

# 전자도서관 환경에서 이용자와 정보 시스템간의 상호작용 연구\*

## Supporting Interaction with Distributed Information Resources: Implications for Digital Libraries

박소연(So-yeon Park)

### 목 차

1 연구의 필요성 및 목적	3.1 이용자 선호도
2 연구방법	3.2 이용자 만족도, 이용 편의성
2.1 연구설계, 데이터베이스, 검색과제, 검색질문	3.3 검색 효율성
2.2 시스템	3.4 정보 검색 행위
2.3 조사 방법, 조사 도구	3 연구방법
3 연구의 결과	4 결 론

### 초 록

본 연구는 전자도서관 환경에서 이용자들이 다수의 데이터베이스를 개별적으로 검색하는 경우(분리 인터페이스)와 통합적으로 검색하는 경우에(통합 인터페이스) 이용자 선호도, 이용자 만족도, 검색 효율성과 정보 검색 행위를 비교 분석하였다. 두 시스템, 분리 인터페이스와 통합 인터페이스를 비교하기 위해 TREC(Text REtrieval Conference)의 데이터베이스와 검색 질문 중 일부를 이용하여 Rutgers 대학에서 실험이 실시됐다. 연구 결과 이 이용자 그룹은 통합 인터페이스 보다 분리 인터페이스를 선호하고 분리 인터페이스의 검색 결과에 더 만족하는 것으로 나타났다. 또 이 이용자 그룹에게는 분리 인터페이스의 검색 효율성이 통합 인터페이스의 검색 효율성 보다 높은 것으로 나타났다. 이러한 결과는 통합 인터페이스가 분리 인터페이스 보다 바람직할 것이라는 데이터베이스 통합분야의 일반적 전제와 상반된다고 할 수 있다. 본 연구는 분산 환경에서 이용자와 정보 시스템 간의 상호 작용에 대한 이해를 넓히고 전자 도서관 환경에서의 시스템 개발에 시사점을 제시한다는 데 그 의의가 있다고 할 수 있다.

### ABSTRACT

This article addresses a crucial issue in the digital library environment: how to support effective interaction of users with heterogeneous and distributed information resources. In particular, this study compared user preference, effectiveness, and searching behaviors in systems that implement interaction with multiple databases through a common interface, and with multiple databases as if they were one(integrated interaction) in an experiment in the Text REtrieval Conference (TREC) environment. The results of this study suggest that: (1) the general assumption of the information retrieval literature that an integrated interaction is best needs to be revisited; (2) it is important to allow for more user control in the distributed environment; (3) for digital library purposes, it is important to characterize different databases to support user choice for integration; and (4) certain users prefer control over database selection while still opting for results to be merged.

키워드 : 상호작용, 전자도서관, 인터페이스 디자인, 데이터베이스 통합

\* 본 논문은 Rutgers대학 박사 학위 논문 일부를 요약한 것임.

■ 논문 접수일 : 2000년 11월 16일

## 1 연구의 필요성 및 목적

전자 도서관(digital library) 환경에서 중요한 연구 과제 중의 하나는 어떻게 하면 이용자가 분산되어 있는 다수의 데이터베이스를 효과적으로 검색하고 이용하게 하는 가이다. 전형적인 전자 도서관 환경 속에서 이용자들은 다수의 상이한 데이터베이스들 중 적절한 데이터베이스를 선택해야 하고, 선택된 데이터베이스들을 검색하고 이용한 후에는 이 데이터베이스들로부터 나온 검색 결과를 통합해야 한다. 분산환경 속에서는 서로 다른 종류의 수많은 데이터베이스가 존재하므로 이용자가 다수의 데이터베이스를 어떻게 탐색하고 상호작용하며 어떻게 탐색하기를 원하는가에 대한 연구는 매우 중요하다 할 수 있다.

정보검색 분야에서의 이용자와 정보시스템간의 상호작용(Interaction)의 중요성이 점차로 부각되고 있음에도 불구하고, 분산환경에서의 이용자와 데이터베이스간의 상호작용에 대한 연구는 거의 전무한 실정이다. 지금까지 대부분의 데이터퓨전(Data Fusion)과 데이터베이스 통합(Database Merging)분야의 연구들은 시스템의 검색 효율성(effectiveness)과 기술적인 문제에만 초점을 맞추어 왔으며 이용자와 시스템간의 상호작용에 대한 문제는 등한시 해왔다(Belkin, Kantor, Cool, & Quatrain, 1994; Belkin, Kantor, Fox, & Shaw, 1995; Callan, Croft, Harding, 1995; Fox & Shaw, 1994; Hawking, Thistlewaite, & Bailey, 1997; Lee, 1995; Lu, Callan, & Croft, 1996; Savoy, LeCalve, & Vrajitoru, 1997; Schiettecatte, 1997; Voorhess, Gupta, & Johnson-Laird, 1995; Voorhess, 1997; Xu

& Callan, 1998). 이 분야의 연구들은 데이터베이스통합이 정보검색시스템의 효율성을 높일 수 있으므로, 유용성이나 이용자 만족도에도 기여할 수 있다고 전제하는 경향이 있다.

본 연구는 정보검색 분야 특히 데이터베이스 통합(database merging) 분야에 널리 퍼져있는 이러한 전제를 시험하는 것을 목적으로 한다. 본 연구는 필자의 럿거스대학교 대학원 박사과정 논문의 일부를 축약한 것인데(Park, 1999), 원 논문은 이용자들이 다수의 데이터베이스를 개별적으로 검색하는 경우(이후 이 시스템을 분리 인터페이스라고 칭함, 그림1)와 다수의 데이터베이스를 통합해서 검색하는 경우(이후 이 시스템을 통합 인터페이스라고 칭함, 그림2)의 이용자선호도, 이용자 만족도, 검색 효율성(effectiveness)과 정보검색행위를 비교하였다. 이러한 주 연구과제를 토대로 다음과 같은 연구가설이 설정되었다.

- 가설1: 분리인터페이스와 통합인터페이스에 대한 이용자의 선호도에 유의적 차이가 있을 것이다.
- 가설2: 두 시스템에 대한 이용자의 만족도에 유의적 차이가 있을 것이다.
- 가설3: 두 시스템 사용 시 이용 편의성에 유의적 차이가 있을 것이다.
- 가설4: 두 시스템 사용 시 검색 효율성에 유의적 차이가 있을 것이다.
- 가설5: 두 시스템 사용시 이용자의 정보 검색 행위에 유의적 차이가 있을 것이다.

분리 인터페이스의 경우, 이용자들은 공통의 인터페이스를 통해 상이한 다수의 데이터베이스들을 개별적으로 사용하고, 상호작용 해야 한

다. 이 인터페이스 사용 시 이용자가 다수의 데이터베이스들 중에서 자신이 원하는 데이터베이스를 선택할 수 있고, 적합성 피드백(relevance feedback)과 같은 시스템의 특성이 쉽게 개발될 수 있다는 장점이 있다. 그러나, 이용자들은 한번에 단일 데이터베이스 밖에 사용할 수 없고, 데이터베이스를 바꾸는 경우에는 검색을 반복해야 한다. 분리인터페이스의 또 다른 단점 중의 하나는 이용자 스스로 수많은 데이터베이스 중에서 적절한 데이터베이스를 찾아야하며, 다수의 데이터베이스로부터 나온 검색결과를 스스로 비교하고 통합해야 한다는 점이다. 다이알로그(Dialog)나 렉시스-넥시스(Lexis-Nexis)등의 상업 온라인 데이터베이스 시스템에서 단일의 데이터베이스를 검색하는 경우가 분리 인터페이스의 예라고 할 수 있다.

반면에 통합 인터페이스의 경우는 이용자가 질의어를 입력하면 시스템이 다수의 데이터베이스들을 동시에 검색한 후 개별적 데이터베이스에서 검색된 문헌들의 순위를 결정하고 이 순위에 근거해 검색결과를 통합한다. 결과적으로 이용자들은 시스템에 의해 통합된 결과판 얻게 된다. 이 인터페이스를 이용하는 경우, 이용자가 데이터베이스를 선택할 필요가 없고 개별적으로 데이터베이스를 검색할 필요가 없으므로 편리하고 효율적이라는 장점이 있다. 그러나 통합인터페이스의 어려움 중의 하나는 데이터베이스들로부터 검색된 문헌들의 순위를 결정하고 통합하는 기술개발이라고 할 수 있다. 통합 인터페이스의 예로는 메타크롤러(Meta Crawler), 쉐비 서치(Savvy Search), 인퍼런스 파인드(Inference Find!)등의 인터넷 메타 검색 엔진을 들 수 있다.

## 2 연구방법

### 2.1 연구설계, 데이터베이스, 검색과제, 검색 질문

두 시스템, 분리인터페이스와 통합인터페이스를 비교하기 위해 TREC(Text REtrieval Conference)의 데이터베이스와 검색과제, 검색 질문 중 일부를 이용하여 미국 Rutgers대학에서 실험이 실시됐다. TREC은 1992년부터 시작된 미국 내 정보검색 분야의 주요학회 중의 하나인데, 새로운 정보검색 시스템 및 기술개발과 평가를 주요 목적으로 하고 있다. (Voorhees and Harman, 1997). TREC에는 각종 연구기관, 정부기관, 학교기관, 기업체들이 참여하는데, 이들은 시스템간의 비교를 위해 공통의 전문(전문)데이터베이스와 공통의 검색질문을 이용하도록 요구된다. TREC안에는 여러 특수 분야가 존재하는데, 그 중 하나가 1996년부터 시작된 시스템과 이용자간의 상호작용(interaction)연구를 전문으로 하는 인터랙티브 트랙(Interactive track)이다. 제5회 TREC에서 interactive track은 보다 심층적인 상호작용 연구를 위해 새로운 정보 검색 과제 "aspectual task"를 제시하고, 그에 적합한 평가 기준도 개발하기 시작했다. 이 과제(aspectual task)는 이용자로 하여금 주어진 질문에 대해 가능한 한 많은 관점 혹은 측면(aspect)을 발견하고, 이러한 관점이 실린 문헌들을 저장하도록 요구한다(Over, 1997). 본 연구에는 TREC5(1996) Interactive track의 검색 과제와 검색 질문 중 일부가 이용되었다. 총 6개의 검색질문이 이용되었는데, 그 중 하나가 부록에 실려 있다. 또한, 본 연구에는 다음의 네 가지 전문 데이터베

이스가 이용되었다:

103회 국회기록기(Congressional Record of the 103rd Congress, 27, 922문헌, 1993);  
Federal Register(55, 630문헌, 1994);  
Financial Times(1991-1994, 210, 158  
문헌);  
Wall Street Journal(1990-1992, 74, 520  
문헌)

## 2.2 시스템

메사추세츠 대학의 Center for Intelligent Information Retrieval에서 개발된 확률검색(Probabilistic Retrieval)에 근거한 Inquiry 3.2 시스템이 본 연구의 기초시스템으로 이용되었다(Callan, Croft, & Harding, 1992). 이 시스템에 기반하여 Rutgers, School of Communication, Information, and Library Studies의 박사 과정 학생인 Shin-jeng Lin이 Tk/Tcl언어를 통해 두 가지 사용자 인터페이스를 개발하였다(Ousterhout, 1994; Welch, 1995). 이 두 인터페이스의 주요 특성은 자동 적합성 피이드백, 사용자 피이드백에 의한 질문 수정(manual query modification), 문헌순위 결정(document ranking), 전문 데이터베이스의 디스플레이(full-text display)등이다. 이 시스템들을 이용하는 사용자들은 불리언 질의문이 아닌 자연어 질의문(unstructured query)을 입력한다. 이용자들은 검색된 문헌들을 적합한 것으로 판정하여 적합성 피이드백에 이용할 수 있고, 적합하다고 판정된 문헌들을 저장할 수도 있다. 또한 이용자들은 검색된 문헌들의 전문을 디스플레이할 수 있고, 전체 문헌 중에서 시스템이 적합하다고 판정해 강조한(highlight) 부

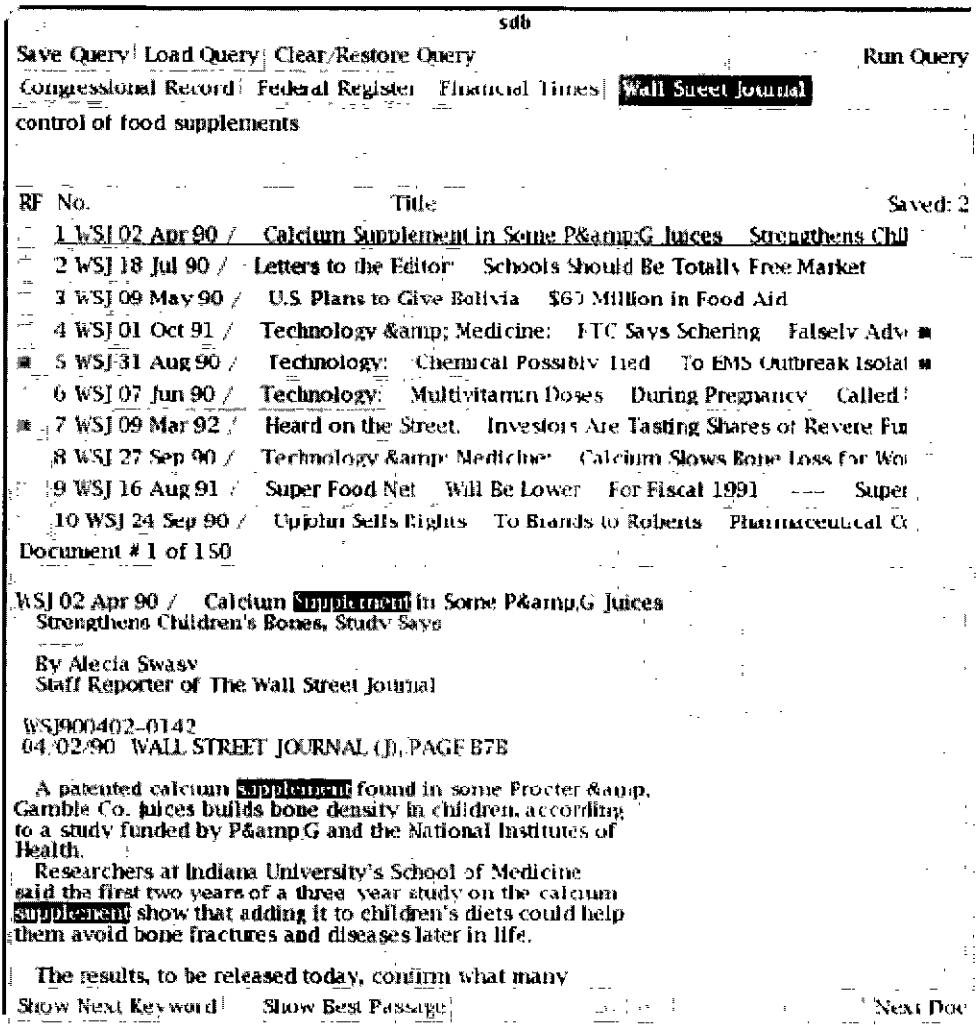
분들을 찾아갈 수도 있다.

분리 인터페이스의 경우 (<그림 1>, 원 논문에서는 허미스 Hermes 로 불리움), 이용자들은 공통의 인터페이스를 통해 상이한 4개의 데이터베이스를 개별적으로 사용하고, 상호작용 해야 한다. 이 인터페이스를 통해 이용자들은 다수의 데이터베이스들 중에서 자신이 원하는 데이터베이스를 선택할 수 있는데, 데이터베이스 선택을 용이하게 하기 위해 각 데이터베이스의 내용, 장르, 크기 등의 메타데이터가 tutorial동안 이용자들에게 제공되었다. 이 시스템을 사용시, 이용자들은 한번에 단일 데이터베이스 밖에 사용할 수 없다. 따라서 데이터베이스를 바꾸는 경우에는 검색과 상호작용을 반복해야 한다. 분리인터페이스의 또 다른 단점 중의 하나는 다수의 데이터베이스로부터 나온 검색결과를 이용자 스스로 비교하고 통합해야 한다는 점이다.

통합 인터페이스의 경우는 (<그림 2>, 원 논문에서는 헤라 Hera로 불리움) 이용자가 질의어를 입력하면 시스템이 다수의 데이터베이스들을 동시에 검색한 후 개별적 데이터베이스에서 검색된 문헌들의 순위를 결정하고 이 순위에 근거해 검색결과를 통합한다. 결과적으로 이용자들은 시스템에 의해 통합된 결과만 얻게 된다. 이 인터페이스를 이용하는 경우, 이용자가 데이터베이스를 선택할 필요가 없고 개별적으로 데이터베이스를 검색할 필요가 없다. 그러나 통합인터페이스의 어려움 중의 하나는 데이터베이스들로부터 검색된 문헌들의 순위를 결정하고 통합하는 기술개발이라고 할 수 있다.

## 2.3 조사 방법, 조사 도구

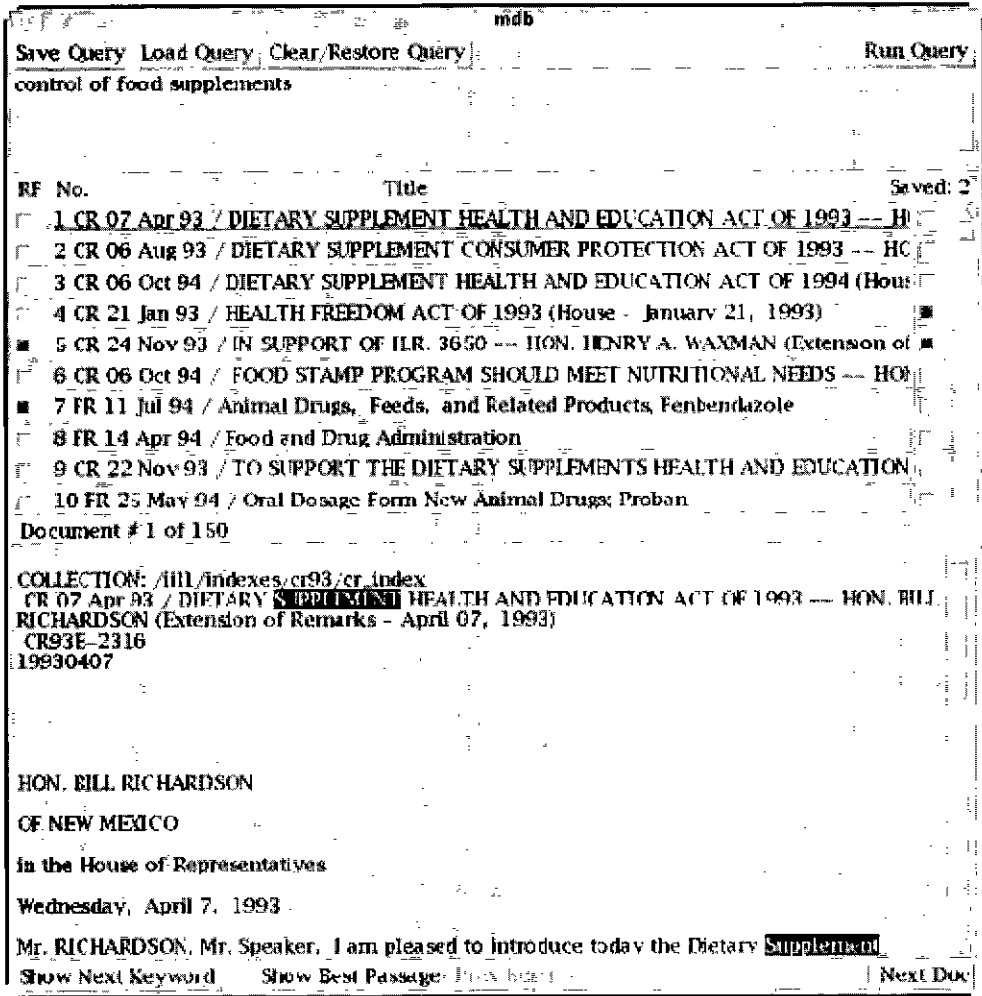
두 시스템, 분리 인터페이스와 통합 인터페이



〈그림 1〉 분리 인터페이스

스를 비교하기 위해 1998년 가을에 미국 Rutgers 대학에서 실험이 실시됐고, 이 실험에는 뉴저지 주립대학 Rutgers, School of Communication, Information & Library Studies에 재학중인 28명의 대학원생들이 자원하였다. 본 연구의 목적이 두 시스템을 직접 비교하는 것이었으므로, 모든 참가자들이 두 시스템을 사용하고 각 시스템마다 3개의 검색질문을

탐색하도록 하였다. 특정 시스템을 먼저 이용해서 발생할 수 있는 순서 효과(order effect)를 방지하기 위해 시스템 순서와 검색 질문 블록의 순서가 참가자마다 바뀌었다. 본 연구는 양적 연구 방법과 질적 연구 방법을 병용하였는데, 데이터 수집을 위한 조사도구로는 설문조사와 면접, 트랜잭션 로그가 이용되었다. 또, 참가자들은 실험에 응하는 동안의 생각을 소리내어 말



〈그림 2〉 통합 인터페이스

하도록 요청되었다(thinking aloud method).

각 검색 세션의 절차는 다음과 같았다:

- 실험동의서 서명
- 첫 시스템에 대한 tutorial (20분 이내)
- 3개의 검색 질문 탐색 (각 질문마다 최대한 20분 허용)
- 매 질문 탐색 후 설문조사
- 첫 시스템에 대한 설문조사
- 휴식

· 두 번째 시스템에 대한 tutorial

· 3개의 검색 질문 탐색 (각 질문마다 최대한 20분 허용)

· 매 질문 탐색 후 설문조사

· 두 번째 시스템에 대한 설문조사

· 인적 사항을 묻는 설문조사

· 종료 설문조사

· 종료 면접

### 3 연구의 결과

#### 3.1 이용자 선호도

종료설문지와 면접에서 주어진 “당신은 어느 시스템을 선호하십니까?”라는 질문에 대해 29퍼센트의 참가자들(8명)이 통합인터페이스를 선호한다고 대답하였고, 64퍼센트의 참가자들(18명)이 분리인터페이스를 선호한다고 대답하였으며, 나머지 2명은 선호도에 차이가 없다고 대답하였다. 이용자 선호도에 관련된 가설1을 검증하기 위해 카이제곱법이 사용되었는데, 두 시스템에 대한 이용자의 선호도에는 통계적으로 유의한 차이가 나타났다.  $\chi^2 (1, N=26) = 3.846, p \leq .05$ . 이러한 선호도의 차이를 분석하기 위해 종료인터뷰에 대한 내용분석이 행해졌다. 참가자들이 특정 시스템을 선호한 이유를 정리하면 <표 1>과 같다.

이용자들이 통합 인터페이스를 선호하는 이유는 다양하였는데, 첫째, 시간과 노력이 덜 든다는 점에서 편리하고 효율적(efficient)이라고 응답하였다. 둘째, 몇몇 이용자들은 데이터베이스를 선택할 필요가 없다는데서, 분리인터페이스보다 사

용하기가 더 용이하다고 응답하였다. 셋째, 통합 인터페이스 사용 시 검색결과가 좀 더 포괄적이고 광범위하다고 응답한 이용자들도 있었다. 또한 한 이용자는 일반적인 탐색에서 시작하여 좁혀 나가는 자신의 평소시의 탐색습관 때문에 통합 인터페이스를 선호한다고 응답하였다.

이러한 통합 인터페이스의 장점에도 불구하고, 더 많은 이용자들이 분리 인터페이스를 선호하였는데, 이들이 분리 인터페이스를 선호한 주된 이유는 데이터베이스를 이용자 스스로 선택하고, 통제할 수 있기 때문인 것으로 나타났다. 즉 이들은 주어진 검색 질문탐색 시 자신들이 원하는 데이터베이스에만 집중하고, 부적절한 데이터베이스를 제외하기를 원하였다.

#### 3.2 이용자 만족도, 이용 편의성

탐색과정과 결과에 대한 이용자 만족도와 이용편의성을 평가하기 위해 참가자들이 각 검색 질문을 마칠 때마다 다음의 질문이 주어졌다.

- 당신의 탐색결과에 얼마나 만족하십니까?
- 이 질문에 대한 검색과정이 얼마나 용이했습니까?

<표 1> 분리인터페이스와 통합인터페이스에 대한 이용자반응 비교

시스템	장 점	단 점
통합 인터페이스	<ul style="list-style-type: none"> <li>· 시간과 노력이 덜 든다는 점에서 편리하고 효율적이다.</li> <li>· 사용하기가 더 용이하다</li> <li>· 검색결과가 좀 더 포괄적이고 광범위하다.</li> <li>· 평소시의 탐색습관과 관련되어 있다 (일반적인 탐색에서 시작하여 좁혀 나간다.)</li> </ul>	<ul style="list-style-type: none"> <li>· 데이터베이스를 선택할 수 없다.</li> <li>· 부적절한 데이터베이스가 탐색에 방해가 된다.</li> </ul>
분리 인터페이스	<ul style="list-style-type: none"> <li>· 데이터베이스 선택시 통제가 가능하다.</li> <li>· 이용자가 원하는 데이터베이스에만 집중할 수 있다.</li> <li>· 부적절한 데이터베이스를 제외시킬 수 있다.</li> </ul>	<ul style="list-style-type: none"> <li>· 불편하고 비능률적이다.</li> <li>· 시간이 많이 걸린다.</li> </ul>

이 설문문항은 리커트식 7단계 척도(7 point Likert Scale)로 설계되었다.

가설 2.3을 검증하기 위해 종속표본의 t-검증(dependent samples t-test)이 사용되었다. 통계 분석 결과, 분리 인터페이스와 통합인터페이스에 대한 이용자의 만족도에는 통계적으로 유의한 차

이가 있었다,  $t(27)=2.117, p \leq 0.05$ . 다시 말해서 참가자들은 통합인터페이스보다 (평균치=3.71) 분리인터페이스(평균치=4.37)의 탐색결과에 훨씬 더 만족하였다. 그러나 이용편의성에 대한 이용자 인식에서는 유의적 차이가 없었는데, 이는 두 시스템이 적합성 피드백과 문헌순위 결정 등 상당수의 특성을 공유하고 있다는 사실로 설명될 수 있다.

### 3.3 검색 효율성

앞에서도 언급되었듯이, 본 연구에서는 평가 척도로 TREC의 aspectual recall (이용자 관점 기반 재현률)이 도입, 사용되었다. Aspectual recall(이용자 관점 기반 재현률)은 제5회 TREC의 interactive track에서 최근에 개발된 평가척도인데, 이 척도가 이용된 이유는 첫째, 제5회 interactive track의 검색과제, 검색질문, 데이터베이스들이 이 실험에 이용되었고, 재현율이나 정확율과 같은 기존의 평가 기준이 상호작용 연구에 초점을 둔 이 실험의 검색과제에는 적합하지 않기 때문이다. 이용자 관점 기반 재현률은 주어진 검색질문마다 이용자가 파악한 관점 (aspect)의 수효와 TREC 평가자들이 파악한 관점의 수효간의 비율로 계산된다 (Belkin et al., 1997). 예를 들어, 만약 주어진 검색질문에 대해 TREC 평가자들이 파악한 관점의 총수

가 10이고, 이용자가 파악한 관점의 총수가 5라면, 이 경우의 이용자 관점 기반 재현률은 0.5가 된다. 이용자 관점 기반 재현률의 가치는 0과 1 사이에 분포한다

이 연구에서 이용자들이 분리 인터페이스 사용시의 평균 이용자 관점 기반 재현률은 (평균=0.31, 분산=0.16) 통합 인터페이스 사용시의 평균 이용자 관점 기반 재현률 (평균=0.22, 분산=0.12)보다 높았다. 검색 효율성에 관한 연구과제 4를 검증하기 위해 종속표본의 t-검증(dependent samples t-test) 이 사용되었고, 이 차이는 통계적으로 유의하였다.  $t(27)=2.485, p \leq 0.05$

〈표 2〉 검색 효율성 비교

	평균	분산	t	p≤
통합 인터페이스	0.2164	0.1159	2.485	0.019
분리 인터페이스	0.3076	0.1573		

### 3.4 정보 검색 행위

이용자와 시스템간의 모든 상호작용이 트랜잭션 로그에 기록되었고, 이 섹션에서는 이렇게 기록된 이용자들의 정보 검색 행위가 다양한 시스템 기능의 이용 빈도의 측면에서 분석되었다. 시스템기능의 이용 빈도를 비교하기 위해 일련의 종속표본의 t-검증(dependent samples t-test) 이 사용되었고, 이 결과가 〈표 3〉에 정리되었다.

이 결과 중 통계적으로 유의한 결과를 일부 요약하자면, 이용자들은 통합 인터페이스에서 보다 분리 인터페이스에서 더 많은 문헌을 저장하였고, 더 많은 수의 전문(全文)을 보았으며,



(표 3) 시스템 기능의 이용 비교

	평균	분산	t	p≤
저장된 문헌수 통합 인터페이스 분리 인터페이스	8.3095 11.8690	4.0885 6.2026	3.777*	0.002
저장이 취소된 문헌수 통합 인터페이스 분리 인터페이스	0 0.6190	0 0.8350	3.923*	0.001
이용자가 본 전문수 통합 인터페이스 분리 인터페이스	24 1786 32.2500	8.4388 14.2325	2.728*	0.011
질의문 검색실행 (iteration) 통합 인터페이스 분리 인터페이스	6.6667 9 0833	3.2049 3.4541	3.880*	0.001
질의문 저장하기 통합 인터페이스 분리 인터페이스	0 3810 1.0714	0.9329 1.1488	4.112*	0.001
저장된 질의문 재생하기 (load query) 통합 인터페이스 분리 인터페이스	0 1.1310	0 1.2905	4 637*	0.001
적합성 피드백을 이용한 질의문 수 통합 인터페이스 분리 인터페이스	3.6071 4.9761	2.3675 3.1647	2.429*	0.022
이용자가 입력한 총 질의어 수 통합 인터페이스 분리 인터페이스	30.6429 44.5952	21.8774 21.2731	3.376*	0.002
이용자가 입력한 개별 질의어 수(unique terms) 통합 인터페이스 분리 인터페이스	11.1429 10.6905	4.8385 3.8787	0.511*	0.613
이용자가 적합하다고 표시한 문헌 수 통합 인터페이스 분리 인터페이스	5.3690 6.5357	4.1130 6.6958	1.137*	0 266
총 소요시간 통합 인터페이스 분리 인터페이스	945 7257 989 7495	217.1359 196.5765	1.000*	0.326
다음 키워드로 가기(next keyword) 통합 인터페이스 분리 인터페이스	1.8333 1.4762	4.8113 3.3450	0.476*	0.638

더 많은 질의문을 실행하였고, 분리 인터페이스에서 적합성 피드백을 더 자주 사용하였다. 두 시스템 모두 "질의문 저장하기(save query)"와 "저장된 질의문 재생하기(load query)"기능을 제공하였는데, 이용자들은 분리 인터페이스에서 이 두 기능을 더 자주 사용하였고, 이 차이는 통계적으로 유의하였다. 흥미롭게도, 분리 인터페이스에서는 "저장된 질의문 재생하기"기능이 평균 1.13회 이용되었으나, 통합 인터페이스에서는 한 번도 이용되지 않았다. 분리 인터페이스에서 이용자들은 데이터베이스를 바꿀 때마다 같은 상호작용을 반복하려는 경향을 보였는데, 실제로 종료 인터뷰에서 대다수의 참가자들이 동일한 질의문을 다른 데이터베이스에도 반복하였다고 응답하였다. 이러한 결과는 분리 인터페이스가 이용자들로부터 더 많은 행위적 노력과 (behavioral effort) 상호작용을 요구하였고, 이용자들이 두 시스템간의 차이와 시스템의 다양한 특성들을 잘 인식하고 있었고, 시스템을 바꿀 때에는 시스템의 기능을 다르게 이용함으로써 새로운 시스템에 적응하려 노력하였다는 점을 암시한다.

#### 4 결 론

본 연구는 이용자가 분산환경에서 다수의 데이터베이스들을 어떻게 이용하고 탐색하는가를 분석하기 위해 두 개의 다른 인터페이스를 비교하였다. 좀 더 구체적으로, 이용자가 데이터베이스들을 개별적으로 검색하는 경우와 다수의 데이터베이스로부터 나온 검색결과가 시스템에 의해 통합되어 이용자에게 전달되는 경우, 이용자 선호도와 만족도, 검색 효율성과 정보 검색행위가 비교되었다. 조사결과, 많은 참가자들이 통합인터페

이스보다 분리인터페이스를 선호하였고 분리인터페이스의 검색결과에 더 만족한 것으로 드러났다. 또 이 이용자 그룹에게는 분리 인터페이스의 검색 효율성이 통합 인터페이스의 검색 효율성보다 높은 것으로 나타났다. 이용의 편의성에 있어서는 두 시스템간에 유의한 차이가 없었다.

통합인터페이스가 편리하고 효율적이라는 장점에도 불구하고, 참가자들이 분리인터페이스를 선호한 주된 이유중의 하나는 데이터베이스 선택시 통제가 가능하기 때문인 것으로 나타났다. 비록 많은 참가자들이 분리인터페이스를 선호하였지만 분리인터페이스의 통제와 통합인터페이스의 편리함을 결합한 시스템 개발을 희망하였다. 다시 말해서 이용자는 데이터베이스 선택을 통제하기 원하지만, 일단 자신들이 원하는 데이터베이스를 선택한 후에는 시스템이 이들 데이터베이스들로부터 나온 결과를 통합하는 것이 바람직하다는 것이다. 이러한 결과는 전자 도서관 환경에서 시스템 개발시 명목적으로 데이터베이스들을 통합하는 것보다는 데이터베이스 통합에 대한 선택권을 이용자에게 부여하는 것이 중요하다는 점을 암시한다. 이용자의 데이터베이스 선택을 용이하게 만들기 위해서는 메타데이터 등의 이용을 통해 개별 데이터베이스의 특성을 잘 묘사하는 것이 중요할 것이다.

본 연구는 표본의 크기와 표집 방법에 있어서 몇 가지 제한점을 가지고 있다. 실험과정이 세 시간 이상 걸린다는 등의 실제적 이유로 인해 참가자 선정이 무작위로 이루어지지 않았으며, 표본의 크기도 30명 정도로 적은 편이다. 또한 참가자 대부분이 교육수준이 높고 정보시스템 검색에 풍부한 경험을 가진 집단이므로 이 조사결과를 다른 집단에 일반화하기는 어렵다. 따라서 실험 환경이 아닌 실제환경에서 표본크기증가와 무작

위 표집을 통한 지속적 연구가 향후과제라고 할 수 있다.

이러한 제한점에도 불구하고 본 연구는 분산 환경에서 이용자와 정보 시스템 간의 상호 작용에 대한 이해를 넓히고 전자 도서관 환경에서의 시스템 개발에 시사점을 제시한다는 데 그 의의를 찾을 수 있을 것이다. 또한 본 연구는 데이터베이스 통합이 검색 효율성을 향상시키고, 이용

자들이 데이터베이스 통합을 선호할 것이라는 정보 검색 분야의 일반적 전제와 상반되는 결과를 제시함으로써, 정보 검색 분야, 특히 데이터베이스 머징 분야에 널리 퍼져있는 이러한 인식이 재검증 받아야 할 필요가 있다는 점을 제시하였다는 데 또 다른 의의가 있다고 할 수 있다.

## 참고 문헌

- Belkin, N.J., Kantor, P., Cool, C., & Quatrain, R. (1994). "Combining evidence for information retrieval." In D. Harman, (Ed.), *TREC-2 proceedings of the second text retrieval conference* (pp.35-44). Washington, DC: Government Printing Office.
- Belkin, N.J., Kantor, P., Fox, E.A., & Shaw, J.A (1995). "Combining the evidence of multiple query representations for information retrieval." *Information Processing & Management*, 31(3), 431-438.
- Belkin, N.J., Cabezas, A., Cool, C., Lin, S., Park, S., Rieh, S., Savage, P., Sikora, C., Xie, H , & Allan, J. (1997). "Rutgers TREC-5 interactive track experience." In D. Harman. (Ed ), *TREC-5 proceedings of the fifth text retrieval conference* (pp.257-266). Washington, DC: Government Printing Office.
- Callan, J.P., Croft, W.B., & Harding, S. M. (1992). "The INQUERY retrieval system." In *DEXA 3: Proceedings of the third international conference on database and expert systems applications* (pp. 83-87), Berlin: Springer Verlag.
- Callan, J.P., Lu, Z., & Croft, W.B. (1995). "Searching distributed collections with inference networks." In E.A. Fox, P. Ingwersen. R. Fidel (Eds.), *SIGIR'95: Proceedings of the 18th annual international ACM SIGIR conference on research and development in information retrieval* (pp.21-28). New York: ACM.
- Fox, E.A., Shaw, J.A. (1994). "Combination of multiple searches." In D Harman (Ed.), *TREC-2, proceedings of the second text retrieval conference* (pp. 243-252). Washington, D.C.: Government Printing Office.

- Hawking, D., Thistlewaite, P., & Bailey, B. (1997). "ANU/ACSts TREC-5 experiments." In D. Harman (Ed.), *TREC-5, Proceedings of the fifth text retrieval conference* (pp.359-375). Washington, D.C: Government Printing Office.
- Lee, J. (1995). "Combining multiple evidence from different properties of weighting schemes." In E.A. Fox, P. Ingwersen, & R.Fidel (Eds.), *SIGIR' 95: Proceedings of the 18th annual international ACM SIGIR conference on research and development in information retrieval* (pp.180-188). New York: ACM.
- Lu, Z., Callan, J.P., & Croft, W.B. (1996). *Applying inference networks to multiple collection searching* (Umass Technical Report TR96-42). University of Massachusetts, MA: Center for Intelligent Information Retrieval.
- Ousterhout, J. (1994). *Tcl and the Toolkit*. Reading, PA: Addison-Wesley.
- Over, P. (1997). "TREC-5 interactive track report." In D. Harman (Ed.), *TREC-5, Proceedings of the fifth text retrieval conference* (pp.29-56). Washington, D.C: Government Printing Office.
- Park, S. (1999). *Supporting Interaction with Distributed and Heterogeneous Information Resources*. Unpublished doctoral dissertation.
- Savoy, J., Le Calve, A., & Vrajitoru, D. (1997). "Report on the TREC-5 experiment: Data fusion and collection fusio." In D. Harman (Ed.), *TREC-5, Proceedings of the fifth text retrieval conference* (pp.359-375). Washington, D.C: Government Printing Office.
- Schiettecatte, F (1997). "Document retrieval using the MPS information server (A report on the TREC-5 experiment)." In D. Harman (Ed.), *TREC-5, Proceedings of the fifth text retrieval conference* (pp.391-404). Washington, D.C.: Government Printing Office.
- Voorhess, E.M., Gupta, N. K., & Johnson-Laird, B. (1995). "Learning collection fusion strategies." In E.A. Fox, P. Ingwersen, & R.Fidel (Eds.), *SIGIR '95: Proceedings of the 18th annual international ACM SIGIR conference on research and development in information retrieval* (pp.180-188). New York: ACM.
- Voorhess, E.M. (1997). The TREC-5 database merging track. In D. Harman (Ed.), *TREC-5, Proceedings of the fifth text retrieval conference* (pp. 103-104). Washington, D.C.: Government Printing Office.
- Voorhees, E.M., & Harman, D. (1997). "Overview of the Fifth Text Retrieval

- Conference(TREC-5)." In Harman, D. (Ed.), *TREC-5, Proceedings of the Fifth Text REtrieval Conference* (pp.1-28). Washington, DC: Government Printing Office.
- Welch, B. (1995). *Practical Programming in Tcl and TK*. Prentice Hall.
- Xu, J., & Callan, J. (1998). "Effective retrieval with distributed collections." In A. Moffat, & J. Zobel (eds.). *SIGIR 98: Proceedings of the 21th annual international ACM SIGIR conference on research and development in information retrieval* (pp.112-120). New York: ACM.

## 부록: 검색주제 274i, 전기 자동차

선생님이 General Motors사에 재직하고 있고, 전기 자동차에 대한 과제 (project)를 막 시작했다고 가정하십시오. 선생님께서는 전기 자동차분야에서 최근에 이루어진 다양한 종류의 개발을 조사한 연구 논문을 쓰도록 요청 받으셨습니다. 20분내에 이 목적에 유용한 문헌들을 찾아 저장하셔야 합니다. 전기 자동차분야에서 "최근에 이루어진 서로 다른 종류의 개발" (each different recent development)을 다룬 문헌을 적어도 하나 저장하여 주십시오. 만약 한 문헌이 여러 종류의 개발을 다루고 있다면, 똑같은 종류의 개발을 다루고 있는 다른 문헌들을 저장하실 필요가 없습니다.