이용자 중심의 국내 탐색엔진의 평가에 관한 연구

A User-Centered Evaluation of Naver and Korea AltaVista

박희진, 오삼균, 성균관대학교 문헌정보학과

Park, Hee-jin and Oh, Sam-kyun
Dept. of the Library and Information Science, Sung Kyun Kwan Univ.

본 연구는 이용자의 정보요구를 최대한 충족시키고 보다 사용하기 편한 이용자 지향의 탐색엔진을 개발하기 위해서 이용자 관점에서 국내 탐색엔진 한글알타비스타와 네이버를 평가하고자 한다. 이용자 중심의 탐색엔진 평가를 위해, 본 연구는 이용자가 자신의 검색질의로 탐색엔진을 통해 직접 검색을 하게 한 후 검색결과에 대해 이용자 스스로 평가하도록 한다. 평가기준은 수와 첸의 탐색엔진의 평가기준을 바탕으로 하여 탐색시간 및 노력, 탐색결과의 정확율, 연결성, 중복률, 유용성, 이용자의 만족도로 한글알타비스타와 네이버의 검색효율성을 비교·평가하고자 한다.

1. 연구의 필요성 및 목적

정보의 보고라고 불리는 인터넷은 현대사회의 중요한 정보원으로써, 정보기술의 발달로 대중화되어 어느 누구나 쉽게 접근하여 사용할 수 있다. 그러나 인터넷에 무수히 산재된 정보는 비조직적이고 비체계적으로 이루어져서 이용자들이 원하는 정보를 정확하고 신속하게 찾는 것이 쉽지 않다. 따라서 이러한 인터넷의 정보를 효과적으로 수집, 활용하도록 도와주는 검색도구인 탐색엔진(search engine)들의 역할이 더욱 중요시되고 있다.

국내에서도 인터넷 이용자가 급증하고 인터넷 시장의 가치가 높아짐에 따라 기존의 국내탐색엔진들도 다양한 부가 서비스를 개발하고 차별 화된 기능을 선보이는 등 검색엔진의 향상에 힘을 쏟고 있다. 또한 야후(Yahoo)와 라이코스(Lycos), 알타비스타(AltaVista) 등 높은 인지도를 보이고 있는 외국 탐색엔진들이 국내로 진출하고 있어서 탐색엔진이 계속적으로 개발되지 않으면 안 되는 현실이다.

탐색엔진을 보다 향상시키고 개발하는 일은, 탐색엔진의 평가를 바탕으로 이루어져야 한다. 탐색엔진의 성능을 비교하고 평가함으로써, 설계자는 이용자가 실제 탐색엔진을 어떻게 사용하고 있으며 자신의 탐색엔진의 장점과 단점이 무엇인지를 파악할 수 있기 때문이다. 또한 정보전문가는 탐색엔진의 평가를 통해 제공한 정보서비스를 이용자가 어떻게 활용하는지 파악해 보다 효과적인 정보를 제공할 수 있으며, 이용자는 탐색엔진의 평가를 통해 자신의 정보요구에 적합한 탐색엔진을 선택하여 효율적으로 검색을 할 수 있으므로 탐색엔진 설계자 뿐 아니라 정보전문가나 인터넷을 사용하는 최종이용자들에게도 탐색엔진의 평가는 매우 중요한 일이다.[1] 이러한 취지에서 탐

1..Mccure, C. "User-based Data Collection Techniques and

색엔진을 비교, 평가하는 연구가 국내외에서 꾸준히 활발하게 이루어지고 있다. 탐색엔진의 평가에 관한 연구들은 각 탐색엔진이 제공하는 탐색기능, 데이터베이스 구성 방법들로 성능을 평가하는 연구와 검색질의로 검색한 후 검색결과로 탐색엔진을 비교·평가하는 연구가 주를 이룬다. 그러나, 이용자의 정보요구를 최대한 충족시키고 보다 이용자가 사용하기 편한 이용자 지향의 탐색엔진을 개발하기 위해서는 탐색엔진을 실제 사용하는 이용자의 현재 정보요구를 해결하고 이용자가 검색결과에 만족하는지 등 이용자 관점에서 탐색엔진을 평가한 연구가 필요하다.

따라서, 본 연구는 수와 첸의 이용자 중심의 평가기준을 바탕으로 하여 국내 탐색엔진 한글알타비스타와 네이버의 검색효율성을 비교·평가하고자 한다. 국내 탐색엔진의 검색효율성을 탐색시간 및 노력, 탐색결과의 정확율, 연결성, 중복률, 유용성, 이용자의 만족도로 비교·평가함으로써 국내 탐색엔진의 기능을 향상시키고, 나아가 이용자에게 적합한 국내 웹 탐색엔진을 개발하는데 기여하고자 한다. 아울러 이용자 중심의 탐색엔진 평가방법을 제시함으로써 앞으로 본 연구가 탐색엔진의 지속적인 평가에 활용되길 기대한다.

2. 국내외 연구동향과 연구배경

1) 국내외 탐색엔진 평가연구의 동향

국내 탐색엔진에 관한 연구들은 대부분 연구자가 임의로 만든 검색질의나 예전에 사용되었던 참고봉사 질의를 사용하여 연구자가 검색결과에 대해 검색질의에 적합한지 판단하여 정확율과 재현율로 평가한 것들로, 실제 탐색엔진을 사용하는 이용자의 현재 정보요구에 충족하고, 이용자가 검색결과에 만족하고, 사용하는데 불편함은 없는지, 이용자의 관점이 결여된 시스템위주의 탐색엔진에 대한 평가가 주를 이루고 있다.

첫째, 기존의 탐색엔진의 평가에 관한 연구는 외국 탐색엔진을 주로 다루고 있다. 야후, 알타비스타, 마젤란, 익사이트 등 외국 탐색엔진에 대한 연구는 국내 뿐 아니라 외국에서도 활발히 이루어지고 있으나, 국내 탐색엔진을 평가한 연구는 그리 많지 않다. 국내 탐색엔진을 다룬 연구는 탐색엔진의 검색기법 및 데이터베이스 구성방법 등 소개 형식에 머물고 있으며 이용자를 실험에 실제로 참여하게 하여 탐색엔진에 대한 검색효율을 비교하고 평가한 연구는 거의 없다.

둘째, 선행연구의 탐색엔진의 평가에 사용된 검색질의는 이용자의 현재 정보요구를 대변하는 실제질의(real question)와 차이가 있을 수 있다. 기존의 탐색엔진의 평가에 사용되었던 검색질의는 과거 이용자의 참고 봉사시 사용되었던 탐색질의나 연구자 임의로 선정한 탐색질의들로, 이를 가지고 평가한 결과는 이용자가 현재 정보검색을 필요로 하는 당면한 실제 질의를 가지고 탐색하여 평가하는 것과 차이가 있을 것이다.

셋째, 선행연구의 탐색엔진 평가는 이용자보다는 연구자에 의해 대부분 이루어졌다. 기존의 연구들은 평가의 주체가 연구자인 문헌정보학 전문가나 컴퓨터 학자로 이용자가 아닌 연구자들이 검색결과에 대한 적합성 여부를 판정하고, 이 적합성을 바탕으로 한 재현율, 정확율을 중요한 평가기준으로 삼고 있어서 실제 탐색엔진을 통해 정보탐색을 하는 이용자의 평가와는 거리가 있다.

이용자 중심의 탐색엔진의 평가에서는 평가에 사용되는 검색질의가 이용자의 현재 정보요구를 대변하는 실제 질의여야 하며, 검색효율 평가를 위한 적합성 판정의 주체가 연구자가 아닌 이용자 스스로에 의해 이루어져야 이용자에 맞는 웹 검색엔진 개발에 보다 도움이 될 수 있을 것이다.

2) 수와 첸의 이용자 관점의 탐색엔진 평가연구

수와 첸은[2] 탐색엔진에 대해 이용자가 스스로 질의를 가지고 검색하게 한 후, 이용자 중심의 탐색엔진 평가에 대해 다음과 같은 5가지 기준들을 제시하

Strategies for Evaluating Networked Information Services," *Library Trends* 4 (1994): 591-607.

2. Su, Louise T. et al. "Evaluation of Web-Based Search Engines from the End-User's Perspective: A Pilot Study," *ASIS'98 Proceedings of the 61st ASIS Annual Meeting* 35 (1998): 348-361.

였다. 첫 번째 기준은 적합성 판단 시간을 포함한 검색에 소요된 전체 검색시간, 두 번째 기준은 검색결과 문헌을 접속하였을 때의 데드링크(deadlink)된 사이트의 수를 알아보는 연결성(connectivity), 세 번째 기준은 이용자가 자신의 질의에 대해 검색결과 문헌이 적합한지 알아보는 적합성(relevance), 네 번째 기준은 시스템의 응답속도, 인터페이스, 온라인도움문서, 결과물디스플레이에 대한 이용자의 만족도(satisfaction)와 정확율, 검색에 대한 전체적인 성공도를 알아보는 것이며, 마지막으로 검색결과가 정보요구에 어떤 가치가 있는가 하는 유용성(utility)을 평가기준으로 내세우고 있다.

이는 탐색엔진 평가에 대한 양적인 연구와 함께 질적인 연구인 이용자 관점의 평가를 실제로 시도했다는데 의의가 있는 것으로, 직접 이용자가 평가에 참여하여 검색질의 뿐 아니라 검색결과에 대해 적합성을 판정하고, 정확율을 측정하도록 하고 있어 이용자가 참여한 체계적인 탐색엔진의 평가라 할 수 있다.

3. 연구의 방법

1) 연구디자인

본 연구에서 주 독립변인은 두 탐색엔진, 한글알타비스타와 네이버이며 종속변인은 정확율, 탐색시간 및 노력, 이용자의 만족도로 하여 비교·평가한다.

실험은 검색자간의 오차를 줄이기 위해 한 피실험자가 동일한 검색질의로 두 탐색엔진 한글알타비스타와 네이버 모두를 이용해 검색을 하도록 한다. 실험에 참가하는 피실험자는 첫 번째 검사의 경험이 같은 내용의 두 번째 검사의 결과에 영향을 미치는 검사효과를 통제시키기 위해 임의로 두집단으로 나눈다. 한 집단은 한글알타비스타로 자신의 검색질의에 대해 실험을 시작하고, 실험이 끝나면 네이버로 같은 검색질의로 실험을 반복하게 한다. 다른 한 집단은 이와 반대로 네이버로 실험을 시작하고, 실험이 끝나면 한글알타비스타로 검색을 반복하게 한다.

피실험자들은 각자 자신들이 인터넷을 통해 찾고자 하는 검색질의를 가지고 탐색을 실행한다. 피실험자들은 실험 전에 한글알타비스타와 네이버의 검색 기능에 대해 충분히 익힌 후, 본 실험에 임하게 한다.

2) 독립변인-탐색엔진의 선정과 특성

국내에서 한글의 고유한 특성을 파악해서 유용한 웹 문서들을 검색하여 제공하는 인터넷 탐색엔진은 모두 15여개가 있다. 최근에는 탐색엔진들간의 검색기능과 네티즌의 눈길을 끌기 위한 부가서비스 개발 경쟁이 대우 치열하다.

한글알타비스타(www.altavista.co.kr)는 알타비스타 개발사인 디지털로부터 한글판 라이센스 계약을 맺은 평창정보통신이 개발, 운영하는 탐색엔진으로 영문 알타비스타와 모든 지원 메커니즘이 동일하며, 빠른 검색속도와 막대한 검색량을 제공한다는 것이 장점이다.[3]

네이버(www.naver.com)는 삼성 SDS 정보기술연구소의 검색엔진 팀인 웹글라이더팀에서 사내 벤처 프로그램으로 개발한 검색서비스로, 300만 건이 넘는 많은 국내 데이터 색인과 이미지, 사운드, 애니메이션 등 멀티미디어 정보검색 서비스에 강점이 있다.[4]

본 연구는 탐색엔진의 성격과 검색기능을 고려해서 이들 국내 인터넷 탐색엔진 중에서 한글알타비스타와 네이버를 그 대상을 선정하여 실제 이용자들의 탐색을 통해 그 성능을 비교·평가하고자 한다.

한글알타비스타와 네이버는 모두 주제 디렉토리 검색과 함께 키워드 검색을 지원하는 통합형 탐색엔진으로, 다른 탐색엔진에 비해 키워드 검색기능이 비교적 강하고 보다 정확한 검색결과를 지향하는 탐색엔진이다. 한글알타비스타와 네이버는 거의 모든 연산자를 지원하고 있으며, 필드검색, 멀티미디어검색, 2차검색 등 대부분의 검색기능이 서로 매우 유사하다.

3) 종속변인

가. 탐색시간 및 노력

실험자가 피실험자의 탐색과정에서 방문한 사이트의 수와 탐색에 소요된 총 시간을 상세히 기록하여 이를 분석한다.

---

3. PCLine편집부. "한글 검색엔진 6종," PCLine(1998년 7월호): 200-213.

4. 한글과 컴퓨터. "야후코리아, 네이버, 심마니, 정보탐정 4개 검색엔진 비교," *인터넷* (1999년 1월호): 120-139.

나. 정확율

피실험자에게 탐색결과가 적합한지, 출력순위 상위 20건에 대한 간단한 설명으로만 적합한지 판단하게 하고, 이를 실제로 클릭해서 해당 페이지에 도달하여 그 문헌의 내용을 보고 적합한 것인지, 두 가지 방법으로 적합성을 판단하게 한다

다. 연결성

실험자가 검색결과의 상위 20건에 대해 데드링크(deadlink)된 사이트의 수를 조사한다.

라. 중복률

실험자가 검색결과의 상위 20건에 대해 동일한 사이트나 문서가 중복되어 검색된 사이트의 수를 조사한다.

마. 유용성

실험이 끝난 후, 피실험자에게 면담과 질문지를 통해 검색의 성공도, 검색질의가 해결되었는지, 검색결과에 대한 유용성을 조사한다.

바. 만족도

실험이 끝난 후, 피실험자에게 면담과 질문지를 통해 탐색엔진의 결과물 디스플레이와 인터페이스에 대한 만족도를 조사한다.

4) 피실험자의 선정

본 연구에서 선정된 피실험자들은 성균관대학교 학생으로 온라인탐색 과목을 이수한 학생으로 구성되어 있으며, 전공학과는 따로 제한하지 않는다. 신청서를 배포하고 내용을 설명한 후 흥미가 있고 참여의사가 있는 사람이 자발적으로 참여하도록 한다. 가능한 한 동질적인 피실험자들을 표집하기 위해서, 탐색엔진을 사용해 본 경험이 있는 사람을 대상으로 하며, 이들은 각자 인터넷을 통해 찾고자 하는 자기의 실제 탐색 질의를 가지고 실험에 임한다.

5) 연구의 절차

가. 피실험자는 실험내용과 절차에 대해 간단한 설명을 받는다. 피실험자에게 네이버와 한글알타비스타의 도움말로 구성된 간단한 도움자료를 제공하고, 탐색할 때 이를 참조할 수 있도록 한다.

나. 피실험자는 자신의 검색질의를 밝히고, 두 탐색엔진 중 하나로 검색을 수행한다.

다. 실험자는 피실험자의 검색질의와 피실험자가 사용한 탐색용어, 연산자, 탐색시간, 탐색단계를 빠짐없이 기록한다.

라. 검색결과가 출력되면 피실험자는 출력건수 상위 20건에 대해 출력된 간단한 내용을 읽고 적합성을 판단하고 실험자는 이를 기록한다.

마. 피실험자는 '라'의 결과를 실제로 클릭 하여 해당 사이트의 내용을 읽어보고, 검색목적에 적합한지 판단하고 실험자는 이를 기록한다.

바. 실험자는 각각 '라'와 '마'의 결과에 대해 자세히 기록하고 후에 이를 분석한다.

사. 피실험자들은 성별, 학력, 전공, 주제에 대한 지식, 연령 등 신상과 만족도에 대한 질문지에 답하도록 한다.

아. 실험자는 검색결과 20건에 대해 데드링크된 사이트의 수와 동일한 사이트나 문서가 중복된 사이트의 수를 조사한다.

자. 한 탐색엔진에 대한 실험이 끝나면 피실험자는 다른 탐색엔진을 통해 같은 검색질의로 실험을 반복한다.

4. 연구결과의 기대효과 및 활용

1) 학문적 기여

이용자의 검색질의에 대해, 직접 검색을 하고 그 검색결과에 대해 평가하는 이용자가 직접 참여한 탐색엔진의 평가연구는 연구자에 의해 계획된 질의로 검색을 하고 그 결과를 평가하는 다른 연구와 어떤 차이가 있는지를 밝혀줄 것이다. 또한 이용자 중심의 탐색엔진 평가에 대한 체계적인 틀인 수와 첸의 연구를 이용해 국내 탐색엔진들을 체계적으로 평가함으로써, 탐색엔진을 평가하는 기준을 정립하는데 도움이 되고자 한다.

2) 실용적 기여

시스템 개발자가 이 연구의 결과로 자신의 인터넷 탐색엔진의 부족한 면을 발견하고 개선시킬 방안의 근거자료로 활용할 수 있을 것이다. 예를 들면, 이 연구에서 검색결과 상위 20건에 대해 간단한 설명으로, 또 그 설명으로 실제 해당 사이트를 보고 적합한지를 판정하게 하는 정확율의 결과로 간단한 설명이 적합성 판정에 도움이 되지 않을 경우, 새로운 방법을 모색해야 할 것이다.