 맺는다.

레이다(Roy Rada) 등은 문헌과 용어의 표현에 대한 추상화의 수준의 차이를 허용하고 탐색깊이의 차이를 허용하며, 휴리스틱 정보탐색방법을 지원하고 어안도(fisheye-view)브라우징 기능을 수행하는 MUCH 시스템을 발표하였다.¹⁴⁾

이 논문은 정보를 조직하고 검색하고 재조직하는 하이퍼텍스트시스템인 MUCH(Many Using and Creating Hypertext)시스템의 사용의 기반이 되는 모델을 제시한다. 이 모델은 전통적인 정보검색시스템에서는 허용되지 않는 휴-리스틱 탐

13) N. J. Belkin, P. G. Marchetti & C. Cool. "BRAQUE: Design of an Interface to Support Interaction an Information Retrieval," *Information Processing & Management*. Vol. 29, No. 3 (1993), pp. 325-344.

14) Roy Rada, Weigang Wang, and Alex Birchall. "Retrieval Hierarchies in Hypertext," *Information Processing and Management*. Vol. 29, No. 3(1993), pp. 359-371.

색(이용자가 계속적으로 탐색전략을 수정할 수 있는 전략: 반복적 탐색(iterative searching)과는 대조되는 전략)을 지원하며 정보에 대한 휴리스틱탐색은 문헌과 용어의 표현에 있어서 추상화의 수준의 차이를 허용한다. MUCH 시스템내에서의 정보의 하부구조는 계층적 의미망이고 그 의미망은 기본적으로 노드들과 링크들에 이름이 부여된, 그 연결성이 인지될 수 있는 그래프다. MUCH 시스템은 텍스트 블록으로부터 색인용어를 자동적으로 선정하고 이 용어들에 의하여 노드를 만들고 용어들 사이의 혹은 노드들 간의 링크-동등 링크, 관련 링크, 계층 링크-를 구축하는 다양한 프로그램을 포함한다. MUCH시스템의 브라우징기능은 어안도 모델에 근거한 계층적인 내용리스트로서 문헌의 부분들에 대한 개요를 보여준다. 연구자들은 하이퍼텍스트의 의미망과 컴퓨터의 신속성은 인간으로 하여금 쉽게 정보를 조직하고 발견하고 재조직할 수 있게 하며 하이퍼텍스트시스템은 정보활동의 재사용을 지원하기에 매우 적절한 기술이라고 피력하였다.

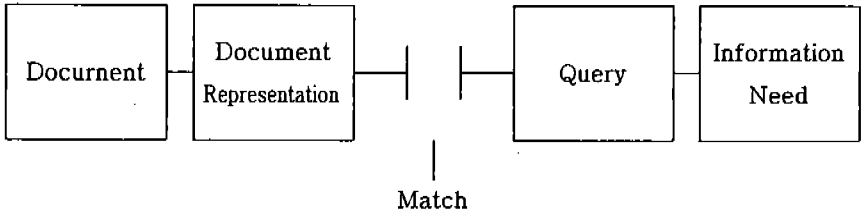
위에서 살펴본 몇편의 논문들은 하이퍼텍스트시스템에 의한 정보탐색과정에 있어서 사용자-시스템간의 상호작용성능을 향상시키려고 노력하는 하나의 연구흐름을 나타내고 있다고 볼 수있다.

이런 연구들은 ‘딸기 따기(berry-picking)’¹⁵⁾ 모델이라는, 상호작용성을 향상시키는 새로운 검색모델로 구현되고 있다고 볼 수 있다. 베이츠는 ‘딸기 따기’ 모델이라는 것을 제안하고 이를 고전적인 정보검색 모델과 대비한다. 그는 이 모델은 전통적인 정보검색보다 더 정확하게 실세계 인간의 정보탐색을 가능하게 한다. 이 새로운 모델의 의의는 온라인 정보시스템의 인터페이스 설계, 브라우징 성능, 정보 탐색 수행에 있어서의 인간-기계간의 관계, 그리고 시험이행의 측면에서 고려될 수 있다.”라고 요약한다.¹⁶⁾

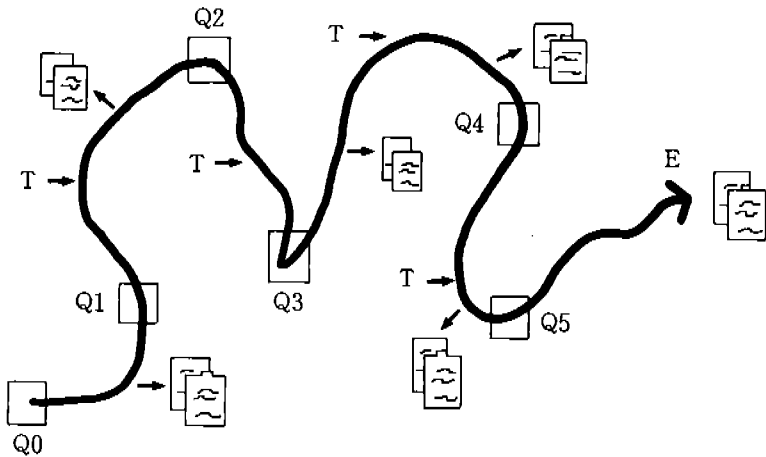
15) Marcia J. Bates, "The Berry-Picking Search: User Interface Design." In Martin Dillon ed. *Interfaces for Information Retrieval and Online Systems: The State of the Art.* (New York: Screenwood Press, 1991), p. 55.

16) Bates, pp. 55-61.

전통적인 정보검색 모델과 딸기따기 탐색 모델은 다음 (그림.1)¹⁷⁾과 (그림.2)¹⁸⁾에 나타나 있다.



〈그림 1〉 전통적인 정보검색 모델(The Classic Information Retrieval Model)



Q = Query variation
 T = thought
 E = exit
 = documents, information

〈그림 2〉 딸기따기 탐색 모델(A Berry-Picking, Evolving Search)

17) Bates, p. 56.

18) Bates, p. 57.

한마디로 (그림. 1)에서는 이용자가 입력한 질의어는 데이터베이스에 대조되어 하나의 탐색결과 집합을 생산하고 필요하면 반복적 탐색을 수행할 수 있다. (그림. 2)에서는 최종 이용자가 하나의 광범위한 주제로 탐색을 시작하여 다양한 정보원들을 옮겨다닐 수 있고 탐색과정 동안에 새로운 정보를 발견할 수 있으며 질의어 수정뿐만 아니라 질의어 자체가 계속 변경되는, 휴리스틱 혹은 진화적(evolving) 탐색이 가능하다.

콘클린(Jeff Conklin)은 전체적으로 딸기따기 개념의 탐색을 위하여 만들어진 것이 하이퍼텍스트 방법¹⁹⁾이라고 하였는데 위의 연구들은 이를 입증하기 위하여 스크린 위에서 상당한 정보의 양을 볼 수 있도록 부분적 혹은 전체적 그림 혹은 어안도와 같은 브라우징을 위한 그래픽 이용자 인터페이스를 개발하여 정보에 노출된 이용자로 하여금 탐구적으로 정보를 탐색할 수 있게 함으로써 시스템과의 효율적인 상호작용을 달성할 수 있는 방안을 제시하고 있다.

이용자-시스템 상호작용을 위한 인터페이스 설계의 세가지 원칙들 즉 첫째, 이용자와 업무들에 초점을 두는 것, 둘째, 원형시스템 사용의 효율을 경험적으로 측정하는 것, 셋째, 피드백에 의하여 개정을 하는 것에 비추어 볼때 지금까지 이 분야에서 이루어져 온 연구들은 전체적으로 첫째의 원칙을 실천하려는 노력이라고 볼 수 있으므로 앞으로는 둘째 및 셋째의 원칙들을 구현하기 위하여 노력이 기울어질 것으로 전망된다.

4. 정보검색 보조도구들의 개발에 대한 연구

하이퍼텍스트 정보검색효율을 보다 증진시키기 위한 의미기반의 하이퍼 색인 작성, 하이퍼텍스트 기반의 디소러스 구축 그리고 하이퍼텍스트 분류 개발에 대한

19) Jeff Conklin, "Hypertext : An Introduction and Survey," *Computer*, 20(9), (1987), pp. 17-41.

연구는 또 하나의 연구흐름을 형성하고 있다.

1991년에 어-(Clifford Urr)는 하이퍼텍스트 지식베이스내에서의 분류표(링크와 노드의 조직)의 중요성 및 위상을 검토하는 논문을 발표하였다.²⁰⁾

도서관 이용자가 어떤 토픽을 연구할 때 당면하는 뚜렷한 두가지 유형의 노력은 정보를 발견하는 것과 발견한 정보를 학습, 소화하여 활용하는 것에 있다고 볼 수 있다. 하이퍼텍스트는 첫번째 유형의 노력을 제거함으로써 두번째 유형의 노력을 증대시킬 수 있게 해야한다. 사용이 용이하고 정보접근 및 발견의 가치를 높이기 위하여 시스템의 접근 구조가 개발되어야 한다. 이용자로부터는 거의 지식을 요구하지 않으면서 사용이 용이한 이러한 구조는 충분히 축적된 지식이 있고 그 형태가 무엇이든지 간에 시스템은 분류구조를 가지고 있어야하며 이 질적인 분류구조는 일관성있고 예측할 수 있는 접근패턴을 제공한다. 저자는 단어들로부터 파일로의 링크를 자동적으로 제공하는 것은 진정한 하이퍼텍스트가 아니며 요어-기반의(concordance-based) 접근 시스템에 가깝다고 지적한다. 또한 저자는 미래의 하이퍼텍스트 경로를 개발하기 위해서는 하이퍼텍스트 분류를 구축해야 하고 훌륭한 분류체계에 의하여 명시적이고 정확한 특성의 지식베이스 구조를 사용할 수 있다고 진술한다. 중요한 것은 하이퍼텍스트 구축자들이 물리적으로 분리되어 있는 많은 정보단위들 사이에 발생하는, 외관상으로는 감추어져 있으나, 그러나 존재하는 서로간의 주요한 관련을 인식하는 것이다. 저자는 앞으로 검토되어야 할 점으로는 노드와 링크가 언제 시작되어 언제 끝나는가의 문제, 그것의 경계는 무엇이며 만약 존재한다면 어떤 보편적 규범이 모든 분류체계에 적절한가의 문제라고 지적한다.

폴라드(Richard Pollard)는 1993년에 서지데이터베이스에 대한 주제 브라우징을 위한 보조도구로서의 하이퍼텍스트-기반의 디소र्स 개발에 대한 연구를 발표하

20) Clifford Urr. "Will the Real Hypertext Please Stand Up ?" *Computers in Libraries*, vol. 11, no. 5. (May 1991), pp.45-49.

였다.²¹⁾

이 논문은 하이퍼텍스트 기반의 디소러스가 서지데이터베이스의 주제영역에 대한 네비게이션 보조도구로 사용될 수 있음을 제의하였고 <Thesaurus of ERIC Descriptors>의 하이퍼텍스트 버전으로의 이행에 관하여 기술하였으며 디소러스를 데이터베이스에 연결하기 위한 전략을 논의하였다. 전통적인 정보검색시스템들은 서지정보검색의 전략에서 점점 중요성이 인식되고 있는 브라우징에 대한 충분한 지원을 제공하지 못했으며 하이퍼텍스트 이용자들은 흔히 적은 정보공간을 브라우징할 때 조차도 방향을 상실하는 약점이 지적되어 왔다. 그래서 이 논문은 상업적인 하이퍼텍스트 개발도구인 GUIDE를 사용하여 서지데이터베이스의 주제영역에 대한 네비게이션 보조도구로서의 디소러스를 하이퍼텍스트 기반의 브라우징 인터페이스로서 개발하였다. 저자는 하이퍼텍스트 기법이 디소러스 표현에 적합하며 그 이유로서 첫째, 하이퍼텍스트는 데이터베이스가 수많은 서로 관련되어 있는 정보 조각들로 구성되어 있고 이용자는 어느때고 적은 분량의 정보만을 필요로 하는 상황에서 특히 적절한데 디소러스는 이 패턴에 매우 잘 부합한다는 점과 둘째, 하이퍼텍스트의 특성이 비선형적으로 정보공간을 네비게이트하는 것이고 디소러스의 구조는 본질적으로 비선형적이라는 점을 들고 있다.

아렌츠(Hans C. Arents)와 보가에르츠(Walter F. L. Bogaerts)는 용어색인을 의미적 하이퍼색인으로 변경함으로써 하이퍼미디어 정보에 대한 개념기반의 검색을 도모하기 위한 연구를 발표하였다.²²⁾

이 연구에서 저자들은 하이퍼텍스트와 하이퍼미디어시스템의 정보검색의 잠재능

21) Richard Pollard, "Hypertext -Based Thesaurus as a Subject Browsing Aid for Bibliographic Database," *Information Processing & Management*. Vol. 29, No. 3(1993), pp. 345-357.

22) Hans C. Arents and Walter F.L. Bogaerts. "Concept-Based Retrieval of Hypermedia Information: From Term Indexing to Semantic Hyperindexing," *Information Processing & Management*. Vol. 29, No. 3(1993), pp. 373-386.

력을 이끌어 내는 열쇠는 하이퍼네트워크내의 정보에 대한 의미기반의 색인을 작성하는데 있으며 이 색인구조는 효과적인 시각화와 네비게이션을 가능하게 할 수 있어야 한다고 지적하고 있다. 이 논문은 Active Library Corrosion 프로젝트의 수행 동안에 개발한 하이퍼미디어 정보의 개념기반의 색인작성에 대한 두가지 다른 방법들을 검토하고 있는데 하나는 3차원적인 색인의 네비게이션을 가능하게 하는 시각화된 용어색인과 또 다른 하나는 버튼 링크(broad button-link) 네비게이션을 가능하게 하는 의미기반 하이퍼색인작성 방법이다. 그리고 저자들은 하이퍼미디어 정보검색의 두가지 가능한 개념화 방안들 중에서 그 하나를 채택함에 있어서 <프리티세 1988년 논문>에서 논의된 질의어에 의한 노트검색 혹은 용어열부함에 의한 검색 ('small-document approach')과 브라우징에 의한 검색 혹은 그래프 횡단 검색 ('graph-traversal approach')중에서 후자를 선택하였다. 그 이유는 컴퓨터 지식이 없는 이 분야의 엔지니어들이 실제로 사용할 수 있는 시스템을 개발해야 한다는 것과 브라우징 탐색이 이용자들에게 보다 덜 심각한 인지적 부담을 부과한다는 선행연구²³⁾에 기인하고 있다.

III. 온라인목록 검색을 위한 하이퍼텍스트 기법의 적용에 관한 연구

온라인목록시스템의 하이퍼텍스트시스템화에 대한 연구, 온라인목록에 대한 포인트-앤드-클릭(point-and-click)기반의 이용자 인터페이스 개발에 대한 연구, 그리고 온라인목록 검색을 위한 하이퍼텍스트 기반의 원격 이용자용 인터페이스 개발에 대한 연구는 점차 관심이 커져가고 있는 새로운 연구동향을 형성하고 있다.

23) Gary Marchionini and Ben Shneiderman, "Finding Facts vs. Rrowsing Knowledge in Hypertext Systems," *Computer*. Vol. 21(Jan. 1988). pp. 70-79.

히에르페(Roland Hierppe)는 하이퍼텍스트기법에 의한 목록시스템개발에 대한 약 4년간의 지속적인 연구의 중간발표로서 그동안의 진척과 발표당시의 현황을 보고하는 논문을 1989년에 제출하였다.²⁴⁾

이 논문은 히에르페가 1986년에 발표한 하이퍼카타로그 프로젝트에²⁵⁾ 대한 전체적인 개념의 구현을 위한 몇년간의 노력후 그 진행사항을 보고한 것으로 그 목적은 문제에 대한 당장의 해결을 제시하려는 것이 아니고 보다 견고한 실용시스템 개발을 위하여 가능한 많은 문제들을 야기시키고 그 문제들을 하나씩 해결해 나가려는 것이다. HyperCat 프로젝트가 지향하는 것은 전통적인 도서관 목록의 확장이고 향상이다. HyperCat는 전통적 목록위에 구축되며 전통적 목록이 정보검색시스템과 결합해서 할 수 있는 모든것 그리고 그 이상을 할 수 있도록 설계되었다. 장기간에 걸쳐 연구되어온 HyperCat에 대한 논의점들이 이 논문에 통합되어 있다. 즉 HyperCat는 첫째, 목록이용의 일차적 모드로서 브라우징과 네비게이션을 지원하고 동시에 전통적 탐색을 제공하며, 둘째, 전통적 목록보다 훨씬 더 고도의 내적 구조를 가진다. 그리고 그 구조속에서 정보와 정보간의 관계를 확립하고 그들간의 궤적을 따라갈 수 있는 도구들을 가진다. 셋째, 사용자 모델링에 근거하여 서로 다른 사용자 유형에 적응하는 시스템이 된다. 넷째, 현재의 목록정보보다 더 많은 정보를 포함한다. 현재 LIBLAB의 주요 강조점은 실용적인 하이퍼카타로그를 구축하기 전에 훌륭한 경험과 배경지식을 얻을 수 있도록 이미 개발된 많은 하이퍼텍스트 시스템들 즉 NoteCard, GUIDE, HyperCard, Links 등의 성능과 제약점을 밝혀내려고 하는데 있다. 이 논문은 이런 의도의 일환으로서 온라인목록과 사용자와의 상호작용을 지원하기 위하여 객체지향적이고 포인트-앤드-클릭 스타일의 인터

24) Roland Hierppe, "HyperCat at LIBLAB in Sweden: (A Progress Report)." in Charles R. Hildreth ed. *The Online Catalog: Development and Directions*. London, The Library Association, 1989, pp. 177-206.

25) Roland Hierppe, "Project HYPERCATALOG." in B. C. Brookes, ed. *Intelligent Information Systems for the Information Society*, North-Holland, Elsevier Science Publisher B. V., 1986, pp. 211-230.

페이스를 이용하고 Zerox의 NoteCard 버전 1.2를 이용하였다.

보베이 (Julie Bobay) 등은 대학도서관 목록의 원격 이용자들을 위한 참조와 교육을 제공하는 이용자편의 시스템인 SULAN/LINKWAY를 소개하였다.²⁶⁾

이 논문은 도서관 자원에 대한 확장된 그리고 용이한 접근을 제공하기 위한 함의를 논의한다. 미국의 인디애나주의 공공도서관, 학교도서관 그리고 특수도서관 이용자들이 NOTIS 온라인 목록에 대한 포인트-앤드-클릭 기반의 인터페이스인 SULAN/LINKWAY를 사용하여 여덟 개의 인디애나 연구도서관의 목록에 접근할 수 있는데 SULAN은 주립대학교도서관자동화망(State University Library Automation Network) 이고 LINKWAY는 IBM 하이퍼텍스트 소프트웨어 제품이다. 이는 사서들의 개입과 설명을 줄일 수 있는 이용자편의의 그래픽스 및 버튼-기반의 인터페이스이며 저자들은 "자유롭고 손쉬운 도서관자원에의 접근을 위해서는 그리고 도서관 장서에 대한 효과적인 원격접근을 제공하기 위해서는 LINKWAY와 같은 진보된 하이퍼텍스트 기술을 이용해야 한다고 주장한다.

앤티elman(Kristian Antelman)은 1992년에 DELCAT MacPAC이라는 온라인 열람목록에 대한 하이퍼텍스트 인터페이스를 소개하였다.²⁷⁾

DELCAT MacPAC은 델라웨어대학(The University of Delaware)이 개발한 NOTIS 온라인목록에 대한 하이퍼카드 기반의 프론트-앤드 인터페이스다. NOTIS 온라인목록에 대한 하이퍼텍스트 기반의 인터페이스를 제공하기 위하여 이루어져 온 다른 두가지 노력사항들로서는 첫째, 메킨토시-기반의 MacNOTIS-Texas A & M University이며 둘째, IBM 기반의 SULAN-LINKWAY-Indiana University이다. MacPAC은 이 두개의 프로그램보다는 덜 정교하지만 메인프레임 기반의 소프트웨어로부터 독립되어 있는 잇점을 제공한다. 그래서 델라웨어 대학은

26) Julie Bobay, ED Stocky & Mary Pagliero Popp, "Library Services for Remote Users with Linkway," *Reference Technologies*. (Fall 1990), pp. 53-57.

27) Kristian Antelman, "Hyping the OPAC: Adopting a Macintosh User Interface to the NOTIS Online Catalog," *Reference Technology*. (Spring 1992), pp. 27-32.

도서관 외부의 메킨토시 이용자들에게 MacPAC을 널리 사용할 수 있게 한다. MacPAC은 NOTIS온라인목록에 대하여 표준의 메인프레임 기반의 터미널 인터페이스가 제공할 수 없는 다음과 같은 몇가지 특징들을 첨가하고 있다. * 마우스지원 * 서지스크린과 색인스크린과 가인스크린 사이를 움직여 다니는 아이콘들, 'previous' 및 'next' 버튼들, 스크린을 노트 파일에 그리고 디스크에 저장하는 능력 등. 저자는 보다 생산적인 OPAC 탐색을 할 수 있도록 그래픽스와 칼라와 마우스지원의 탐색이 OPACs 에 통합되고 있으며 브라우징과 궤적 그리고 사용자 모델링과 같은 개념들을 하이퍼텍스트로 부터 빌려오는 온라인목록 탐색이 성행하게 될것이라고 꾀력한다.

페이터센(Annelise Mark Pejtersen)은 공공도서관에 있어서 이용자 대출의 50% 이상을 구성하고 있는 소설자료 검색에 연루되는 문제를 해결하기 위하여 임의의 초보이용자들을 지원하는 상호작용적이고 멀티미디어 기반의 OPAC 시스템인 The Book House의 개발에 대하여 설명한다.²⁸⁾

이질적인 인구를 이용대상자로 갖고 있는 도서관들은 메뉴형태, 도움말 디스플레이, 통보적 에러메세지, 그리고 브라우징 설비 등으로 온라인목록 이용자들을 도와 줄 수 있는 OPACs의 이용자편의 성능을 향상하기 위하여 노력을 기울여 왔으나 많은 연구들은 너무 많은 탐색실패를 보고하고 있다. 이 탐색실패의 주요 원인은 기존의 OPAC 시스템이 가지고 있는 문제점들 즉 성능이 훌륭하지 못한 인터페이스와 사용자-시스템간의 대화 지원의 부재, 질의어 언어와 서지레코드 어휘간의 부합 정도의 미약, 그리고 이용자와 시스템간의 상호작용의 빈약함에 기인된다.²⁹⁾ 저자는 이러한 문제들을 극복하기 위하여 임의의 초보이용자들을 도울 수 있는

28) Annelise Mark Pejtersen, "New Model for Multimedia Interfaces to Online Public Access Catalogues," *The Electronic Library*. Vol. 10. No. 6. (Dec. 1992), pp. 359-366.

29) Charles R. Hildreth, "General Introduction; OPAC Research: Laying the Groundwork for Future OPAC Design," in Hildreth ed. *The Online Catalogue: Developments and Directions*. pp. 1-24.

The Book House를 하이퍼텍스트와 멀티미디어 기술을 통합한 TOOLBOOK 에 의하여 개발하였다. 이를 위하여 이용자의 목표분석과 소설자료 검색영역의 업무분 석에 근거하여 이용자의 정보요구를 충족시키고, 소설에 대한 새로운 분류체계를 설계해야할 필요성을 피력한다. 저자는 또한 The Book House의 하이퍼텍스트 기능과 멀티미디어 데이터베이스가 기존의 검색방법과 데이터베이스 보다 더 우수한 상호작용적인 최종 이용자용 OPAC 시스템을 개발할 수 있는 이유를 밝히고 이는 앞으로 곧 실용될 것이라는 전망을 피력한다.

IV. 인공지능의 하이퍼미디어에의 적용에 대한 연구

할라즈(F. G. Halasz)는 다음 세대의 하이퍼미디어시스템은

- * 하이퍼미디어 네트워크상에서 향상된 탐색과 질의어를 제공할 수 있는 설비를 지원하고
- * 기본적인 노드의 링크모델을 증대시키는 구성방법을 지원하고
- * 자주 변경되는 정보를 다룰 수 있는 실제적 구조를 지원하고
- * 하이퍼미디어 네트워크상에서 평가들 가능하게하는 설비를 지원하고
- * 변형(versioning)과 접근 통제를 지원하고
- * 협동적 업무를 지원하고 그리고
- * 완전한 확장성과 적응성(tailorbility)을 갖추어야 한다.³⁰⁾고 지적하였다.

할라즈가 예견한 이상의 문제들은 하이퍼미디어에 더 많은 지능을 구축함으로써 즉 개념과 하이퍼미디어 분야의 기법과 인공지능 기법을 통합하는 혼성시스템을 설계함으로써 해결될 수 있으리라는 관점에 착안하여 이루어져 오고 있는 많은 연

30) F. G. Halasz, "Reflections dn NoteCards : seven issues for the next generation of hypermedia systems," *Comm. of ACM*, Vol. 31, No. 7(1988), pp. 836-852.

구들은 최근의 새로운 연구동향을 형성하고 있다.

앤더-선(R. E. Anderson) 등은 자연언어처리 기법을 사용하여 하이퍼텍스트응용을 향상시킬 수 있는 방안에 대한 연구를 발표하였다.³¹⁾

이 연구는 CD Word Inc.에서 개발한 CD Word 시스템을 소개한다. 이 시스템은 하이퍼텍스트환경에서 성경공부를 할 수 있도록 신학영역의 사전적인 키-워드 혹은 구를 입력하여 이에 부합하는 관련 데이터를 검색하는 하나의 응용시스템이다. CD Word는 CD-Rom에 축적된 문헌을 문헌의 단서가 되는 사전적인 항목의 입력에 의하여 탐색하고 이용자에게 제시되는 정보의 질을 향상시킬 수 있도록 4단계의 자연언어처리 기법을 적용하였다. 가장 낮은 단계에서는 질의어와 문헌간에 부합된 항목의 중요성을 표시하기 위하여 단순히 그 중요도를 발견해 낸다. 그 다음으로 낮은 단계에서는 목표로 삼은 단어가 나타나는 텍스트에 함께 나타나는 어떤 단어의 빈도가 가장 낮은 단계에서의 부합의 중요도를 나타내는지를 보여준다. 그 다음 단계에서는 ATN 문법을 사용하여 구문분석을 한다. 가장 상위 단계에서는 단어의 의미들(동의어 등)에서 시작하여 문장분석에 이르기까지 의미분석을 수행한다. 결론적으로 이 논문은 이용자가 정보공간을 네비게이션할 때 접근한 정보의 진가를 효과적으로 판단하기 위하여 멀티미디어 응용에 있어서 텍스트데이터 기반의 하이퍼텍스트를 적용하기 위해서는 반드시 자연언어처리 기법을 통합해야 한다고 주장한다. 앞으로의 연구과제로는 전문적인 성경학자들이 그들이 사용하고 싶어하는 검색질문의 유형을 더 정확히 확인하고 그 질문들을 적절히 처리할 수 있는 구문분석기(parser)의 능력을 신장하는 실험을 이행하는 것이라고 전한다.

리틀포드(Alan Littleford)는 미래세대의 하이퍼미디어시스템은 인공지능기법과 하이퍼미디어기법이 통합된 혼성시스템이 될 것이라고 주장하며 인공지능기법이

31) R. E. Anderson, P. J. Sallis and W. K. Yeap, "Enhancing a Hypertext Application Using NLP Techniques," *J. of Information Science*, Vol. 17(1991), pp. 49-56.

하이퍼미디어 도큐멘테이션을 증대시킬 수 있는 방법들을 설명하기 위하여 정보검색 원칙과 하이퍼미디어의 장·단점, 그리고 인공지능 기능을 서명하는 논문을 발표하였다.³²⁾

이 논문의 요지는 다음과 같다. 즉 인쇄문헌 저자는 자연스런 방식으로 문헌의 지면배치를 할 수 있고 좋은 색인과 좋은 목차를 제공하는 것으로 충분하며 그 다음에는 독자들이 스스로 그것들을 활용하게 된다. 하이퍼미디어는 이용자의 정보탐색을 위하여 적절한 수준의 상세성 정도를 링크뒤에 감출 수 있고 필요할 때 호출될 수 있으며 관련된 토픽으로 부터 다음에 무엇을 읽을 것인지를 선택할 수 있다는 점에서는 신축적이지만 이것은 아직까지는 전통적인 데이터베이스를 얼마간 향상시킨 것에 불과한 것이다. 인공지능을 하이퍼미디어 도큐멘테이션에 적용하는 것은 문헌-특히 정책 및 절차편람, 참조 및 이용자 메뉴얼, 학술논문, 리뷰기사, 유지지침서 등이 고정된 문헌(fixed document)이라는 개념을 벗을 수 있게 준다. 고정된 문헌이라는 개념을 추방해야하는 이유는 첫째, 같은 수준의 지능을 가진 독자들이라도 같은 정보를 다른 방식을 이용할 수 있고, 둘째, 이용자가 선택하는 링크들은 이용자의 정신구조, 현재의 관심, 기술적 수준에 의존할 수 있고, 셋째, 독자들이 실수를 할 수도 있고 실수를 한다. 는 것으로 지적되고 있다. 저자는 인공지능기법이 하이퍼미디어 도큐멘테이션을 증대시키기 위해 사용하는 네 가지 방법을 다음과 같이 제시하고 있다.

첫째, 조직에 대한 지식(organizational knowledge-문헌의 조직에 대한 의미론으로 구성됨)을 하이퍼베이스에 'IF-THEN'문의 추론규칙으로 심는 것,

둘째, 전문영역에 대한 지식(domain knowledge-문헌의 내용에 대한 지식)을 역시 추론규칙으로 하이퍼베이스에 심는 것인데, 여기서 사용할 수 있는 한가지

32) Alan Littleford, "Artificial Intelligence ND HYPERMEDIA," in ed. Emily Berk and Joseph Delvin. *Hypertext/Hypermedia Handbook*. New York, McGraw-Hill Publishing Company, Inc, 1991, pp. 357-377.

적합한 인공지능기법은 개념계층이 될 수 있다.

셋째, 부분 에이전츠(local agents)기법은 이상의 지식을 효율적으로 활용하기 위한 핵심 개념으로 이용자의 어젠다(agenda)를 추적하고, 매립된 지식을 이용하여 요구된 정보를 내 보내는, 이용자를 위해 작용하는 프로그램이다. 넷째, 전체적 에이전츠(global agents)기법은 역동적인 web을 위해 일한다. 즉 문헌이 추가되면 이것은 정보를 개념계층에 의하여 연결하고 부분적 에이전츠에게 새로운 web 상태를 통보한다. 마지막으로 저자는 이론을 실천할 수 있는 방안으로서 인공지능과 하이퍼미디어가 통합되어 있는 상당히 성능이 우수한 제품인 HyperBase를 설명하고 있다.

아렌츠(Hans C. Arents)와 보가에르츠(Walter F. L. Bogaerts)는 미래 세대의 지능적 정보참조시스템-지식업무 종사자들이 지능적으로 정보를 참조할 수 있도록 도와주는 시스템-이 어떻게 3세대 하이퍼미디어(3rd order hypermedia)를 사용하여 구축될 수 있는지를 연구 발표하였다.³³⁾

저자는 IIC(Intelligent Information Consultation) 시스템의 필요성을 강조하고 이와같은 시스템을 구축하기 위하여 새로운 세대의 하이퍼미디어가 필요하다고 지적하며 새로운 세대 혹은 제 3 세대 하이퍼미디어는 문헌의 내용을 표현할 뿐 아니라 그것을 능동적으로 조작할 수 있는 하이퍼미디어이고 참조와 표현 그리고 네비게이션을 보다 지식기반적이게 한 하이퍼미디어라고 정의한다. 저자는 지식기반의 하이퍼미디어는 기본적인 객체지향의 페러다임을 확장하는 정보표현을 가능하게 하여 하나의 객체(텍스트, 그래픽, 그림, 소리)는 그 속성과 그 형태(방법)에 의해서 뿐만 아니라 다른 객체들에 대한 의미관계에 의해서도 표현될 수 있다고 설명한다. 아울러 인공지능의 지식기반 시스템과 하이퍼미디어시스템의 성공적인

33) Hans C. Arents and Walter F.L. Bogaerts, "Building Intelligent Information Consultation systems using third-order hypermedia systems," in ed. S. W. I. F. T., *Intelligent Information Access*, London, Elsevier Science Publishers B. V., 1992, pp. 29-39.

통합은 내용 의미론(어떻게 한 정보가 다른 정보에 관련되어 있는가?), 네비게이션 의미론(어떻게 정보들이 서로서로 의미있게 관련지어 지는가?), 그리고 표현 의미론(어떻게 정보가 최적의 상태로 디스플레이 될 수 있는가?)의 세가지 사이의 명백한 개념적 구분이 유지될 수 있어야 한다고 피력한다. 저자는 이런 개념적 구분을 가능하게 하는 하나의 구조(architecture)를 정의하려고 시도했는데 그것은 Model-Map-View-Praxis(MMVP)이다. MMVP구조는 4가지 기능의 층들- 정보 내용과 의미를 나타내는 층(the model layer), 네트워크 구조와 그 사이의 횡단을 나타내는 층(the map layer), 네트워크 네비게이션과 표현을 나타내는 층(the view layer), 시스템의 참조기능과 조작기능을 운영하는 층(the praxis layer)으로 이루어진다. 이 논문은 MMVP구조에 대한 이론을 구현하기 위하여 IKON(Intelligent Knowledge Objects Navigator)라고 부르는 의미지각의 하이퍼미디어 시스템에 대한 연구프로젝트의 어떤 특성을 설명함으로써 제 3세대하이퍼미디어 시스템의 배후의 아이디어를 명료하게 밝히고 있다.

하이퍼텍스트/하이퍼미디어에 대한 연구는 이제 그 독자적인 특성과 기술상의 차이를 입증하려는 단계를 훨씬 넘어서서 기존의 검색기법들을 통합하고 하이퍼텍스트에 의미를 심음으로써 질의기법과 네비게이션 기법, 그리고 인공지능 기법에 내재한 한계점들을 지양하려는 보다 복잡하고 정교한 체제로 나아가고 있다. 1991년 현재 인공지능 측면과 하이퍼미디어를 결합하고 있는 몇몇 상업적인 제품들이 출현하고 있는데 <표. 1> 그들은 각기 다른 방법으로 때로는 다른 이유 때문에 운영되지만 두 가지 기법의 혼용을 시도하고 있다.³⁴⁾

34) Littleford, p. 376.

제 품 이 름	제 작 기 관	비 고
Nexpert Object	Neuron Data	하이퍼카드 스택과 전문가시스템의 연관을 허용함
Hypertrans	Texas Institutions, Inc.	인공지능 기법을 사용하여 문헌분석을 수행하고 가장 적절한 하이퍼텍스트 링크를 구축할 수 있게함. 지식데이터를 설명하는 생성규칙기반의 전문가 기반의 전문가시스템과 하이퍼텍스트시스템을 통합한 것임.
KnowledgePro Windows	Knowledge Gardens, Inc.	마이크로소프트 윈도우환경을 위한 응용구축 도구이며 결정나무(decision tree)와 하이퍼텍스트를 지원한다.

이상의 어떤 것도 지식창고(knowledge warehouse) 개념을 지원하기에는 충분치도 않고 나름대로의 약점을 노출하고 있다.³⁵⁾

V. 결 언

1. 서지데이터베이스의 검색과 수치데이터베이스의 검색, 지식베이스의 검색 그리고 하이퍼텍스트의 검색으로 이어지는 다양하고 새로운 검색기법들이 연구되고 개발되어 오고 있으며 하나의 새로운 기법의 개발은 기존의 검색모델과 통합됨으로써 더 나은 성능이 입증되었다.

35) Littleford, p. 376.

2. 이제 하이퍼텍스트연구자들은 전통적인 질의어 검색기법을 통합하고, 온라인 열람목록의 인터페이스로 적용되며, 인공지능 기법을 통합함으로써 대규모 하이퍼텍스트의 네비게이션 과정에서 야기되는 방향상실의 문제, 신속하게 적합정보에 접근해야 하는 문제를 해결하고, 최종이용자와 시스템간의 상호작용 성능을 향상시키고, 개개 이용자의 정보탐색 행태에 적응할 수 있는 의미기반-검색의 가능성을 찾기위한 연구를 계속하고 있다.

3. 하이퍼텍스트에 관한 연구는 개념과 하이퍼미디어 기법과 인공지능 기법의 통합을 지향할 것이며 이것을 위하여 궁극적으로 필요한 것은 전문지식의 의미론을 하이퍼텍스트에 표현하는 것이지만 이것을 수작업으로 행하는 것은 시간 소모적이며 지식표현기법을 모르는 저자들에게는 불가능한 일이다. 따라서 문헌의 의미를 하이퍼텍스트에 자동적으로 심을 수 있는 컴퓨터보조문헌공학(Computer Aided Document Engineering-CADE)³⁶⁾과 같은 적절한 개발환경이 필요하다.

36) Littleford, p. 377.

Research Trends in Hypertext Information Retrieval

Lee Young-Ja*

〈Abstract〉

The purpose of the study is to understand the research trends in the hypertext information retrieval. Around 30 related papers were investigated, from which three distinctive streams of research trends are grasped: 1) a trend of incorporating the traditional retrieval models, especially the query-based searching model into the hypermedia system, 2) a trend of applying the hypermedia system as an interface to the OPAC system, 3) a trend of incorporating the artificial intelligence techniques into the hypermedia techniques.

The research on the hypermedia is going on, and the research directions will be increasingly intend to incorporate the traditional retrieval models and artificial intelligence techniques into the hypermedia system.

* Professor, Department of Library and Information Science, Kyung Pook National University.