

통합 검색 환경에서 이용자 적합성 판단 기준에 관한 탐색적 연구

Users' Relevance Criteria in Universal Search in Korea : An Exploratory Study

박정아(Jung Ah Park)*

초 록

본 연구는 한국 통합 검색 환경에서의 이용자 적합성 판단 기준에 관한 탐색적 연구이다. 이를 위해 10명의 참가자들을 대상으로 반구조화(semi-structured) 인터뷰를 수행하여 데이터를 수집하였다. 참가자들은 네이버, 다음 등과 같은 통합 검색 환경에서 본인들이 관심 있거나 필요로 하는 다양한 검색을 수행하고, 그 과정에서 문서가 적합한지와 그 판단 기준에 대해 기술하였다. 연구 결과 8개의 적합성 판단 기준과 비적합성 판단 기준, 그리고 검색 환경이 변화하여도 이용자가 적합성을 판단하는 기준들이 크게 변화하지는 않지만 데이터 증가와 이용자 요구의 고도화로 특수성과 구체성이 중요한 적합성 판단 기준으로 부각되는 점을 발견하였다.

ABSTRACT

This study is an exploratory research on the user relevance criteria in Korean search service environments that provide integrated search results. Data were collected from 10 participants using a semi-structured interview technique. The participants conducted a web search using integrated search services, such as Naver or Daum on a self-selected topic. They were asked to judge the relevance of retrieved documents and to report their relevance criteria. As a result, the research indicated 8 user-defined relevance and non-relevance criteria. The research shows that specificity and richness are the two most important criteria yet, the user's relevance criteria have not changed much despite the change in search environment.

키워드: 정보검색, 적합성, 이용자 중심 적합성, 적합성 판단 기준, 이용자 연구
information retrieval, relevance, user-oriented relevance, relevance criteria, user study, user relevance

* 다음커뮤니케이션 검색 팀장, Ph.D.(sunseed9@gmail.com)

■ 논문접수일자: 2012년 5월 17일 ■ 최초심사일자: 2012년 5월 27일 ■ 게재확정일자: 2012년 6월 14일
■ 정보관리학회지. 29(2). 113-133, 2012. [http://dx.doi.org/10.3743/KOSIM.2012.29.2.113]

1. 서론

인터넷의 발전으로 많은 정보가 급격하게 증가하고 있고 필요한 대부분의 정보들을 인터넷에서 찾게 되었다. 그러나 이렇게 넘쳐나는 정보들은 결과적으로 이용자들에게 중압감을 느끼게 하는 정보 과부하 현상을 발생시키는 원인이 되기도 한다. 수많은 데이터 속에서 적합한 정보를 찾는 것은 건조 더미에서 바늘을 찾는 것에 비유될 정도로 점점 더 어려운 일이 되어 가고 있다. 정보검색의 주요 목적은 적합한 문서를 찾아주는 것으로, 적합성은 정보 검색의 핵심 개념이다. 초기의 적합성은 정확률(precision)과 재현률(recall)과 같은 시스템 중심의 개념으로 논의되었다. 그러나 시스템 중심의 적합성만으로는 부족하다는 주장이 제기되면서 적합성 개념에 대한 논의가 재기되었다. 그리고 이용자 중심 관점의 연구가 활발해 지기 시작했다. 이용자 중심 관점에서 적합성은 다차원의 인지적 개념으로써 이용자의 정보 인식과 정보 이용자의 정보 요구 상황에 상당 부분 의존한다.

적합성에 대한 중요한 연구 주제 중 하나는 이용자의 적합성 판단 기준이다. 즉 이용자가 문서가 적합한지 아닌지를 판단하는 속성들이다. 기존 많은 연구들은 이용자가 어떤 기준으로 정보를 적합하다고 판단하는지, 즉 이용자 적합성 판단에 영향을 미치는 다양한 기준들을 연구해 왔다(Barry & Schamber, 1998; Barry, 1994; Bateman, 1998; Choi & Rasmussen, 2002; Cuadra & Katter, 1967a, 1967b; Park, 1993; Rees & Schultz, 1967; Savolainen & Kari, 2006; Schamber, 1991; Wang & Sorel, 1998).

적합성 판단 기준에 관해, 대부분의 기존 연

구들은 탐색적으로 접근하여 인터뷰와 같은 정성적 연구 방법을 사용하여 왔다(Barry, 1994; Hirsh, 1999; Park, 1994; Schamber et al., 1990; Wang & Soergel, 1998). 적합성 판단 기준에 관한 기존 연구들은 다양한 적합성 판단 기준들 - 예를 들어 최신성(recency), 신뢰성(reliability), 주제성(topicality) 등 - 을 제시하였다. 상당수의 이용자 적합성 판단 기준에 대한 기존 연구들이 전통적인 검색 환경 그리고 연구 또는 논문 검색 도메인에서 이루어졌다. 오늘날과 같은 웹 검색 환경에서의 연구는 Xu와 Chen(2005), 그리고 Savolainen과 Kari(2006) 연구 정도를 들 수 있다. 더욱이 한국과 같은 통합 검색 환경에서의 적합성 판단 기준에 관한 연구는 찾기 어려운 실정이다.

통합 검색은 웹 문서, 지식(Q&A), 블로그, 카페, 뉴스, 디렉토리, 쇼핑, 전문 자료 등 다양한 형태의 자료들을 동시에 검색해 주고 형태별로 구분하여 검색 결과를 제공해 주는 검색 서비스를 말한다. 이는 주로 동일한 형태의 자료들 - 예를 들어, 논문, 이미지, 웹 문서 등 - 만을 검색 대상으로 하는 전통 검색 환경이나 웹 검색 환경과 차이가 있다. 통합 검색 서비스는 2000년 8월 네이버가 인터넷 검색 업계에서 세계 최초로 개발한 이후 국내 타 검색 포털들도 구현, 제공하고 있는 서비스이다. 통합 검색 서비스는 국내에서 가장 대중적이고 인기가 있는 서비스라고 할 수 있다(박소연, 이준호, 2008). 해외 검색 포털들의 경우, 2007년 6월 구글이 통합 검색과 기본 개념이 유사한 “Universal Search”를 도입, 공개하여 점차 통합 검색을 벤치마킹 하고 있는 상황이다. 비즈니스 위크지는 구글이 NHN에게서 배워야 할 점으로 통

합 검색 서비스를 주목하기도 하였다(Moon & Woyke, 2006). 그러나 현재까지 한국 통합 검색 환경에서의 이용자 적합성 판단 기준에 관한 연구는 찾기 힘들다.

본 연구는 이러한 한국에서 가장 대중적으로 사용되는 통합 검색 환경에서의 이용자 적합성 판단 기준을 알아보고자 하였다. 한국 통합 검색 환경에서 사용자가 적합성 여부를 판단하는 기준에는 무엇이 있는지 조사해 보고, Barry와 Schamber(1998)의 주장처럼 검색 환경의 변화에도 크게 변하지 않는 핵심 적합성 판단 기준들이 존재하는지 그리고 통합 검색 환경에서 특징적인 적합성 판단 기준은 무엇인지를 알아보려고 하였다.

2. 이론적 배경

2.1 적합성 개념

적합성은 정보 검색의 주요 연구 분야이다(Borlund, 2003; Mizzaro, 1997; Schamber, 1994). 적합성은 정보 검색 시스템의 기능과 평가에 있어 기본적인면서도 중요한 것으로 알려져 있다(Borlund, 2003). 적합성에 관한 많은 연구에도 불구하고, 적합성 개념은 아직까지 명확하게 정의하기 어렵다(Toms et al., 2005). 기존 연구들에 의하면 적합성은 복합적이고 다차원적이면서 또한 상황에 따라 달라지는 동적인 개념으로, 인지적, 정서적, 사회 문화적 요인들에 영향을 받는다고 알려져 왔다(Schamber, 1994).

초기의 적합성은 정확률(precision)과 재현율

(recall) 같은 시스템 중심의 개념으로 논의되었다. 그러나 시스템 중심의 적합성만으로는 적합성 개념을 충분히 설명하기 부족하다는 주장이 제기되면서 적합성 개념에 대한 논의가 제기되었다(Saracevic, 1975; Schamber, Eisenberg, & Nilan, 1990). 적합성에 관한 시스템 중심 접근은 실질적인 이용자 요구를 고려하지 못하는 한계를 가진다(Fruend, 2008). 그리고 적합성 판단에 영향을 주는 상당수의 주관적이고 상황적인 변수들을 고려하지 못한다(Barry, 1994).

적합성에 관한 연구는 점차 시스템 중심에서 점차 이용자 중심으로 넘어왔다(Borlund, 2003; Cosijn & Ingwersen, 2000; Mizzaro, 1997; Saracevic, 1975; Schamber, 1994). 이용자 중심 적합성 개념에 관한 연구는 1960년대 후반에 시작되어 활발해 지기 시작했다(Borlund, 2003; Cosijn & Ingwersen, 2000; Saracevic, 1970; Schamber et al., 1990). 이용자 관점에서 적합성은 “다차원의 인지적 개념으로써 이용자의 정보 인식과 정보 요구 상황에 상당 부분 의존한다”고 정의될 수 있다(Borlund, 2003, p. 913).

1970년대와 1980년대에 행해진 많은 연구들은 적합성 판단이 이용자의 지식과 인식에 의존적이고, 적합성 평가가 이용자의 목적, 지식 수준과 믿음, 정보의 속성, 시간 제한, 정보 획득에 소요되는 노력과 비용 등과 같은 다양한 요인들에 영향을 받는 역동적 과정이라는 것을 보여주었다. 이용자 중심 관점에서 적합성은 정보 요구와 문서와의 관계에 대한 이용자의 주관적 인식이다(Borlund, 2003; Cosijn & Ingwersen, 2000; Mizzaro, 1997; Saracevic, 1975; Schamber, 1994). Schamber, Eisenberg

와 Nilan(1990, p. 773)은 “이용자가 정보를 평가하는데 사용하는 기준 등을 연구함으로써 적합성에 대한 더 구체적인 이해를 할 수 있게 될 과 동시에 시스템 디자인에도 도움을 줄 수 있을 것이다”라고 하여 이용자 중심 적합성 연구의 중요성을 강조하기도 하였다.

2.2 적합성 판단 기준에 관한 기존 연구

적합성에 대한 중요한 연구 주제 중 하나는 이용자 적합성 판단 기준이다. 즉 문서가 적합한지 아닌지를 판단하는 속성들이다. 이용자는 무엇으로 정보를 적합하다고 판단하는가? 적합성 판단에 영향을 미치는 기준들은 오랫동안 주요 연구 주제였다. 예를 들어 Cuadra와 Katter(1967a, 1967b) 그리고 Rees와 Schultz(1967)의 연구를 들 수 있다. 이 연구들은 적합성 판단에 영향을 줄 수 있는 40개의 변수들을 정의하였다. 특히 1990년대 후반 이후 웹 검색을 중심으로 정보이용자들을 대상으로 적합성 판단 기준에 대한 경험적 연구들이 집중되면서, 주제성 외에도 수십 가지의 적합성 판단 기준들이 정의되어 왔다(Mizzaro, 1997; Saracevic, 2007; Schamber, 1994).

Schamber(1991)는 30명의 직장인들을 대상으로 실제 날씨 정보 검색을 사용하는 상황에서 인터뷰를 하였다. 참가자들은 그들이 어떻게 다양한 형태의 날씨 정보를 평가하는지에 답하였다. 총 811개의 다양한 기준이 정보를 평가하는데 사용된 것으로 나타났다. 이 기준들은 총 10개의 범주로 분류되었다. 정보의 정확성, 현재성, 지리적 근접성, 신뢰성, 접근용이성, 검증가능성, 정보제시의 명확성, 역동성 또는 이용자와

시스템의 상호 작용성, 오락적(entertainment) 가치와 정보에 대한 정서적 반응, 화면 품질 등이다. 이러한 범주 중에서 화면 품질, 현재성, 신뢰성이 가장 많이 언급된 반면 접근성, 정확성, 명확성이 가장 적게 언급되었다.

Barry(1994)의 연구에서는 적합성 판단 기준으로 사용되는 핵심 기준 집합이 존재한다고 주장되었다. 이 연구에서는 주립 대학의 18명의 학생들과 교수들을 대상으로 23개의 적합성 판단 기준을 정의하였다. 연구 결과 적합성 판단 기준으로 수준/범위, 최신성, 비용 등이 있었고, 이 중 수준/범위가 가장 자주 언급되었다(14%). 새로운 출처가 두 번째(12%) 그리고 주관적 정확성/타당성 순이었다. 비용과 같은 판단 기준은 매우 드물게 언급되었다. Barry는 이 연구를 통해 예상대로 주제성 외에도 상황적 요소가 영향을 많이 미치지만, 상황에 따라 변하지 않고 적합성 판단 기준으로 사용되는 적합성 판단 기준 집합 자체는 크게 변하지 않는다고 주장하였다. 이밖에도 응답자가 10명을 넘어가면 언급되는 적합성 판단 기준이 중복되면서 더 이상 새로운 기준을 찾기 힘들다고 하였다(Barry, 1994). Barry는 다른 검색 목적, 다른 종류의 정보 등 자신의 연구와 다른 환경에서의 적합성 판단 기준에 관한 연구, 그리고 적합성 판단 기준간의 중요도, 적합성 판단 기준과 문서의 특징과의 연관성을 찾는 것과 같은 향후 연구의 필요성을 제안하였다.

상황이나 도메인의 변화가 있더라도 적합성 판단 기준 집합 자체는 크게 변하지 않는다는 Barry의 주장처럼 기존 연구들에서 정의된 적합성 판단 기준들이 상당 부분 겹치는 현상에 주목한 연구들이 있다. Schamber와 Bateman

(1996)는 기존 연구에서 밝힌 적합성 판단 기준을 비교 및 분류하여 5개의 주요 적합성 기준, 즉 명확성(clarity), 현재성(currency), 신뢰성(credibility), 접근가능성(availability), 주제성(aboutness)으로 압축한 바 있다. Barry와 Schamber(1998) 또한 정보 이용자나 상황에 상관없이 이용자 정의 적합성 판단 기준의 유한한 집합이 존재한다고 하였다. 이들 또한 각자의 연구에서 밝혀진 적합성 판단 기준을 비교 및 분류하여 10개의 주요 적합성 판단 기준을 정리하였다. 10개의 적합성 판단기준으로는 수준/범위/특수성, 정확성/타당성, 명확성, 현재성, 현실성, 출처/품질, 접근용이성, 접근 가능성, 검증 가능성, 정서, 주제 적절성 등이 있었다.

2000년대에도 다양한 도메인을 대상으로 이용자 적합성 판단 기준을 정의한 연구가 계속되었다. 다양한 적합성 판단 기준들이 제시되었는데, 특히 주제성은 대부분의 연구에서 공통적으로 언급된 적합성 판단 기준이었다(Bateman, 1998; Hirsh, 1999; Schamber & Bateman, 1996; Wang & Soergel, 1998). 주제성은 적합성 판단의 기본 조건으로 정의되기도 하였다(Greisdorf, 2003).

미국 역사 이미지 검색 도메인에서의 이용자 적합성 판단 기준을 연구한 Choi와 Rasmussen(2002)의 연구에서는 시간대(time frame), 정확성(accuracy), 주제성(topicality)이 이용자가 적합성을 판단할 때 가장 중요한 기준으로 나타났다. 이들은 비적합성 판단 기준도 함께 연구하였는데 판단 기준은 순서만 다르고 동일하게 나타났다. 즉 주제성, 시간대, 정확성의 순이었다. 동영상 검색에서의 적합성 판단 기준

을 연구한 Yang과 Marchionini(2004)의 연구는 동영상 정보 검색에 특화된 다양한 적합성 판단 기준을 정의하였다. 예를 들어 장르, 기간, 영화 촬영법, 움직임 등이었다. 이 연구에서도 이전 연구들에서 주요 기준으로 자주 언급된 주제성이 가장 중요한 기준으로 드러났다. 이 연구의 결과는 정보 검색의 도메인과 상관없이 주제성이 매우 중요한 적합성 기준이라는 것을 말해주는 것이다. 또한 주제성이 전통 검색 환경과 웹 검색 환경에서 모두 중요한 기준이라는 것을 의미한다(Savolainen & Kari, 2006).

아이들을 대상으로 한 적합성 판단 기준에서도 주제성은 주요 적합성 판단 기준으로 나타났다. Hirsh(1999)는 10명의 5학년 아이들을 대상으로 적합성 판단 기준을 연구하였다. 이 연구에서 적합성 판단 기준은 주제성(49%), 신선성(15%), 흥미성(10%), 품질(8%), 관심사(7%), 편리성/접근용이성(5%), 최신성 등의 시간 관련 이슈(3%), 권위(2%), 언어(1%) 순으로 나타났다. 아이들 대상으로 웹 검색 엔진 Yahoo!를 연구한 Bilal(2000)의 연구에서도 주제성이 주요 적합성 기준으로 나타났다.

웹 검색 환경, 그리고 특히 주로 연구나 학업과 관련된 검색이 대상이었던 기존 연구들과 달리 일상생활 검색을 대상으로 한 적합성 판단 기준에 관한 Savolainen와 Kari(2006)의 연구에서는 특수성과 주제성이 적합성 판단과 비적합성 판단 모두에서 가장 높은 비중을 차지한 기준으로 나타났다. Savolainen와 Kari(2006)의 연구에서는 특수성을 정보가 이용자의 요구에 얼마나 초점을 잘 맞추었는지로 정의하였고, Taylor 등(2007)에서는 특수성과 수준/범위를 구분하여 특수성을 “일반적이거나 전문적인 정

도”와 “기술적인 정도”로, 수준/범위를 “문서의 특정한 범위 정도”와 “문서 내 다양한 주제에 대한 정의”로 구분한 바 있다. Taylor 등의 연구에서는 모든 검색 단계에서 특수성이 가장 많은 비중을 차지하였고, 그 다음으로 수준/범위였다.

이러한 적합성 판단 기준에 관한 대부분의 연구들이 정성적 연구 방법에 기반한 탐색적 연구인데 반해, Xu와 Chen(2006)은 적합성 판단 기준에 관한 정량적 연구를 시도하였다. 이들은 Grice의 커뮤니케이션 이론에 기반하여 주제성(topicality), 신선성(novelty), 신뢰성(reliability), 이해가능성(understandability), 범위(scope)의 5가지의 주요 적합성 판단 기준을 도출하였다. 이들 연구에서 5가지 적합성 판단 기준 중 범위를 제외한 나머지 4가지가 통계

적으로 유의미하게 나타났고, 주제성과 신선성이 주요 기준으로 밝혀졌다. Xu와 Chen(2006)의 연구는 통계적으로 적합성의 핵심 기준들을 정의하였다는데 의의가 있고, 알려져 온 주제성 외에 신선성의 중요성을 밝히기도 하였다.

이상에서 살펴본 선행연구들을 고찰해 보면, 적합성 판단 기준에 관한 연구는 다양한 도메인과 상황에서 꾸준히 이루어져 왔으며 다양한 적합성 판단 기준들이 제시되어 왔음을 알 수 있다(〈표 1〉 참조). 그리고 상당수의 적합성 판단 기준은 명칭은 다를지라도 많은 부분이 겹친다는 것을 알 수 있다. 또한 적합성 판단 기준 중에서 특히 주제성은 정보 검색의 도메인과 상관없이 전통 검색 환경과 웹 검색 환경에서 모두 중요한 기준으로 제시되어 왔다는 것을 알 수 있다.

〈표 1〉 기존 연구들의 적합성 판단 기준

적합성 판단 기준	출처
topicality	Wang & Soregel(1998), Wang & White(1999), Hirsh(1999), Choi & Rasmussen(2002)
related	Park(1997)
about my topic	Bateman(1998)
topic relatedness	Tang & Solomon(1998)
specificity	Schamber(1991), Barry & Chamber(1998), Fitzgerald & Gallway(2001)
depth/scope	Barry(1994), Maglaughlin & Sonnewald(2002), Xu & Chen(2006)
specific, suitably general or specific	Park(1997), Fitzgerald & Gallway(2001)
focused	Bateman(1998)
insufficient information	Park(1993), Fitzgerald & Gallway(2001)
variety/volume	Schamber(1991), Savolainen & Kari(2006)
suitable length	Bateman(1998)
detailed, overview	Bateman(1998)
not enough information	Spink et al.(1998)
novelty	Barry(1994), Wang & Soregel(1998), Wang & White(1999), Hirsh(1999), Maglaughlin & Sonnewald(2002)
new information, repetitive information	Park(1993)
duplicated information	Spink et al.(1998), Tang & Solomon(1998)

적합성 판단 기준	출처
unique/only, original, new	Bateman(1998)
recency	Barry(1994), Park(1993), Wang & Soregel(1998), Wang & White(1999), Tang & Solomon(1998), Hirsh(1999), Maglaughlin & Sonnewald(2002)
currency, dynamism	Schamber(1991)
timeliness	Spink et al.(1998)
understandability	Barry(1994), Bateman (1998), Spink et al.(1998)
clarity	Schamber(1991), Barry(1994), Hirsh(1999)
too technical, readability	Park(1993), Spink et al.(1998)
well-written, provides background/graphics	Bateman(1998), Maglaughlin & Sonnewald(2002)
(wrong) language	Spink et al.(1998), Tang & Solomon(1998), Hirsh(1999)
reliability	Schamber(1991)
accuracy	Schamber(1991), Barry(1994), Bateman(1998), Maglaughlin & Sonnewald(2002), Choi & Rasmussen(2002)
verifiability	Schamber(1991), Barry(1994)
authoritative source, author's/journal's status	Barry(1994), Spink et al.(1998), Maglaughlin & Sonnewald(2002)
credible, reputation	Bateman(1998)
authority, quality	Hirsh(1999), Maglaughlin & Sonnewald(2002)
entertainment value	Schamber(1991)
interesting, like, enjoyable	Bateman(1998)
excites me	Spink et al.(1998)
appeal of information	Choi & Rasmussen(2002)
affectiveness	Barry(1994)
accessibility	Schamber(1991), Hirsh(1999), Choi & Rasmussen(2002)
availability	Schamber(1991), Barry(1994), Park(1993), Park(1997), Wang & Soregel(1998), Wang & White(1999)

3. 통합 검색 환경에서의 이용자 적합성 판단 기준 조사

3.1 데이터 수집 및 분석 방법

본 연구는 10명의 한국 인터넷 기업 종사자들을 대상으로 데이터를 수집하였다. 참가자들은 본인들이 관심 있거나 필요한 내용을 자유롭게 검색하였다. 인터넷 검색은 주로 네이버, 다음과 같은 통합 검색을 사용하도록 요청되었고,

검색하는 과정 동안 반구조화(semi-structured) 인터뷰가 수행되었다. 인터뷰는 문서를 클릭한 이유와 클릭하지 않은 이유, 문서를 보고 나서 문서가 적합하다고 생각하는지, 만족스러운지, 적합하거나 만족스러운 이유가 무엇인지, 반대로 적합하지 않거나 만족스럽지 않은 이유가 무엇인지 등과 같은 질문으로 이루어졌다. 응답자가 구체적으로 답변하기 어려워하는 경우 기존의 연구에서 밝혀진 기준들 즉 이해하기 쉬운지, 새로운 내용인지, 믿을만한지, 원하는 내용에 초

점이 맞춰져 있는지 등의 질문에 답하도록 하였다. 참가자들은 한 문서에 대해 복수개의 적합성 판단 기준을 답할 수 있었다. 사전 동의하에 검색 과정의 모든 화면이 Camstudio 프로그램을 이용하여 녹화되었고 인터뷰 또한 모두 음성 녹음되었다.

본 연구는 10명의 참가자를 대상으로 하였고, 범주화 및 코딩 과정에서 8명을 넘어가면서 대부분 이미 도출된 적합성 판단 기준과 중복되는 것을 볼 수 있었다. 총 10명의 참가자 중 6명이 남자, 4명이 여자였으며 평균 32.4세였다. 10명 모두 인터넷 기업 종사자들이었다. 이들은 총 55개의 주제를 검색하였고, 178번의 검색을 수행하였으며, 총 450개의 문서를 평가하였다. 검색한 주제는 뉴스(예: 신종플루), 쇼핑(예: 블랙베리, 넷북, 보이스 레코더 추천 등), 영화/책/여행에 대한 리뷰(예: 애자, 천년의 금서, 몬베츠), 업무 관련(예: mysql_pconnect, maximum entropy model, 구글 검색 지표), 생활 지식(예: 청약예금, 청약부금) 등으로 다양했다. 검색 과정에서 브라우저 등은 제한하

지 않았으며 참가자들이 평소에 사용하는 환경과 동일한 것을 사용하도록 권장되었다. 한 사람당 60분에서 100분 동안 자유롭게 검색을 하였다.

음성 녹음 및 녹화된 인터뷰 자료는 녹취 과정을 통해 응답자들의 분석 자료로 사용되었다. 이 자료를 기반으로 코딩 및 범주화 과정을 거쳐 참가자들의 적합성 판단 기준과 비적합성 판단 기준을 도출하였다. 도출된 적합성 판단 기준은 <표 3>에 정의되어 있다. 코딩 및 범주화 과정에서 기존 연구들(Barry, 1994; Barry & Schamber, 1998; Bateman, 1998; Choi & Rasmussen, 2002; Hirsh, 1999; Maglaughlin & Sonnewald, 2002; Park, 1993; Park, 1997; Savolainen & Kari, 2006; Schamber, 1991; Spink et al., 1998; Tang & Solomon, 1998; Wang & Soregel, 1998; Wang & White, 1999; Xu & Chen, 2006)에서 정의한 적합성 판단 기준들을 참고하였다. 연구자 본인과 인터넷 기업에 종사하는 정보 검색 전문가 총 2명이 코딩 작업을 수행하였다. 코더 간 신뢰도 계수는

<표 2> 본 연구 참가자

참가자	성별	나이	검색한 주제수	검색 회수	평가한 문서수
1	남	35	1	10	28
2	여	26	6	24	49
3	남	31	10	10	36
4	남	34	12	23	59
5	여	38	5	11	44
6	여	31	4	10	45
7	남	34	3	13	22
8	남	32	5	23	83
9	남	33	3	18	33
10	여	30	6	36	51
합계			55	178	450

0.82(82%)이었다. 코더간 신뢰도(inter-coder reliability)는 두 코더간 일치도를 백분율로 계산하였다(Stempel, 1981). 일반적으로 탐색적 내용 분석 연구의 경우, 최소한 0.80(80%)를 넘는 것이 권장된다(Krippendorff, 1980).

3.2 분석 결과

3.2.1 적합성 판단 기준 집합

한국 통합 검색 환경에서 이용자는 적합하다고 판단하는데 있어, 주제성, 특수성, 구체성, 신선성, 이해가능성, 신뢰성, 흥미성의 7개 적합성 판단 기준을, 적합하지 않다고 판단하는 데 있어서는 접근용이성을 포함하여 8개의 적합성 판단 기준을 사용하는 것으로 나타났다. 통합 검색 환경에서 적합성 여부를 판단하는 기준으로 나타난 주제성, 특수성, 구체성, 신선성, 이해가능성, 신뢰성, 흥미성, 접근 용이성의 총 8개 기준은 기존 연구에서 정의되어 온 적합성 판단 기준에 모두 부합되는 것으로써, 한국 통합검색 환경에서도 상황이나 환경에 관계없이 적합성 판단 기준 집합이 크게 변화되지 않음을 확인할 수 있었다. 적합성 판단 기준 각각의 설명과 예시는 <표 3>에 정리되어 있고, 각 적합성 판단 기준을 살펴보면 다음과 같다.

1) 주제성

검색 이용자들은 현재 검색하고 있는 주제와의 관련성이나 문서의 내용이 현재 검색 주제에 관한 것인지를 보고 적합성 여부를 판단하는 것으로 나타났다. 예를 들어 '소니 ux81'에 관한 정보를 찾고 있던 한 검색이용자는 검색 결과 중 한 문서를 살펴 보고 나서 해당 문서가

소니 ux81에 관한 내용이 아니라는 이유로 해당 문서가 검색결과로 적합하지 않다고 판단하였고, '신종플루'에 관한 정보를 찾고 있는 검색 이용자 또한 검색 결과문서에 신종플루라는 단어가 들어 있기는 하지만, 실제로는 신종 플루를 말하고 있는게 아니기 때문에 적합하지 않다고 판단하였다. 이외에도 '블랙베리 트위터 클라이언트' 프로그램에 대한 정보를 찾고 있던 검색 이용자도 찾은 문서가 트위터에 대한 이야기일 뿐, 해당 프로그램에 대한 설명이 아니어서 적합하지 않다고 하였고, '라이카 m8'을 찾고 있던 검색 이용자도 라이카 m8과 관련된 내용이 아니기 때문에 적합하지 않다고 판단하였다. 검색 이용자들은 주로 제목이나 내용 요약 부분에 나타난 단어로 주제성 여부를 판단하였다. 예를 들어 검색어가 문서에 나왔지만 서로 멀리 떨어져 있을 경우 관련 내용이 아니라고 보았다.

다양한 정보 검색 도메인에서 공통적으로 언급되어 온 적합성 판단 기준이기도 한 주제성은, 한국 통합 검색 환경에서 적합성 판단에서는 많이 언급되지 않은 반면(6.96%), 부적합성 판단에서는 특수성(37.54%) 다음으로 많은 빈도(22.53%)로 언급되었다. 이는 주제성이 적합성의 기본 기준, 즉 필요조건 이지만 충분조건은 아니다(Burgin, 1992; Cooper, 1971; Froehlich, 1994)라는 주장에 부합된다고 볼 수 있다. 즉 적합성 판단일 경우, 주제성은 기본적으로 만족하므로 별도로 언급되지 않고 다른 기준들이 결정적으로 적합성 판단에 영향을 미치는 반면, 기본적으로 주제성이 만족되지 않을 때 부적합성 판단을 내리게 된다고 유추해 볼 수 있겠다.

〈표 3〉 이용자 적합성 판단 기준 정의와 응답 예

기준	정의	예시
주제성	검색하고 있는 주제와의 관련성으로 문서의 내용이 현재 검색 주제에 관한 것인지 여부	“이건 소니 ux81에 관한 내용이 아니예요” [p5, 소니 ux81] “이건 트위터에 대한 설명이군요.” [p1, 블랙베리 트위터 클라이언트] “관련 없는 내용이에요” [p6, 직장인 공인 중개사 준비] “신종플루를 말하는 게 아닌데요” [p4, 신종플루]
특수성	다루는 범위가 적절하거나 원하는 특정 내용을 가리키고 있는지 또는 원하는 범위에 얼마나 초점을 잘 맞추고 있는지 여부	“블랙베리가 맞긴 한데 OS 버전이 달라요” [p1, 블랙베리 OS] “영화나 감독 얘기보다 배우 얘기인 거 같아요” [p2, 애자] “원하는 건 삼성인데 아수스 넷북에 관한 거네요” [p4, 넷북] “코드예제를 보고 싶었는데 바로 그거네요” [p8, mysql_pconnect]
구체성	내용이 다양하거나 상세하고, 충분한 정보를 다루었는지 여부	“내용이 다양하고 상세해서 만족스럽다” [p1, 블랙베리 리뷰] “자세하게 잘 적어놓아서 여행 갈 때 도움이 많이 될 거 같아요” [p2, 일본 도깨비 여행] “사이트 링크도 되어있고, 심지어 동영상도 있고, 사진 멋지고, 영화 정보까지 풍부해서 완전 좋아요” [p8, 엘 시스템아] “제작발표회 현장, 사진, 캐릭터 소개 등이 자세하 있어 좋다” [p5, 뮤지컬 영웅]
신선성	내용이 새롭거나, 독창적이거나 고유한 정보를 다루고 있는지 또는 최신의 정보이거나 적절한 시간대의 문서인지 여부	“넷북 사려고 했는데 몰랐던 새로운 정보, 새로운 후보군을 줘서 매우 만족스럽다” [p4, 넷북 추천] “몰랐던 걸 알게 되서 좋았다” [p8, mysql_connect] “날짜가 4월, 5월이라서 안본다. 컴퓨터는 주기가 짧아서” [p4, 넷북 추천] “다 비슷한데 최신이라서 봤어요” [p6, 피넛걸] “이미 다 알고 있는 내용이에요” [p10, 신종플루] “제목만 다를 뿐, 아까 본거랑 똑같은 내용이에요” [p8, 몬베츠]
이해가능성	내용이 명확하거나, 가독성이 좋은지, 잘 정리되어 있거나 또는 그림 등이 제공되는지 여부	“한눈에 훑어볼 수 있어 좋다” [p4, kindle case] “Q&A 형식으로 보기 좋게 정리를 잘 해놓았다” [p9, 신종플루 증상] “보기 편하게 정리 잘 되어 있네요” [p10, 여성용 통기타] “정보가 매우 깔끔하네요” [p5, 소니 보이스레코더 ux-81] “이해하기 쉬워서 1/5.. 이는 내용 중심으로 지식 습득 가능하니가 어느 정도는 이해할 수 있어서 좋았어요” [p7, maximum entropy model]
신뢰성	내용이 정확하거나 믿을만한 정보인지, 권위 있는 출처 또는 저자인지 여부	“답변수가 많아서 괜찮은 정보일 것 같네요” [p6, 피넛걸] “이거는 다시 한번 재확인 하는 것도 맘에 듬. 신뢰하거나 정리할 수 있어서” [p6, 직장인 공인 중개사 준비] “알고 있는 내용이지만 확신시켜줘서 만족해요” [p1, 블랙베리 리뷰]
흥미성	문서가 흥미롭거나 흡입력 있는지 여부	“별다른 정보는 없지만 웃음 줘서 어느정도 만족해요” [p2, 찹깨 다이어 몬드] “몰랐던 영화 에피소드에 대한 내용이라 흥미롭네요” [p3, 정기훈] “딱 원했던 건 아니지만 재밌긴 해요” [p8, powerpoint template]
접근용이성	정보를 접근하는데 있어 편의가 제공되는지 여부	“검색결과가 크롤링 당시 내용이라 달라서 찾을 수가 없어요” [p1, 블랙베리 트위터] “이미지가 깨져있네요” [p2, 이블팩] “로그인 해야 볼 수 있네요” [p7, general iterative scaling IIS] “가입해야 볼 수 있군요” [p8, 템플릿]

2) 특수성

검색 이용자들은 또한 특수성 즉, 문서에서 다루는 범위가 적절한지, 원하는 특정 내용을 가리키고 있는지, 원하는 범위에 얼마나 초점을 잘 맞추고 있는지를 적합성 판단의 기준으로 사용하는 것으로 나타났다. 특수성에 관한 응답 예를 보면, '여행박사'에 관한 정보를 찾고 있던 한 검색 이용자는 해당 문서가 '여행박사'에 대한 내용이라는 하지만, 본인이 원하는 여행박사에 대한 사이트 정보가 나와 있지 않았기 때문에 해당문서를 적합하지 않다고 판단하였다. 또 '애자'라는 영화 정보를 찾았던 검색 이용자는 해당 영화에 대한 이야기이긴 하지만 영화나 감독에 대한 정보보다 영화에 출연한 배우에 관한 이야기가 대부분이라 적합하지 않다고 보았다. 'mysql_pconnect'로 검색하던 한 이용자는 코드 예제를 보고 싶었는데 바로 그 코드 예제가 나와 있는 문서를 찾아 적합하다고 판단하였다. 이외에도 검색 이용자들은 뮤지컬 배우에 관한 정보를 찾으며 실제로 그 배우가 공연하는 정보를 원하거나(p5, '홍광호') 넷북에 관한 정보를 찾으면서도 특정 브랜드에 관한 정보를 원하는 등(p4, '넷북') 검색어 어느 정도 구체적인 검색 니즈를 가지고 이에 부합하는 문서를 발견하였을 때에 적합하다고 판단하였고, 그렇지 않은 경우 해당 문서가 찾고 있는 정보에 대한 내용을 다루고 있다 하더라도 적합하지 않다고 판단하였다.

특수성은 적합성 판단에서는 28.26%의 빈도로, 비적합성 판단에서는 37.54%의 빈도로 언급됨으로써, 양쪽에서 가장 높은 비중을 차지하였다. 비적합성 판단의 경우, 관련된 내용이긴 하지만(주제성), 원하는 형태나 브랜드가 아

니어서(비특수성) 적합하지 않다고 판단하는 경우가 많은 것이 특징적이었다.

3) 구체성

검색 이용자들이 적합성을 판단하는 기준의 하나는 내용의 다양성, 상세함, 충분한 양과 같은 구체성에 관한 것이었다. 구체성과 관련해 적합하다고 판단한 검색 이용자들의 응답들은, 내용이 다양하고 상세하다, 자세하다, 다양한 정보들이 나와 있고, 풍부하다 등이 있었다. 이외는 반대로, 내용이 너무 짧고 불만한 내용이 없다는 이유로 적합하지 않다고 판단하는 것으로 나타났다. 구체성은 적합성 판단에서는 특수성에 이어 두 번째로 많이 언급된 기준인 반면(21.74%), 비적합성 판단에서는 주제성, 특수성, 신선성에 이어 빈도순으로 4번째로 언급된 기준이었다(10.58%). 이는 적합성 판단에 있어서 상세하고 다양하며 충분한 내용이 중요하지만, 비적합성 판단의 경우 주로 관련이 없는 내용이거나, 특정한 형태가 맞지 않거나, 이미 보거나 알고 있는 기준들에서 걸러지는 것으로 유추해 볼 수 있겠다.

4) 신선성

적합성 판단 기준으로는 또한 새롭거나 독창적인 정보, 최신 정보, 적절한 시간대의 문서 등에 관한 신선성이 사용되는 것으로 나타났다. 신선성과 관련하여 검색 이용자들은 몰랐던 내용을 알려주었기 때문에 만족스러웠고, 적합하다고 판단하였다. 예를 들어 넷북을 사기 위해 필요한 정보를 찾던 검색 이용자는 '넷북을 사려고 했는데 몰랐던 새로운 정보, 새로운 후보군을 줘서 매우 만족스럽다'고 응답하였다. 이

밖에도 관심 영화에 대한 정보를 찾는 검색 이용자도 '몰랐던 영화 찍는 과정 중 에피소드 내용을 알게 되서 좋았다'고 응답하였다. 신선성과 관련하여 검색이용자들은 또한 최신의 정보들을 적합하다고 판단하였다. 응답의 예를 보면, 컴퓨터는 주기가 짧기 때문에 서너달 전의 정보들은 보지 않는다(p4, '넷북 추천'), 대체로 최신의 리뷰를 선호한다(p5, '소니 보이스 레코더'), 대부분 비슷하지만 최신이기 때문에 보았다는 답변 등이 있었다(p6, '피넛겔'). 반대로 적합하지 않다고 판단한 경우로는, 이미 다 알고 있는 내용이다(p10, '신중플루'), 이전에 본 내용과 동일하다 등과 같은 응답들이 있었다.

신선성은 적합성 판단에서는 15.65%의 빈도로, 비적합성 판단에서는 17.06%의 빈도로 언급됨으로써, 양쪽에서 모두 3번째로 많이 언급된 기준이었다. 이는 신선성이 적합성 판단과 비적합성 판단 모두에서 모두 중요한 기준임을 암시한다고 볼 수 있다. 신선성은 Xu와 Chen(2006)의 연구에서 기준에 많이 알려진 주제성과 함께 중요하게 부각된 적합성 판단 기준이다.

5) 이해가능성

검색 이용자들은 보기 편하게 되어 있는지, 정리가 잘 되어 있는지 등과 같은 이해할 수 있는 정도를 또한 적합성 여부를 판단하는 기준으로 언급하였다. 이해가능성과 관련하여 검색 이용자들은 이는 내용이긴 하지만 보기 편하게 잘 정리되어 있고(p3, 'google wave'), 정보가 깔끔하다(p5, '소니 보이스레코더 ux-81'), 한눈에 훑어볼 수가 있다(p4, 'kindle case'), Q&A 형식으로 보기 좋게 정리를 잘 해놓았다(p9, '신중플루 증상'), 그리고 이해하기 쉬워

서 이는 내용 중심으로 지식 습득 가능하다(p7, maximum entropy model)는 이유로 적합하다고 판단하였다. 반대로 무슨 이야기인지 잘 모르거나 이해되지 않기 때문에 해당 문서를 적합하지 않다고 판단하였다.

이처럼 사람들은 보기 편하게 정리가 잘 되어 있는 문서 또는 한눈에 훑어보기 편한 문서를 선호하였다. 때로는 이미 알고 있는 내용이더라도 정리가 잘 되어 있으면 검증하는데 도움이 된다고 긍정적으로 받아들이기도 하였다. 또한 사람들은 전문 지식 수준보다, 비슷한 지식 수준 가진 사람들이 정리하거나 풀이해 준 것을 선호하였다. 이 또한 이해가능성의 차원으로 해석될 수 있다. 예를 들어 신중 플루에 대해, 질병관리 본부나 보건복지 가족부에서 전문 용어를 이용하여 정의한 내용보다는(예: 발열(37.8℃), 콧물, 인후통, 기침) 포털 지식 검색에 일반 사람들이 이해할 수 있는 용어로 설명된 내용을 선호하였다. 이해가능성은 적합성 판단에서는 11.74%, 비적합성 판단에 있어서는 6.14%의 빈도로 언급되었다.

6) 신뢰성

검색 이용자들은 정보의 정확성, 믿을만한 정보, 확신, 검증가능성 권위 있는 출처 또는 저자, 신용/평판, 명성 또는 품질 등과 같은 신뢰성을 적합성 판단의 기준으로 사용하는 것으로 나타났다. 신뢰성과 관련된 응답으로는 답변수가 많아서 괜찮은 정보일 것 같다(p6, '피넛겔'), 해당 정보가 다시 한번 재확인할 수 있게 해 줌으로써 신뢰하거나 정리할 수 있다(p6, '직장인 공인 증개사 준비'), 이미 알고 있는 내용이지만 확신시켜준다(p1, '블랙베리 리뷰')

등과 같은 이유를 들어 적합하다고 판단하였다. 신뢰성은 적합성 판단에서는 11.74%, 비적합성 판단에 있어서는 3.07%의 빈도로 언급되었다. 특이한 점은 신선성과 반대로 알고 있는 내용이더라도 구체적으로 정리가 잘 되어 있는 경우, 이는 내용을 검증할 수 있다는 점에서 만족하고 해당 검색결과를 적합하다고 판단하는 것으로 나타났다.

7) 흥미성

적합성 판단의 기준의 또 하나는 흥미성으로, 검색 이용자들은 내용이 흥미롭고, 재미있거나 즐겁다는 이유로 적합하다고 판단하는 것으로 나타났다. 응답의 예를 보면, 참깨 다이어몬드에 관한 정보를 찾던 검색 이용자는 검색결과 문서를 보고 별다른 정보가 없지만 웃음을 주기 때문에 만족한다(p2, '참깨 다이어몬드'), powerpoint template 정보를 찾던 검색 이용자 또한 딱 원했던 것은 아니지만 재밌긴하다는 이유로 해당 문서를 적합하다고 판단하였다. 흥미성은 적합성 판단에서는 3.91%, 비적합성 판단에서는 0.34%의 비중으로 언급되어 큰 비중을 차지하지는 않았다.

8) 접근용이성

검색이용자들은 정보에 접근하거나 보기 어렵다는 점을 들어 적합하지 않다고 판단하였다. 접근용이성에 대한 응답 예로는, 로그인을 해야 해서(p7, 'general iterative scaling IIS'), 가입을 해야 정보를 볼 수 있어서(p8, 템플릿), 이미지가 깨져 있어서(p2, 이블랙), 검색결과가 크롤링 당시 내용과 달라서 찾기 어렵다(p1, 블랙베리 트위터) 등이 있었다.

접근용이성의 경우 예전에는 인터넷이 느리거나, 웹 문서에 이미지가 많거나 하는 경우 해당 웹 문서를 보기 힘든 경우가 많았으나, 오늘날 인터넷 환경의 급격한 발달로 대부분 좋은 하드웨어와 소프트웨어로 인하여 그 중요성이 약화된 것으로 유추할 수 있다. 본 연구에서 접근용이성은 비적합성 판단에서만 언급된 기준으로 2.7%를 차지하였다.

3.2.2 적합성 여부 판단 기준 빈도 분석

적합성 판단에 있어서 가장 두드러진 판단 기준은 특수성(28.26%)이었다. 응답자들이 찾는 문서의 형태(사이트, 이미지)에 해당하거나 원하는 특정 브랜드나 세부 내용에 초점을 잘 맞추고 있는 경우 적합하다고 판단하는 비중이 높았다. 그 다음으로는 구체성(21.74%)이었다. 즉 문서의 내용이 다양하고 상세하면서 충분한 경우 적합하다고 판단하였다. 세 번째는 신선성(15.65%)이었다. 예전에 몰랐던 새로운 정보를 제공할 경우 적합하다고 판단하였다. 네 번째는 이해가능성(11.74%)과 신뢰성이 같은 비중으로 언급되었다. 다섯 번째 기준은 주제성(6.96%), 여섯 번째 기준은 흥미성(3.91%)이었다.

<표 4> 적합성 판단 기준 빈도 분석

적합성 판단 기준	빈도	비율(%)
특수성	65	28.26
구체성	50	21.74
신선성	36	15.65
이해가능성	27	11.74
신뢰성	27	11.74
주제성	16	6.96
흥미성	9	3.91
합계	230	100.00

반면, 비적합성 판단에서 가장 두드러진 기준은 적합성 판단과 마찬가지로 특수성이었다(37.54%). 그리고 주제성(22.53%)이 그 뒤를 이었다. 주제성은 적합성 판단에서와 달리 비적합성 판단에 있어 매우 중요한 기준으로 사용되었음을 볼 수 있다. 응답자들은 주로 찾는 내용이 아니거나(주제성) 찾는 내용에 관한 것이라도 특정한 문서의 형태가 맞지 않거나 그 내용 중에서도 세부 주제가 다른 것이 아닐 때(비특수성) 적합하지 않다고 판단하였다. 예를 들면 “정하웅 GIS에 대한 내용이 맞긴 한데 기사가 아닌 논문을 원했어요”를 들 수 있다. 세 번째는 신선성(17.06%)이었다. 이미 본 문서이거나 아는 내용일 경우, 비적합성 판단 기준으로 사용되었다. 네 번째는 구체성(10.58%)으로 내용이 짧거나, 불만한 내용이 없을 때 응답자들은 또한 비적합성으로 판단하였다. 구체성은 적합성 판단에서와 달리 비적합성 판단에서는 약화된 것을 알 수 있다. 그 다음은 이해가능성(6.14%), 신뢰성(3.07%), 접근용이성(2.73%), 흥미성(0.34%)의 순이었다. 접근용이성은 비적합성 판단에서만 언급되었다.

〈표 5〉 비적합성 판단 기준 빈도 분석

비적합성 판단 기준	빈도	비율(%)
특수성	110	37.54
주제성	66	22.53
신선성	50	17.06
구체성	31	10.58
이해가능성	18	6.14
신뢰성	9	3.07
접근용이성	8	2.73
흥미성	1	0.34
합계	293	100.00

분석 결과를 종합해 보면, 한국 통합 검색 환경에서 이용자가 적합성 여부를 판단하는 기준은 주제성, 특수성, 구체성, 신선성, 이해가능성, 신뢰성, 흥미성, 접근 용이성의 총 8개로 나타났다. 이들을 기존 연구, 특히 적합성 판단 기준들을 비교 및 분류하여 10개의 주요 적합성 판단 기준을 정리한 Barry와 Scamber(1998), 그리고 비교적 최근 웹 검색 환경에서의 적합성 판단 기준을 연구한 Savolainen와 Kari(2006)의 연구와 비교하여 〈표 6〉으로 정리하였다.

본 연구에서 밝혀진 8개의 적합성 판단 기준은 기존 연구에서 정의되어 온 적합성 판단 기준에 모두 부합되는 것으로써 상황이나 환경에 관계없이 적합성 판단 기준 집합이 크게 변화되지 않음을 알 수 있었다. 그리고 많은 연구들에서 공통적으로 언급된 적합성 판단 기준인 주제성은, 본 연구에서 상대적으로 적합성 판단 기준에서보다(전체 7개의 기준 중 빈도 순으로 6번째), 비적합성 판단에서 높은 빈도(전체 8개 중 2번째) 많이 언급된 것으로 나타났다. 이를 통해 주제성은 적합성 판단의 기본 기준, 즉 필요 조건이지만 충분 조건은 아니다(Burgin, 1992; Cooper, 1971; Froehlich, 1994)라는 주장에 부합되는 것을 확인할 수 있었다.

마지막으로 본 연구에서 특수성이 적합성 판단과 비적합성 판단 둘 다에서 가장 많이 언급된 기준으로 드러났다. 이는 지식(Q&A), 블로그, 카페, 뉴스, 디렉토리, 쇼핑, 전문 자료와 같은 다양한 형태의 데이터를 동시에 검색해 주고, 문서 형태 별로 구분하여 검색결과를 제공해 주는 통합 검색 환경의 특징과도 연관시켜 볼 수 있다. 즉, 문서 형태로 나누어 정보가 제공됨으로써 이용자는 특정 형태 또는 내용의 정

〈표 6〉 기존 연구와의 적합성 판단 기준 비교

본 연구	Barry & Schamber(1998)	Savolainen & Kari(2006)
특수성	specificity depth/scope	specificity
주제성	-	topicality
신선성	currency	novelty currency
구체성	-	variety
이해가능성	-	ability to understand clarity
신뢰성	accuracy/validaity quality of sources verification	reliability validity
접근용이성	accessibility availability of information	accessibility
흥미성	-	curiosity
기타	affectiveness	affectiveness familiarity cost language security usability time constraints

보를 찾는 데 익숙할 수 있다. 또한 본 연구에서는 특별히 적합성 판단 기준에서 내용의 다양함, 상세함, 충분함 등의 언급이 두드러짐을 확인하였다. 구체성은 특수성 다음으로 가장 많이 언급된 기준으로 나타났다. 이는 데이터 형태의 다양성과 데이터 양 자체의 증가 등으로 인하여 이용자 요구가 더욱 특수화되고, 상세하며 구체화 되어감에 따른 현상과 연관 지어 생각해 볼 수 있다. 즉 다양한 형태 및 좋은 품질의 문서가 증가함으로써 정보 이용자들에게 특수성과 구체성이 점점 더 중요한 적합성 판단 기준으로 부각되는 것으로 해석 가능하다고 볼 수 있겠다.

3.2.3 기타 검색 행위

사람들은 이미지나 스크린 캡처 사진을 매우 유용하게 생각하였고 키워드를 아주 상세하게는 입력하지 않았다. 예를 들어 궁금한 내용이 참깨 다이아몬드의 가격이었는데도 검색어는 “참깨 다이아몬드”를, 블랙베리의 어플리케이션을 찾으면서도 “블랙베리 리뷰”를 입력하였다.

검색 결과에 관해서는, 통합 검색 결과 상단에 표시되는 스폰서나 광고 관련 검색 결과는 거의 보지 않았고, 주로 제목을 보고 클릭해서 볼 문서를 결정하였다. 중복 문서나 본 문서는 거의 클릭하지 않았다. 때로는 굳이 문서를 클릭하지 않고 검색된 문서의 요약 내용만 보고서도 원하는 내용을 알게 되기도 하였다.

어떤 주제에 대해서는 특정한 문서 형태를 원하기도 하였다. 예를 들어 영화 리뷰에 대한 내용은 “블로그” 정보를 골라서 보았다. 특별히 내용이 구체적이거나 새로운 정보를 추천해주는 문서를 만족스러워 했다. 또한 아는 내용이더라도 보기 편하게 잘 정리되어 있는 문서에도 만족도를 나타내었다. 이런 경우는 문서 자체의 품질이 좋은 이유로 만족스러워 하는 것으로 해석할 수 있을 것이다.

문서는 전문 지식 수준으로 정리된 문서보다, 비슷한 지식 수준의 사람들이 정리하거나 풀이해 준 것을 선호하였다. 이는 이해가능성의 차원으로 해석 가능하다. 본인의 상황이나 지식 수준에 맞는 글들을 찾을 수 있어서 일반 이용자들이 만든 블로그, 카페, 지식, 게시판 등과 같은 UGC(User Generated Contents) 정보를 선호하기도 하였다. 예를 들어 신종플루에 대해, 질병관리 본부나 보건복지 가족부에서 전문 용어를 이용하여 정의한 내용보다는 (예: 발열(37.8℃), 콧물, 인후통, 기침) 지식에 일반 사람들이 이해할 수 있는 용어로 설명된 내용을 선호하였다.

쇼핑 검색의 경우, 이미지를 보고 클릭 여부를 결정하였고, 내용이 상세하고 사실감 있는 문서를 선호하였다. 그리고 검색 옵션을 사용하는 대신 검색어로 정보를 필터링 하였다(예를 들어 “이불패드”를 입력한 후 “언더배드 이불패드”를 재입력). 쇼핑 검색 결과에서 맘에 드는 이미지의 모델명을 알고 나서 통합 검색에서 다시 모델을 검색하기도 하였다.

10명 대부분이 첫 번째 검색 결과 페이지를 보고 원하는 걸 찾지 못할 경우 두 번째 검색 결과 페이지를 보는 대신 검색어를 변경하였다.

원하는 내용이나 만족스러운 정보를 찾으면 대체로 신뢰하였다. 그리고 적합하지 않다고 생각하는 문서가 검색결과로 나오는 것에 크게 거부감을 갖지 않았다. 다른 사람에게는 적합할 수 있고 추가 정보를 줄 수 있다고 생각하였다.

대부분 구글과 다음이나 네이버와 같은 통합 검색을 용도에 따라 다르긴 하지만 함께 사용하였다. 구글에서는 객관적인 정보를 찾는 반면, 다음이나 네이버와 같은 통합 검색에서는 지식과 같은 주관적인 정보나 다양한 정보를 원할 때 이용한다고 답한 이도 있었다. 구글에서는 능동적으로 정보를 자세히 보게 되고, 다음이나 네이버는 보기 편한 만큼 훑어볼 수 있는 정보들 위주로 본다고 답변하기도 하였다. 예를 들어 다음이나 네이버와 같은 통합 검색에서는 웹 문서를 잘 보지 않는다고 답변하기도 하였다. 통합 검색의 두드러진 특징이라고 볼 수 있는 컬렉션 정보는 검색 후 특별한 컬렉션만의 정보를 보거나 또는 문서를 클릭할 때 어느 컬렉션의 문서인지를 보는 정도로 활용하였다. 대부분 원하는 내용을 찾지 못하면 바로 검색어를 바꾸거나 검색 엔진을 바꾸어 이용하였다.

4. 결론 및 제언

본 연구에서는 한국 통합 검색 환경에서 이용자가 적합성을 판단하는 기준들을 인터뷰라는 정성적(Qualitative) 연구 방법을 통하여 탐색적으로 알아보려고 하였다. 연구 결과 적합성 판단 기준으로는 주제성, 특수성, 구체성, 신선성, 이해가능성, 신뢰성, 흥미성의 7개가, 비

적합성 판단 기준으로는 접근용이성을 포함하여 총 8개의 적합성 판단 기준이 확인되었다. 이러한 연구 결과는 기존 연구에서 밝혀진 적합성 판단 기준에 모두 부합되는 것으로써, 상황이나 도메인의 변화가 있더라도 적합성 판단 기준 집합 자체는 크게 변하지 않는다는 기존 연구의 주장(Barry, 1994)에 부합되는 것을 확인할 수 있었다.

적합성 판단에서는 특수성, 구체성, 신선성, 이해가능성이 주요 기준으로, 비적합성 판단에서는 특수성, 주제성, 신선성, 구체성, 이해가능성이 주요 기준으로 사용된 것을 알 수 있었다. 특수성은 적합성 판단과 비적합성 판단에서 모두 가장 많은 비중을 차지한 판단 기준으로 나타났다. 적합성 판단 기준으로는 특수성과 구체성 순, 비적합성 판단 기준으로는 특수성과 주제성 순으로 나타났다.

많은 연구들에서 공통적으로 적합성 판단 기준으로 언급되어 온 주제성은, 한국 통합 검색 환경에서도 적합성 여부 판단 기준으로 사용되는 것을 알 수 있었다. 다만 적합하다고 판단하는 경우 비중이 매우 낮은 반면, 적합하지 않다고 판단하는 경우에 매우 높은 빈도로 나타나는 것을 통해, 주제성이 적합성 판단의 기본 기준, 즉 필요 조건이지만 충분 조건은 아니다(Burgin, 1992; Cooper, 1971; Froehlich, 1994)라는 기존 주장에 부합되는 것을 확인할 수 있었다. 즉 적합한 문서를 판단할 때, 주제성은 기본적으로 만족하는 조건으로써 별도로 언급되지 않고 다른 기준들이 결정적으로 적합성 판단에 영향을 미치는 반면, 기본적으로 주제성이 만족되지 않을 때 비적합성 판단을 내리게 된다고 유추해 볼 수 있었다.

본 연구에서는 특수성이 적합성 판단과 비적합성 판단 둘 다에서 가장 많이 언급된 기준으로 드러났다. 이는 지식(Q&A), 블로그, 카페, 뉴스, 디렉토리, 쇼핑, 전문 자료와 같은 다양한 형태의 데이터를 동시에 검색해 주고, 문서 형태 별로 구분하여 검색결과를 제공해 주는 통합 검색 환경의 특징과도 연관시켜 볼 수 있다. 즉, 문서 형태로 나누어 정보가 제공됨으로써 이용자는 특정 형태 또는 내용의 정보를 찾는 데 익숙할 수 있다. 또한 본 연구에서는 특별히 적합성 판단 기준에서 내용의 다양함, 상세함, 충분함 등의 언급이 두드러짐을 확인하였다. 구체성은 특수성 다음으로 가장 많이 언급된 기준으로 나타났다. 이는 데이터 형태의 다양성과 데이터 양 자체의 증가 등으로 인하여 이용자 요구가 더욱 특수화되고, 상세하며 구체화 되어감에 따른 현상과 연관 지어 생각해 볼 수 있다. 즉 다양한 형태 및 좋은 품질의 문서가 증가함으로써 정보 이용자들에게 특수성과 구체성이 점점 더 중요한 적합성 판단 기준으로 부각되는 것으로 해석 가능하겠다.

본 연구는 한국 통합 검색 환경에서의 이용자 중심 적합성 연구라는 점에서 의의를 가진다. 한국은 통합 검색이라는 독특한 검색 환경을 가지고 있고, 이는 국내에서 가장 많이 사용되는 검색 서비스이면서 동시에 국외에서도 주목 받고 있는 검색 서비스이다(박소연, 이준호, 2008). 그럼에도 불구하고 아직까지 한국 통합 검색 환경에서의 이용자 중심 적합성에 관한 연구를 찾기 어려운 실정이다. 본 연구는 한국 통합 검색 환경에서 이용자 적합성을 연구를 통해, 이제껏 연구되지 않은 새로운 환경(context)에서 이용자 적합성 판단 기준을 알아봄으로써 이

용자 적합성 연구 상황을 확장하였다는 점에서 의의를 가진다.

향후 연구로는, 정보의 유형이나 정보 검색 과제 등 컨텍스트를 고려한 이용자 적합성 판단 기준에 관한 확장 연구를 통해, 이용자 적합성 연구를 시스템 개선에 적극 활용할 수 있을 것이다. 예를 들어 적합한 정보 노출을 결정하는 정보 검색 시스템의 랭킹 모델에 이용자 적

합성 판단 기준을 고려하는 등과 같이, 이용자 연구 접목을 통한 시스템 개선 노력을 통해 궁극적으로 사용자의 경험을 향상시킬 수 있을 것이다. 결론적으로 본 연구 결과는 한국 통합 검색 환경에서 이용자 적합성 판단 기준을 제시함으로써 이용자 중심 적합성에 대한 이해를 높이고, 향후 효과적인 정보 검색 시스템 디자인 및 개발에 활용될 수 있을 것으로 기대된다.

참 고 문 헌

- 박소연, 이준호 (2008). 주요 검색 포털들의 통합 검색 서비스 비교 평가. 한국도서관·정보학회지, 39, 265-278.
- Barry, C. L. (1994). User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science*, 45, 149-159.
- Barry, C. L., & Schamber, L. (1998). Users' criteria for relevance evaluation: A cross-situational comparison. *Information Processing & Management*, 34, 219-236.
- Bateman, J. (1998). Changes in relevance criteria: A longitudinal study. In: *ASIS Proceedings*, 1998, 23-32.
- Bilal, D. (2000). Children's use of the Yahoo!igans! web search engine: I. Cognitive, physical and affective behaviors on fact-based search tasks. *Journal of the American Society for Information Science*, 51, 646-65.
- Borlund, P. (2003). The concept of relevance in IR. *Journal of the American Society for Information Science and Technology*, 54, 913-925.
- Choi, Y., & Rasmussen, E. M. (2002). Users' relevance criteria in image retrieval in American history. *Information Processing & Management*, 38, 695-726.
- Cooper, W. S. (1971). A definition of relevance for information retrieval. *Information Storage and Retrieval*, 7, 19-37.
- Cosijn, E., & Ingwersen, P. (2000). Dimensions of relevance. *Information Processing & Management*, 36, 533-550.
- Cuadra, C. A., & Katter, R.V. (1967a). Experimental studies of relevance judgments: Final report.

- Volume 1: Project summary (TM-3520/001/00). Santa Monica, CA: System Development Corp.
- Cuadra, C. A., & Katter, R.V. (1967b). Opening the black box of "relevance". *Journal of Documentation*, 23, 291-303.
- Fitzgerald, M. A., & Galloway, C. (2001). Relevance judging, evaluation, and decision making in virtual library: A descriptive study. *Journal of the American Society for Information Science and Technology*, 52, 989-1010.
- Freund, L. (2008). Exploiting task-document relations in support of information retrieval in the workplace. PhD thesis, University of Toronto.
- Froehlich, T. J. (1994). Relevance reconsidered: Towards an agenda for the 21st century: Introduction to special topic issue on relevance research. *Journal of the American Society for Information Science*, 45, 124-134.
- Greisdorf, H. (2003). Relevance thresholds: A multi-stage predictive model of how users evaluate information. *Information Processing & Management*, 39, 403-423.
- Hirsh, S. G. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science*, 50, 1265-1283.
- Hjørland, B., & Christensen, F. S. (2002). Work tasks and socio-cognitive relevance: A specific example.
- Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills, CA: Sage.
- Maglaughlin, K. L., & Sonnewald, H. (2002). User perspective on relevance criteria: A comparison among relevant, partially relevant, and not-relevant. *Journal of the American Society for Information Science and Technology*, 53, 327-342.
- Mizzaro, S. (1997). Relevance: The whole history. *Journal of the American Society for Information Science*, 48, 810-832.
- Moon, I., & Woyke, E. (2006, January 30). NHN: The little search engine that could. *Businessweek*.
- Park, T. K. (1993). The nature of relevance in information retrieval: An empirical study. *Library Quarterly*, 63, 318-351.
- Park, T. K. (1994). Toward a theory of user-based relevance: A call for a new paradigm of inquiry. *Journal of the American Society for Information Science*, 45, 135-141.
- Park, H. (1997). Relevance of science information: Origins and dimensions of relevance and their implications to information retrieval. *Information Processing & Management*, 33, 339-352.
- Rees, A. M., & Schultz, D. G. (1967). A field experiment approach to the study of relevance

- assessments in relation to document searching, 2. Cleveland, OH: Center for Documentation and Communication Research, School of Library Science, Case Western Reserve University.
- Saracevic, T. (1975). Relevance: A review of and a framework for the thinking on the notion in information science. *Journal of the American Society for Information Science*, 26, 321-343.
- Saracevic, T. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part III: Behavior and effects of relevance. *Journal of the American Society for Information Science and Technology*, 58, 2126-2144.
- Savolainen, R., & Kari, J. (2006). User-defined relevance criteria in web searching. *Journal of Documentation*, 62, 685-707.
- Schamber, L. (1991). Users' criteria for evaluation in a multimedia environment. *Proceedings of the American Society for Information Science*, Washington, DC, 126-133. Medford, NJ: Learned Information, Inc.
- Schamber, L. (1994). Relevance and information behavior. In M.E. Williams (Ed.), *Annual review of information science and technology (ARIST)*, 3-48. Medford, NJ: Learned Information, Inc.
- Schamber, L., & Bateman, J. (1996). User criteria in relevance evaluation: Toward development of a measurement scale. *Proceedings of the 59th Annual Meeting of the American Society for Information Science*, 33, 218-225. Medford, NJ: InformationToday.
- Schamber, L., Eisenberg, M. B., & Nilan, M. S. (1990). A re-examination of relevance: Toward a dynamic, situational definition. *Information Processing & Management*, 26, 755-775.
- Taylor, A. R., Cool, C., Belkin, N. J., & Amadio, W. J. (2007). Relationships between categories of relevance criteria and stage in task completion. *Information Processing & Management*, 43, 1071-1084.
- Wang, P., & Soergel, D. (1998). A cognitive model of document use during a research project: Study I. Document selection. *Journal of the American Society for Information Science*, 49, 115-133.
- Wang, P., & White, M. D. (1999). A cognitive model of document use during a research project: Study II. Decisions at the reading and citing stages. *Journal of the American Society for Information Science*, 50, 98-144.
- Xu, Y., & Chen, Z. (2006). Relevance judgment- What do information consumers consider beyond topicality? *Journal of the American Society for Information Science and Technology*, 57, 961-973.
- Yang, M., & Marchionini, G. (2004). Exploring users' video relevance criteria - A pilot study.

Proceedings of the 67th annual meeting of the American Society for Information Science and Technology (ASIST '04), 229-238. Medford, NJ: Information Today.

- 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

Park, Soyeon, & Lee, Joon-Ho. (2008). Comparative Evaluation of the Unified Search Services Provided by Major Korean Search Portals. Journal of Korean Library and Information Science Society, 39, 265-278.

