

# 적합성 평가기준 변화에 관한 실험 연구

## A Study on the Variation of Criteria for Relevance Judgement of Retrieved Documents

김 홍 렬(Hong-Ryul Kim)\*

### 〈 목 차 〉

- |                   |                          |
|-------------------|--------------------------|
| I. 서 론            | IV. 적합성 평가기준 변화에 대한 실험조사 |
| 1. 연구의 필요성 및 목적   | 1. 가설의 설정                |
| 2. 연구의 범위 및 방법    | 2. 실험조사방법                |
| II. 적합성평가의 이론적 배경 | 3. 실험 및 결과분석             |
| 1. 시스템 지향적 평가     | V. 실험결과에 따른 가설검증         |
| 2. 이용자 지향적 평가     | 1. 적합여부와 평가기준과의 상관관계     |
| III. 적합성 평가기준의 설정 | 2. 적합여부평가의 인식검증          |
| 1. 적합성 평가기준의 분석   | 3. 적합성 평가기준의 인식검증        |
| 2. 평가기준의 설정       | VI. 결 론                  |

### 초 록

본 연구는 이용자 지향적 평가의 의의와 타당성을 규명하기 위하여 연구과정에서 개인의 지식과 상황 및 인지적 요인의 변화에 따라 이용자가 인식하는 검색문헌의 적합성을 평가하는 기준들의 변화를 규명하는데 그 목적이 있다. 연구의 방법은 이론가설을 설정하고 이의 검증을 위하여 실제 연구문제를 가지고 있는 5명의 이용자를 대상으로 그들의 연구과정에 따라 적합성을 평가하는 기준들과 그 기준들의 변화양상을 관찰하는 실험을 수행하였다. 실험결과 분석에는 빈도분석, 상관분석, 일원분산분석이 활용되었다. 그 결과 연구자들은 연구과정에 따라 적합성을 평가하는 기준이 다르게 나타났으며, 평가자에 따라 서로 다른 평가기준이 사용되고 있음을 확인할 수 있었다.

주제어 : 검색성능, 적합성 평가, 적합여부, 정보검색시스템 평가, 평가기준

### Abstract

This research was an attempt to observe the variation of criterions for relevance judgement of retrieved documents regarding the changing nature of the individual's information need situations and dimensional view users' information problem, and its subsequent impact on the various stages of research. Users were asked to mark the portions of the document representations and evaluating criteria for relevance of documents that indicated to the users whether they would or would not pursue the information within documents according to their research stages. Responses were analysed using the SPSS for windows version, and the perception differences of evaluation criterions at various research stages were examined by correlation analysis and ANOVA.

Key Words : evaluating criteria, information retrieval system, relevance judgement

\* 동원대학 문헌정보과 전임강사. kykim@tongwon.ac.kr

## 1. 서 론

### 1. 연구의 필요성 및 목적

정보검색시스템의 성능은 일반적으로 재현율과 정확률에 기반을 두고 있다. 이러한 재현율과 정확률은 검색문헌의 적합성 평가에 의해 측정된다. 일반적으로 적합성 평가는 이용자의 정보요구와 검색된 문헌간의 유사 정도를 측정하는 것으로서 재현율과 정확률은 정보검색시스템의 성능에 결정적인 영향을 미치게 된다. 적합평가에 관한 연구들은 적합성을 다양한 관점에서 정의하고 있다. 먼저 전통적인 시스템 지향적 관점에서 보는 적합성의 개념은 이용자의 정보요구를 표현한 질문과 검색된 특정 문헌이 서로 주제적으로 어느 정도 관련이 있는가에 따라서 결정되는 것으로 객관적 적합, 또는 topicality로 표현되기도 한다.

한편 이용자 지향적 관점에서의 적합성의 개념은 질문서에 표현되지 않은 이용자의 정보요구와 검색문헌과의 관련성으로 보고 있다. 일반적으로 이것은 적합평가에 있어서 개인적 특성을 고려하는 것으로서, 연구자에 따라 적절성, 상황적 적합, 심리적 적합, 유용성, 이용자 만족도 등으로 다양하게 표현되기도 한다. 이러한 관점을 지지하는 연구자들은 정보를 요구한 이용자가 검색시스템의 성능을 평가하는 주체가 되어야 한다는 것을 강조하고 있다. 그러나 이용자에 의한 적합성 평가는 이용자의 지식과 인식에 의존하는 인지적 처리과정과 이용자의 정보요구가 상황에 따라 지속적으로 변화하는 동적인 과정에서 이루어지고 있다는 데 문제가 있다(Barry 1994). 즉 정보탐색과 이용상황에 따라 적합평가가 달라질 수 있다는 것이다. 특정한 질문에 대한 동일한 검색문헌일지라도 이용자의 개인적인 상황이나 검색문헌이 제시되는 시점에 따라 적합문헌으로 평가될 수도 있고 부적합 문헌으로 평가될 수도 있다는 것이다. 예를들면, 이용자가 연구를 시작하여 완료하기까지 정보를 추구하는 과정에서 요구 정보에 대한 초기의 지식상태는 변화될 것이며, 이러한 지식상태의 변화로 인하여 연구를 시작하는 때에 판단한 적합평가는 연구를 종료하는 때 판단되는 적합평가와 달라질 것임에 틀림없다는 것이다. 이러한 이용자의 상황적인 요인들에 의한 영향은 적합 평가의 일관성 확보를 매우 어렵게 만들기 때문에 종래의 검색성능 평가방법에 대한 근본적인 문제가 제기되고 있다(Harter 1992). 따라서 이용자의 인식상태, 지식등의 상황적 요인에 의하여 발생하는 적합성 판단의 변화정도를 규명하여 시스템 지향적 평가의 대안으로 제기되는 이용자 지향적 평가의 의의와 타당성을 규명해야 할 필요가 있는 것이다. 본 연구는 이러한 문제에 착안하여 연구과정에서 개인의 지식과 상황 및 인지적 요인의 변화에 따라 이용자가 인식하는 검색문헌의 적합성을 평가하는 기준들의 변화를 규명하는데 그 목적이 있다.

## 2. 연구의 범위 및 방법

본 연구에서는 연구의 목적을 달성하기 위하여 다음과 같은 이론 가설을 설정하였다. 첫째, 연구과정에 따른 문헌의 적합여부와 평가기준들사이의 유의적인 상관관계를 보일 것이다. 둘째, 연구과정에 따라 적합여부에 영향을 미치는 평가기준들의 인식도에는 차이를 보일 것이다. 이 가설들을 검증하기 위하여 본 연구는 문헌조사와 관찰에 의한 실험연구를 병행하였다. 먼저 문헌조사를 통하여 본 연구문제의 이론적배경과 이용자 지향 적합성 평가기준들을 종합적으로 분석하고, 실험에 사용할 평가기준들을 도출하였다. 관찰에 의한 실험연구에서는 이용자의 연구과정을 정보입수자가, 관련정보를 수집하는 최초의 과정(이후 시작과정)과, 수집정보를 정리하고 초점을 형성, 집필을 시작하는 과정(이후 진행과정), 본격적인 집필과정으로 인용문헌을 결정하는 과정(이후 완성과정)으로 나누어 평가기준의 변화를 관찰하였다. 적합성 판단행위의 관찰 대상은 실제 연구문제를 가지고 있는 이용자중에서 실험에 자발적인 참여 의사를 밝힌 5명을 선정하였다. 그 이유는 연구문제를 해결하는 방법으로 문헌조사를 수행하는 이용자만을 대상으로 그들의 연구과정에 따라 지속적으로 적합성 평가기준과 인식 변화를 관찰해야 하는 본 실험의 성격으로 인하여 다수의 이용자를 대상으로 실험을 수행하기란 매우 어렵기 때문이다. 따라서 자발적으로 실험에 응한 5명을 대상으로 KINITI의 BIST 데이터베이스를 통하여 직접 검색한 문헌을 평가하게 하였다. 적합성 평가에 사용된 척도는 적합/부분적합/부적합의 3등급척도를 사용하였으며, 관찰기간은 1998년 12월부터 1999년 5월까지 약 6개월에 걸쳐 진행되었다. 결과데이터의 분석에는 SPSS Windows 버전 7.0을 사용하여 빈도분석, 상관관계분석, 분산분석(ANOVA)을 실시하였다. 본 연구에서는 정보검색시스템의 성능을 평가하는데 사용되는 적합성 평가관점중에서 시스템 지향적 관점의 적합성 판단은 제외하고 차후의 연구과제로 미루었다.

## II. 적합성평가의 이론적 배경

일반적으로 적합성은 끊임없는 논쟁을 불러일으킬 수 있는 다양한 의미를 지닌 개념이다. 따라서 연구자들 사이에서 적합성의 의미와 개념에 대한 일치된 이해를 구하기는 매우 어렵다. 일반적으로 이해되고 있는 적합성은 이용자의 정보요구와 요구를 해결할 수 있는 정보사이의 관계로 보는 이론적인 개념이면서 동시에 정보시스템으로 부터 검색된 정보를 받아들

일 것인지 아니면 거부할 것인지의 여부에 대한 이용자의 결정과 관련있는 행위적인 개념이다. 그러나 이용자가 어떠한 측면 또는 관점에서 적합정보를 판단하는지에 대해서는 말하기 어렵다는 것이다.

따라서 적합성평가와 관련한 대부분의 연구자들은 적합성의 의미와 관련하여 다른 종류의 적합성 개념을 구축하고 사용하였다. 물론 이러한 원인에는 초기의 적합성 연구에서 적합성의 개념을 결정하는 구성요소로 문헌과 이용자로 구분하고, 이 두가지 구성요소를 개별적으로 연구한데서 비롯된다. B. Vickery(1959)는 이들 요소를 주제에 관한 적합성(후에 시스템 적합성, 정보소스에 대한 적합성, 문헌적합성으로 명명)과 이용자 적합성(후에 주관적 적합성으로 명명)으로 나누어 언급하였다. 그 후 Vickery가 제시한 두가지 요소를 대변하는 적합성 개념, 즉 시스템 지향적 적합과 이용자 지향적 적합이 정보검색시스템의 성능평가에서 수행하는 특징과 차이점, 그리고 관련성을 규명하기 위한 연구들이 지속적으로 수행되었다.

본 장에서는 정보검색실험에서 적합성 판단행위를 살펴보는데 이론적 배경이 되는 시스템 지향적 평가와 이용자 지향적 평가의 관점 및 장단점 등에 대해 고찰하고자 한다.

## 1. 시스템 지향적 평가

1950년대에 정보과학이 논리학, 철학에 이어서 적합성의 개념을 수용한 이래, 적합성의 의미와 관련하여 그 정의를 밝히려는 노력이 1960년대에 들어 매우 다양하게 시도되고 있는데 그 주요 내용은 다음과 같다. C. Cuadra와 R. Katter(1967)는 적합성을 정보요구와 이용성향 사이에 관련되는 그 무엇으로 보고, 이용자의 정보요구를 반영한 정보요구서와 정보요구서에 적합한 자료를 수록하고 있는 아티클 사이의 일치로 정의하고 있다. Rees & Schultz (1967)는 적합성을 이용자의 정보요구 및 유용성과 검색된 문헌과의 관계, 그리고 이용자의 개인적인 특성과 검색된 문헌과의 관계로 정의하고 있다. 즉 정보요구에 대한 검색된 객체와의 관계로 적합을 보고 있는 것이다.

한편 이러한 다양한 적합성의 개념적 정의들의 출현 현상에 대하여 Saracevic(1975)은 적합성이라는 용어가 안고 있는 의미의 불분명성에 기인하는 것으로 보고 있다. 그러나, 적합성 개념에 대한 정의상의 혼란에도 불구하고 초기의 정보검색연구에서 사용된 다수의 관점은 이용자의 정보요구를 표현한 질문과 검색된 문헌과의 관계를 나타내는 시스템 지향 관점이었다. 그러한 배경에는 1958년에 개최된 국제학술정보회의(ICSI)에서 적합성의 개념을 시스템 지향적인 관점에서 주장하고 있는데서 비롯된다. 이 회의에서는 적합성을 질문에 나타나는 용어가 문헌에 출현하는 용어사이의 대조로 보고 있는데, 이는 일반적으로 'topicality'로

알려지고 있다. 이것은 검색된 문헌의 주제와 검색질의가 서로 주제적으로 부합될 경우 검색 질의와 검색문헌이 서로 적합한 것으로 간주하는 것이며, 적합성 평가의 결과는 시스템과 관련된 요인들에 의존하고 있다. 이는 정보검색시스템의 성능 평가에 대한 최초의 연구로 간주되는 크랜필드 실험(Cranfield test)에서 적합성을 ‘논제상의 (on the topic)’, ‘주제상의 (on the subject)’, 그리고 ‘aboutness’ 의미로 보는 것과 상통한다. Eisenberg & Schamber(1988)는 이러한 적합성의 개념을 ‘주제적 적합성’이라고 명명하였는데, 검색된 문헌의 주제와 요구한 질문의 주제가 서로 부합되는가의 여부를 측정하는 측면이다. 따라서 주제적으로 적합하다고 보는 것은 검색질문이 문헌의 주제를 다루고 있기 때문이며, 주제는 가장 정확하게 검색문헌의 적합성을 보편적으로 대변한다는 가정을 기초로 하고 있다.

전통적으로 서지검색시스템에서 사용된 주된 검색 메커니즘은 문헌을 기술하기 위하여 사용된 주제어와 정보요구를 기술하기 위하여 사용된 주제어의 매칭에 기반을 두고 있다. 이들 시스템의 평가는 이용자의 요구를 만족시키는 시스템의 능력에 기초한다. 즉 평가는 이용자의 현재의 정보요구 상황에 적합한 문헌들의 레퍼런스를 검색할 수 있는 시스템의 성능을 검증하는 것이다. 이러한 시스템의 성능에 대한 평가행위를 정당화시키기 위해서는 다음과 같은 가정이 전제되어야 한다고 Barry(1994)는 밝힌 바 있다. 첫째, 질의에 사용된 주제어는 이용자의 정보요구상황을 적절하게 기술할 수 있어야 한다. 둘째, 문헌에 포함된 주제어는 문헌의 내용을 적절하게 기술할 수 있어야 한다. 셋째, 적합문헌의 검색에서 반드시 주제는 검색결과들과 매칭해야 한다. 이러한 가정은 적합성의 정의에서 주제적 적합성에 대한 기초를 형성하고 있다. Saracevic(1970)은 이를 정보검색의 ‘시스템 관점’으로 설명하고 있다. 이 관점에서 질문과 검색문헌의 관련여부는 시스템의 내적인 메커니즘의 속성에서 결정되며, 질의의 주제어와 문헌에 할당된 주제어 사이에 대조의 결과로 언급되고 있다. 따라서 이용자에게 적합한 모든 문헌은 시스템에 의하여 정확하게 검색되며, 검색되지 않은 문헌은 적합문헌이 아닌 것으로 믿고 있다. 이것은 이용자가 지닌 정보요구의 내적인 의미는 구체화시킬 수 없고 관찰이 불가능하기 때문에 내적의미의 정확한 표현에 의존할 수밖에 없다는 사실에 근거한다(Belkin 1980). 따라서 주제용어는 문헌에 포함된 정보와 질의로 공식화된 이용자 정보요구의 의미로 사용할 수 있고, 이러한 용어들은 이용자의 정보요구에 매치시키기 위하여 시스템이 사용할 수 있다는 것이다. 그러므로 시스템 관점은 정보를 요구하는 이용자의 역할을 무시하고 오로지 질문과 검색되는 문헌요소와의 효율적인 매칭에 기반하는 시스템의 내적 메커니즘에 초점을 맞추고 있다. 이러한 시스템 지향의 주제적 적합 관점은 오늘날까지 정보검색 효율성을 평가하기 위한 주요한 기준으로 사용되고 있다. 따라서 주제적 적합을 개선시키는 것이 효율적인 정보검색을 달성하고 시스템의 성능을 개선시키기 위한 최우선적인 과제이다. 그러므로 정보검색 연구의 초기부터 정보학자들은 이용자에게 적합한 문헌을 검색하기 위한 최적의 정보시스템의 설계는 색인과 탐색조작에서 주제용어들 사이의 관계에 보다

많은 주의를 기울임으로써 달성되어질 수 있다는 전제하에 행동해 왔다. 오늘날에도 역시 'topicality'를 개선시키기 위하여 불리언 연산자의 사용이나 용어 가중치, 용어의 클러스터링 등의 기법을 통하여 최선의 용어 매칭에 더욱 관심을 기울이고 있는 것도 이 때문이다. 그러나 시스템 지향 관점은 더욱 복잡한 모델의 정보검색시스템을 개발하고 문헌들간의 관계를 보다 효율적으로 표현할 수 있다는 장점은 가질 수 있지만, 정보검색시스템이 실제 이용자의 정보요구에 대해서는 질의에 표현된 정보요구 이상으로 이용자를 충족시킬 수 없는 단점이 있다(Belkin 1980)는 주장에 주목할 필요가 있다. 그러므로 정보검색이 문헌과 질문의 논리적인 관계 이상으로 이용자와 시스템사이, 정보와 정보요구사이의 역동적인 상호과정으로 해석된다면 시스템 지향적 관점의 평가요소인 'topicality'는 분명히 적합성의 개념적인 정의로는 불충분한 것으로 볼 수 있다. 실제 정보검색에서는 이용자가 시스템에 표현하는 요구를 검색된 문헌이 만족시키는지 아니면 그 요구에 유용한 것인지는 전적으로 이용자만이 결정할 수 있다. 또한 이용자에게 보다 정확하고 적합한 정보를 제공해야 하는 정보검색시스템의 궁극적인 목적을 달성하기 위해서도 이용자측면의 요소들을 적합성 평가에 고려해야 한다. 이러한 점으로 인하여 주제적 적합에 의한 검색실험의 신뢰성이 이용자 지향 적합 평가를 주장하는 사람들에 의하여 상당한 의문을 받게 되었던 것이다.

이상에서 살펴본 바에 따르면, 시스템 지향 관점은 질문과 검색문헌과의 관련여부는 질문과 가공된 색인어간에 시스템의 매칭기능을 통해 적합과 부적합이 결정됨을 의미하고 있다. 따라서 검색시스템이 수행하는 적합성 판단을 인간이 내리는 경우 타당성이 결여될 수밖에 없다는 측면으로서, 단지 검색시스템의 매칭성능을 극대화시킬 수 있는 노력만이 부적합 문헌의 검색은 최소화시킬 수 있다는 정보검색과 서비스 측면을 강하게 주장하고 있다.

그러나 검색시스템은 정보를 요구하고 검색된 정보를 활용하는 이용자를 만족시켜야 한다는 또 다른 측면도 동시에 고려해야만 한다. 즉 이용자의 요구를 시스템이 반영할 수 있을 때 검색시스템은 더욱 높은 성능을 가질 수 있을 것이다. 다음은 적합성 평가에서 인간의 주관적인 판단을 반영하는 이용자 지향적 적합성 관점에 대하여 살펴보고자 한다.

## 2. 이용자 지향적 평가

Cuadra & Katter(1967)은 적합성 평가과정과 연관없이 단지 일련의 용어들의 관계로서 구조화된 적합성의 개념은 별로 유용한 정의가 아니라고 언급하고 있다. 따라서 지난 수십년간 많은 연구자들이 적합성의 개념을 인간의 적합평가과정과 관련지어 이용자 지향의 적합성 정의를 밝히려는 시도를 해 왔지만, 일치된 정의의 도출은 커녕 개념에 대한 다양한 용어의 사용으로 혼란만 가중시키는 결과를 가져왔다. 이들 연구에서 주로 기술되는 적합성의 정의

는 개념적 관련성, 유용성, 상황적 적합성, 유의성, 적절성, 만족 등이다. '개념적 관련성'은 정보요구에서의 개념들과 검색문헌내의 개념들 사이의 근본적인 연결 및 합치로 언급된다. 이러한 개념적 관련성을 적합성의 정의에 포함하고 있는 주요한 연구자들은 Cuadra & Katter(1967), Rees & Schultz(1967), Saracevic(1970)등이다. 또한 개념적 관련성을 정보가 단순한 topicality 이상으로 무엇에 관한 어떤 아이디어를 줄 수 있다는 의미에서 'aboutness'의 의미와 상통하고 있다. 'Aboutness'는 시스템에서 주제적 검색을 목적으로 문헌의 내용에 대한 용어들을 색인자들이 부여할 때 발생할 수 있는 개념이다. Cuadra & Katter(1967)는 적합성에 영향을 미치는 하나의 요인으로 유용성을 고려하고 있다. 그들은 이를 내적인 이용지향, 또는 문헌의 의도된 이용에 대한 평가자들의 태도로 정의하고 있다. Rees & Schultz(1967)는 적합성과 유용함은 개별적이지만 그 범주가 중복된다고 언급하면서, 적합성은 문헌과 평가자의 정보문제영역사이의 관계로, 그리고 유용함을 평가자의 관심, 지식, 경험 등을 포함하는 문헌, 연구, 평가자 사이의 관계로 규정하였다. Saracevic(1970)은 유용성을 실용적이거나 비용효과의 관점에서 표현되는 적합성의 정의로 논의하였다. 이 관점에서 그는 정보검색시스템은 적합한 정보를 제공할 수는 있지만, 그 정보는 행위에 직접적으로 영향을 미칠 수 있고, 직면한 문제를 해결하는 데 도움을 주는 정보, 즉 유용성을 가진 정보여야 한다는 것이다. Cooper(1971)는 '논리적 적합성'을 주장하면서 적합성과 유용성 사이의 특징을 구분하였는데, 적합성은 'aboutness (pertinence 또는 topic-relatedness)'와 관련있으며, 유용성은 논제 연관성 뿐만 아니라 질적인 측면, 새로움, 중요성, 신뢰성 등의 개념들을 포함하는 광범위한 의미로 보고 있다. 따라서 그는 검색시스템은 단지 적합한 문헌이 아니라 유용성이 높은 문헌을 검색할 수 있는 능력을 가져야 함을 주장하였다. 또한 P. Wilson(1973)은 Cooper의 논리적 적합성 개념에 개인의 관심, 기호, 지식상태를 통합한 '상황적 적합성'을 제기하였다. 그는 이용자는 다양한 관점과 상황에 기초하여 정보를 평가하므로 이용자의 상황에 적합하면 그 정보는 주관적으로 적합한 것으로 평가하는 경향이 있다고 주장하였다. D. Swanson(1986)은 정보를 요구한 개인의 정신적 경험으로 주관적 적합성을 설명하고 있는데, 이는 검색된 문헌과 정보요구자의 정신적 경험과의 관계 측면에서 적합성의 해석을 시도하는 것으로 볼 수 있다. 또한 이용자 지향 적합과 적합평가는 유용성과 이용자의 주관적 지식 상태에 따른 판단을 결합한 적절성이란 용어로 해석하여 이용자의 정보요구 상황에 적절하지 않으면 적합정보가 아니라고 해석하는 연구자도 있다. Foskett(1972)에 의하면 이용자는 개인적 지식에 기초하여 검색결과를 해석하고, 이용자가 원하는 문헌은 이용자의 상황적 관점이나 요구에 도움을 줄 수 있는 문헌이라야 한다는 것이다. 그는 또한 적절성과 적합성 사이의 서로 다른 특성을 각각 개인적 지식과 공공의 지식에 비유하여 제시하고 있다.

이와 같이 적합성 평가의 관점이 시스템 지향적 평가에서 이용자 지향적 적합로 변화함에 따라 적합성 평가에 관심을 가지는 연구자들은 주제적 적합이외에 다양한 주관적인 요인들

이 평가과정에 영향을 미치고 있음을 밝히고 있다. 이러한 사람들은 적합성 평가를 이용자의 정보탐색 행위에 초점을 맞추고 있다. 즉, 탐색과정에서 이용자가 지각하거나 느낄 수 있는 모든 상황적 요소들을 종합적으로 포함할 수 있는 정보검색시스템의 평가가 이루어져야 한다는 것이다. 요약해볼 때, 시스템 지향적 관점이 검색질문과 관련하여 판단할 수 있는 문제라면 이용자 지향적 적합은 이용자의 정보요구와 관련한 개인적 특성과 검색문헌의 인식을 모두 관련시켜 판단할 수 있다. 즉, 영역의 지식, 교육 및 주제배경, 탐색목적, 연구과정 등의 정보요구자의 특성과 문헌요소의 종류, 검색문헌의 판단기준등 이용자의 적합성 평가와 관련된 영향요인들을 종합하여 판단할 수 있다고 보는 것이다. 그러나, 이용자 지향적 적합의 신뢰성에 의문을 던지는 측면도 없지 않다. 유재옥(1996)은 이용자 지향적 적합은 이용자 측면에서 검색문헌의 적합여부를 검토하는 것임에는 분명하지만, 이를 뒷받침 할 수 있는 이론적인 배경이 상당히 부족하다고 보고 있다. 또한 이용자가 적합하다고 선언하면 적합문헌이 된다는 것은 이용자 지향적 적합의 모호성을 단적으로 표현한다는 지적이 있다. 즉 특정 이용자가 검색된 문헌을 적합하다고 느끼는 것은 그 이용자의 독특한 인식일 뿐이지, 검색된 문헌이 내포하는 정보에 의해서 일어나는 것은 아니라는 것이다(Schamber, Eisenberg and Nilan 1991). 종합해 볼 때, 이용자 지향적 적합성 개념은 검색시스템의 성능평가에 분명히 영향을 미치고 있으며, 이들 영향 요인을 반영한 성능평가가 이루어지는 것은 지극히 당연하다. 그러나, 이용자의 정보요구와 관련된 모든 요인들을 고려하여 정보검색실험을 수행하기란 결코 쉬운 일이 아니다. 즉 이용자 지향적 적합성 개념은 이용자의 정보요구와 관련된 요인들이 변할 수 있다는 가정을 기반으로 하기 때문에 이를 검색시스템의 효율성 평가에 통합시키기란 대단히 어렵다는 것이다. 따라서 현재까지의 정보검색실험에서는 전통적인 시스템 지향적 관점에 의존한 검색실험을 수행하였고, 이용자 지향적 적합성 개념을 통합하여 검색실험을 수행한 연구가 거의 나타나지 않는 것을 보아도 알 수 있다.

### III. 적합성 평가기준의 설정

적합성 평가기준이란 일반적으로 이용자가 검색된 문헌의 적합성을 판단할 때 사용하는 요소들을 말한다. 적합성을 평가할 때 이용자의 상황적 요인에 의하여 변화될 수 있기 때문에 고정된 판단기준을 도출하기란 매우 어렵다. 이러한 이유로 종래의 검색실험에서는 적합성 평가시 이용자가 직접 언급한 기준들을 밝히는 연구들이 주를 이루었다. 그러나 이용자가 언급하는 평가요소들은 연구의 성격이나 목적에 따라 다양하게 언급될 수 있다. 따라서 본



연구의 목적인 검색과정에서 동일한 평가요소들에 대한 인식변화를 관찰할 수 없다는 문제가 제기될 수 있다. 이러한 점에서 본 연구는 관련연구에서 도출된 적합성 평가기준을 활용하여 평가실험을 수행하기로 하였다.

## 1. 적합성 평가기준의 분석

이용자 차향적 관점에서 적합성을 평가한 연구들에서 발견되는 개별적인 적합성 평가요소들은 대체로 통합적으로 기술되어 있는 것이 대부분이다. 그러나 인간의 적합성 판단행위에 영향을 미치는 요소들을 기술하고 있는 초기의 연구들(Cuadra and Katter 1967, Rees and Schultz 1967)과 Schamber(1994)의 연구에서 이들 요소들을 몇 개의 범주로 나누어 적합성 평가의 모델로 제시하고 있다. 그들은 다양한 요소들을 판단자와 요구(이용자특성과 관련있는 요소), 문헌(정보내용과 관련있는 요소), 정보시스템과 환경조건(물리적 환경과 관련있는 요소), 척도의 선정(적합성 측정도구)등 여섯 개의 범주에서 명명하였다. 본 연구에서는 평가기준군의 특성과 차이점을 세부적으로 관찰하기 위하여 동일한 범주에 속하는 평가요소들을 위에서 제시한 범주로 나누어 분석하고자 하였다. 다만, 범주의 명칭을 본 연구에서는 편의상 정보내용기준군, 이용자특성기준군, 물리적환경기준군으로 명명하였고, 평가요소와 관련없는 척도의 선정은 관찰 대상에서 제외하였다. 또한 적합성 평가시 평가기준의 단서로 활용되는 중요성을 고려하여 문헌요소를 관찰 대상에 포함하였다. 본 연구에서 사용하기 위한 평가기준을 도출하기 위하여 위에서 제시한 4개 평가기준군에 대한 관련연구를 분석한 바, 이를 요약하여 정리하면 다음의 <표 1>과 같다.

<표 1>을 분석정리해 보면, 문헌요소군에서는 12개의 요소가 확인되었으며 이 중에서 표제, 초록, 색인어, 저자명 등은 많은 실험에서 적합문헌을 적절하게 지시할 수 있는 문헌요소로 확인되고 있다. 또한 정보내용기준군에서 14개의 요소가 확인되었으며, 이중에서 주제적 적합, 범위와 깊이, 최신성등의 요소가 많은 관련연구에서 도출되었다. 이용자특성기준군에서는 9개의 요소가 확인되었으며 정보소스의 신뢰성, 명성, 경험및 배경 등이 대체로 많은 실험에서 도출되었다. 물리적 환경기준군에서는 대략 9개의 요소가 확인되었다. 이 중에서 입수가능성, 지리적 근접성, 소요비용 등이 다수의 실험에서 확인되고 있다. 그러나 이러한 기준요소들은 이전의 적합성 평가에서 다수 언급된 평가기준으로만 인식해야지 언급된 평가기준의 많고 적음이 실제 적합성 판단실험에도 그대로 적용될지는 의문이다. 왜냐하면, 이용자의 상황적 요인과 특성들이 적합성 평가시점과 환경에 따라 많은 변화를 보이기 때문이다. 따라서 기존의 실험연구에서 도출된 평가기준을 실제 적합성 실험에 활용하기 위해서는 평가기준의 신중한 선정이 요망된다.

<표 1> 적합성 평가관련 기준의 종합분석표

실험연구자		Saracevic	Marcus	Janes	Burgin	Park	Barry	Wang	Schamber	Cool	Bruce	Cooper	Taylor
평가기준		1969	1978	1991	1992	1993	1994	1994	1991	1993	1994	1991	1996
문헌요소군	표제명	○	○	○	○	○	○	○					
	초록	○	○	○	○	○	○	○					
	색인어		○	○	○	○	○	○					
	국가명							○					
	저자명			○		○	○	○					
	소속기관					○	○	○					
	출판사명						○	○					
	잡지명					○	○	○					
	출판일자			○		○	○	○					
	문헌유형			○		○	○	○					
정보내용기준군	문헌길이						○	○					
	권/호							○					
	사용언어							○					
	주제적적합					○	○	○	○	○	○		
	깊이/범위					○	○	○	○	○	○		
	타당정확성						○	○	○	○	○		
	최신성					○	○	○	○	○	○		
	명확성						○		○		○		
	현실성						○	○	○				
	해석성						○						
이용자특성기준군	새로움					○	○	○			○		
	외적타당성						○						
	유용성									○			
	유추성					○					○		
	역동성								○				
	효율성						○						
	완전성										○		
	정보소스신뢰성					○	○	○	○	○			
	정보소스명성					○	○	○	○	○			
	정보소스유형					○			○	○			
물리적환경기준군	감정/정서						○	○		○			
	경험/배경					○	○	○	○				
	이해능력						○						
	활용가능성						○		○	○			
	주관적타당						○						
	저자 관계						○	○					
	입수가능성					○	○	○	○				○
	지리근접성					○	○		○			○	○
	소요시간							○				○	
	소요비용							○	○	○		○	
연구제약기준군	연구제약						○					○	
	문헌포괄성												○
	소스연결성												○
	응답속도												○
예상결과물					○								

## 2. 평가기준의 설정

본 연구에서 제시한 가설을 검증하는데 사용한 평가기준들은 다음과 같이 설정되었다. 먼저 관련연구의 분석을 통하여 밝혀진 4개 평가기준 군에 속하는 45개 평가요소들을 일차적으로 선정하였다. 그러나 기존의 평가실험에서 밝혀진 평가요소들이 우리나라 정보검색 환경에서도 효용성이 있을지는 단정하기 매우 어렵다. 따라서 보다 합리적이고 타당성 있는 평가기준을 선정하기 위하여 실험에서 도출된 평가요소는 연구기관의 실제 이용자와 주제전문가와 면담을 통하여 새롭게 재구성하였다. 면담의 과정은 먼저 기존 평가실험에서 도출된 4개 기준 군에 속하는 45개 평가요소를 조작적 정의와 함께 그들에게 제시한 후, 평가요소가 실제 검색문헌의 적합성 평가에 활용되는 정도를 구술하게 하였다. 대략적인 구술내용과 선정의 이유는 다음의 개별 항목 속에 기술하였다.

### 1) 문헌요소의 설정

문헌요소는 적합성 평가 시 최초의 관심을 유발시키는 요소로서 이용자 평가 기준의 단서로 사용된다. 따라서 이용자의 적합성 평가에 상당한 영향을 미치고 있는 요소임에는 분명하다. 그러나 문헌요소들이 검색시스템에 어느 정도 그리고 어떤 문헌요소가 구비되어야만 이용자의 적합성 평가행위에 도움을 줄 수 있는가에 대해서는 신중하게 고려해야 한다. 예비적인 면담을 통해서도 문헌요소의 과다가 적합성 평가에 영향을 미친다기 보다는 검색대상 DB를 구성하는 모든 문헌요소가 적합성 평가를 위한 단서로 활용될 수 있다고 면담자들의 다수는 밝히고 있다. 종래의 평가에서도 적합성 평가에 도움을 줄 수 있는 문헌요소의 적정성을 규정한 논의는 없었다. 따라서 본 연구는 탐색대상 데이터베이스에 수록된 요소 가운데 문헌의 내외적 속성 모두를 선정하였다. 다만 데이터제작년월, 데이터제작기관 등의 문헌요소와 관련없는 요소는 제외하였다.

### 2) 정보내용기준군의 설정

실질적으로 적합성 평가기준을 밝히는 연구의 대부분은 이용자가 적합성을 평가할 때 검색된 문헌의 어떤 내용적 기준으로 적합성을 평가하는가를 밝히려는 것이 핵심이다. 따라서 이들의 대부분은 정보내용기준에 초점을 맞추고 있다. 그러나 종래의 실험에서 밝혀진 정보내용기준군에 속하는 평가기준들을 본 연구의 실험에 활용하기 위해서는 신중히 고려해야 한다. 왜냐하면 개별 평가기준들이 실험에서 언급된 빈도의 많고 적음이 국내 이용자들의 적

합성 평가시에도 효용성이 그대로 적용될지 의문이기 때문이다. 따라서 본 연구에서는 기존의 실험에서 밝혀진 정보내용기준군에 속하는 14개 평가요소를 예비적인 면담을 통해 검증한 결과, 그들은 주제적 적합, 내용의 깊이와 범위, 객관적 타당성/정확성, 최신성, 명확성, 해석성, 새로움 등의 기준이 다수의 면담자에 의하여 거론되었다. 그러나, 기존 실험에서 언급된 빈도가 높았던 현실성, 외적정보소스에 의한 타당성, 유추성 등은 전혀 거론되지 않았다. 특이한 점은 이들 기준보다 언급빈도가 낮았던 이해성은 그 반대의 현상을 보였다는 점이다.

### 3) 이용자특성기준군의 설정

적합성 평가 행위에서 가장 중요하고 기본적인 요소는 정보요구를 가지고 적합 문헌을 탐색하는 이용자이다. 이와 관련한 연구의 대부분은 이용자의 정보요구에 대한 인식 및 사상에 기초하여 정보를 평가하는 이른바 인지적 측면의 역할과 평가기준의 영향요인으로서 중요성을 밝히는 연구에 초점을 맞추고 있다. 따라서 이용자의 인식, 과거경험과 교육적 배경, 감정, 연구단계 등은 오로지 적합성 평가의 영향요인으로만 고려하였고 어떤 이유로 적합성을 평가했는가 하는 평가기준으로는 고려하지 않았다. 본 연구는 적합성을 평가하는 과정에서 개별 이용자가 독창적으로 인식하는 요소들을 평가기준으로 보고 이의 도출을 시도하였다. 기존의 연구에서 밝혀진 이용자특성기준군에 속하는 9개의 기준을 예비적인 면담조사에서 제시한 후 이의 타당성을 평가하게 하였다. 그들은 9개의 요소중에서 정보소스의 신뢰성인식, 정보소스의 명성 인식, 정보소스의 유형에 대한 인식, 개인적인 감정/취향, 과거경험이나 연구배경, 정보소스의 이해능력, 개인적 활용가능성 기준이 면담자에 의하여 거론되었다. 이 중에서 기존 실험에 언급된 빈도가 매우 적은 정보소스의 이해능력은 거론된 반면 정보소스의 이해능력보다 언급빈도가 높은 저자와의 관계는 전혀 거론되지 않았다. 이는 대체로 국내 연구자들이 국외연구자들보다 국외 저명학술지에 기고하는 빈도가 낮은 학술환경에 기인한다고 볼 수 있다.

### 4) 물리적환경기준군의 설정

적합성 평가 기준과 관련한 연구에서 발견되는 물리적환경기준들은 보통 정보내용기준, 이용자특성기준들과 함께 통합적으로 언급되고 있는 것이 일반적이다. 기존 연구에서 도출된 물리적환경기준군에 속하는 평가기준은 9개였다. 본 연구의 실험에 사용한 물리적환경기준을 설정하기 위하여 9개의 평가기준을 면담자들에게 제시한 후 이의 타당성을 평가하게 하였다. 그들은 9개의 요소 가운데서 문헌의 입수가능성, 입수 소요시간, 소요비용, 연구시간압박감등 네개의 기준들을 거론하였고, 그 나머지 기준들은 거론하지 않았다. 기존 연구에서 많이 언

급된 지리적 근접성 기준은 거론하지 않은 반면, 이보다 언급빈도가 적었던 연구시간압박감과 입수소요시간은 면담자에 의하여 거론된 점이 특이하다.

이상과 같이 본 연구의 실험에 사용하기 위하여 최종적으로 설정된 평가기준을 종합하여 정리하면 다음 <표 2>와 같다.

<표 2> 선정된 평가기준의 종합표

평가기준군	평가 기준
문헌 요소군	표제명
	저자명
	색인어
	초록
	잡지명
	출판연도
	자료구분
	문헌길이
	주제분류
	사용언어
	기사구분
정보 내용 기준군	정보내용의 주제적 적합
	정보내용의 깊이와 범위
	정보내용의 객관적 타당성 및 정확성
	정보내용의 최신성
	정보내용의 명확성
	정보내용의 해석성
	정보내용의 참신성
이용자 특성 기준군	정보소스의 신뢰성
	정보소스의 명성
	정보소스의 유형
	개인적인 취향/감정/정서
	이용자의 과거경험이나 배경
	내용의 이해능력
물리적 환경 기준군	개인적인 활용가능성
	문헌의 입수가능성
	문헌의 입수에 소요되는 시간
	문헌의 입수에 소요되는 비용
	연구에 소요되는 시간압박감

## IV. 적합성 평가기준 변화에 대한 실험조사

본 장에서는 이용자가 연구과정에서 인식하는 적합성 평가기준의 변화를 규명하기 위하여 연구가설들을 설정하고, 연구환경에서 실제 이용자들의 적합성 판단기준들을 도출하고자 하였다. 이를 위하여 이용자들이 직접 검색한 문헌들을 대상으로 연구과정에 따라 적합여부 평가할 때 사용하는 기준들을 관찰하는 실험연구를 수행하였다.

### 1. 가설의 설정

본 연구는 검색문헌의 적합성 평가기준에 대한 이용자의 인식 변화를 규명하기 위하여, 연구과정에 따른 문헌의 적합여부와 평가기준들 사이의 상관관계를 식별하고, 적합여부에 영향을 미치는 평가기준들의 인식도 차이를 파악하기 위하여 필요한 연구가설들을 다음과 같이 설정하였다.

본 연구의 첫째 목적인 연구과정에 따른 문헌의 적합여부와 평가기준들 사이의 상관관계를 식별하기 위하여 연구가설 1을 설정한다.

가설1 : 연구과정에 따라 문헌의 적합여부와 평가기준들사이에는 유의적인 상관관계를 보일 것이다.

본 연구의 둘째 목적인 연구과정에 따른 문헌의 적합여부에 영향을 미치는 평가기준들의 인식도 차이를 검증하기 위하여 다음과 같은 연구가설 2, 3을 설정한다.

가설2 : 검색문헌의 적합여부평가에 대한 이용자들의 인식도는 연구과정에 따라 차이가 있을 것이다.

가설3 : 검색문헌의 평가기준에 대한 이용자들의 인식도는 연구과정에 따라 차이가 있을 것이다.

### 2. 실험조사방법

본 연구의 핵심내용은 실제 정보문제를 가지고 있는 이용자들이 연구과정에 따라 검색문헌의 적합성을 판단할때 사용하는 평가기준의 변화를 관찰하는 실험에 있다. 이러한 실험에는 다음과 같은 전제조건이 충족되어야 한다. 첫째, 실제 연구문제를 가지고 있는 이용자를

대상으로 해야 한다. 둘째, 이용자들은 문헌조사 연구방법을 수행해야 한다. 셋째, 이용자들은 온라인을 통하여 적합문헌을 탐색하고, 연구의 시작부터 완료되기까지 탐색을 지속적으로 수행해야 한다. 넷째, 실험은 장기적으로 수행되어야 하므로 관찰에 의한 실험을 이용자가 허용해야 한다. 다섯째, 관찰은 연구과정에 따라 지속적으로 이용자들의 적합성 평가기준 인식변화를 살필 수 있어야 한다. 이와 같은 실험성격은 다수의 이용자를 대상으로 실험을 수행하기 어렵다는 문제점이 제기된다. 따라서 실험조건을 충족시키고 실험에 참여할 의사가 있는 이용자로 한정하여 실험을 수행할 수밖에 없다. 이러한 이유로 인하여 본 연구에서는 실험에 자발적인 참여 의사를 밝힌 5명을 대상으로 그들의 연구과정에 따른 적합성 평가기준을 조사하였다. 그러나 1명은 중도에 포기하였고, 결과분석에는 4명의 평가결과만 활용되었다. 조사 대상자들의 연구과정에 따른 적합성 평가기준을 밝히기 위하여 구체적으로 수행된 실험조사방법은 다음과 같다. 먼저 그들에게 본 연구에 대한 대략적인 목적, 그리고 기존 연구와 주제전문가의 면담을 통해 추출된 적합성 평가기준, 체크 방법 등의 사전 설명을 실시하였다. 둘째, 정보요구를 표현하고 있는 질의어를 직접 작성하게 하고 이를 산업기술정보원의 데이터베이스 사업부에서 제작하고 있는 국내 최대의 과학기술 문헌정보 데이터베이스인 BIST를 통하여 검색하도록 하였다. 셋째, 검색된 문헌의 상위 50개 문헌만을 대상으로 적합여부와 적합여부평가를 위하여 참고하는 문헌요소항목 및 인식한 적합성 평가기준을 별도로 배부한 과정별 평가기준 체크 리스트를 통하여 모두 표시하도록 요구하였다. 적합의 여부는 적합/부분적합/부적합의 3등급의 척도를 사용하였다. 그 이유는 적합여부평가를 위하여 선정되는 평가기준들 사이에는 특징적 차이가 있을 것으로 예상하고, 검색문헌을 적합으로 평가하는 기준, 부분적합으로 평가하는 기준, 부적합으로 평가하는 기준을 확인하고 연구과정별 변화를 관찰하기 위하여 채택되었다. 평가기준 체크 리스트는 총 4페이지로 구성되었다. 첫 페이지는 평가기준에 대한 조작적 정의와 기입방법을 개설하고, 나머지 페이지는 연구진행에 따른 연구과정별 체크리스트 3장이 동일한 포맷으로 개별적으로 구성되어 있다. 각 과정별 체크 리스트에는 검색된 문헌번호를 의미하는 동일한 문헌번호, 적합/부분적합/부적합의 적합여부, 평가문헌요소, 정보내용기준군, 이용자특성기준군, 물리적환경기준군에 속하는 평가기준등이 표시될 수 있도록 하는 공간을 마련하였다. 적합성 평가는 최초에 검색된 동일한 문헌을 대상으로 연구과정에 따라 시작, 진행, 완성의 3개 과정으로 나누어 실시하고, 각 과정사이는 대략 2개월의 간격을 두고 평가자 자신이 편리한 시기에 수행하도록 하였다. 여기에서 시작과정은 연구의 초기과정으로서 정보의 입수 필요성을 자각하고 정보요구에 대한 관련 정보를 검색하고 수집하는 과정으로 정의하였다. 진행과정은 관련된 수집 정보를 정리하고 문제영역에 초점을 맞추는 과정으로서 연구보고서를 집필하는 초기과정으로 정의하였다. 완성과정은 연구보고서를 집필하는 본격적인 과정으로서 연구보고서를 완성하고 인용문헌을 결정하는 과정으로 정의하였다.

### 3. 실험 및 결과분석

#### 1) 실험개요

본 연구의 평가실험에 마지막까지 참여한 4명의 이용자에 대한 구체적인 연구문제, 검색질문, 검색한 문헌건수는 <표 3>과 같다.

<표 3> 평가대상자들의 검색실험개요

평가자	연구문제	질문식	검색건수	평가건수
평가자1	건설분야 멀티미디어 활용	(건설 토목 건축) and (멀티미디어 컴퓨터)	313	50
평가자2	멀티미디어 데이터베이스의 콘텐츠 개발에 관한 연구	멀티미디어 and (DB 데이터베이스)	225	50
평가자3	한국교회건축의 사적 변천	교회 and 건축	42	42
평가자4	전자저널의 수집 연구	전자잡지 and (수서 수집)	70	50

본 연구에서는 평가자들이 검색한 모든 문헌을 실험대상으로 포함시키지 않고, 초록을 포함하고 있는 검색된 문헌이 출현하는 순서상의 상위 50개 문헌만을 평가대상으로 삼았다. 그 이유는 검색건수의 차이와 검색건수 차이에서 오는 평가의 불합리성 때문이다. 평가자 1, 2는 검색건수가 각각 300건, 200을 초과하는 반면에 평가자 3은 검색건수가 42건 밖에 나오지 않는다. 따라서 평가건수 차이에서는 오는 대상자간 평가의 불합리성을 최대한 줄이기 위하여 최저의 평가건수에 가까운 50건을 평가대상 문헌으로 삼았다. 그러므로 평가자 1, 2, 4는 50건이 평가대상이 되며, 평가자 3은 42건이 평가대상이 된다. 선정된 이들 검색문헌을 대상으로 연구과정을 3개 과정로 나누어 각각의 문헌에 대한 개별적인 적합성 평가를 독립적인 환경에서 수행하고 평가기준을 체크하도록 요구하였다.

#### 2) 적합여부 평가결과의 분포현황

본 연구는 가설을 검증하기에 앞서 본 평가실험에서 밝혀진 연구과정에 따른 적합여부 평가결과를 빈도분석을 통하여 다음 <표 4>와 같이 외향적인 변화를 관찰하였다.



&lt;표 4&gt; 연구과정에 따른 적합여부

평가자	시작과정						진행과정						완성과정					
	적합		부분적합		부적합		적합		부분적합		부적합		적합		부분적합		부적합	
	건	%	건	%	건	%	건	%	건	%	건	%	건	%	건	%	건	%
평가자1	35	70.0	9	18.0	6	12.0	23	46.0	13	26.0	14	28.0	19	38.0	4	8.0	27	54.0
평가자2	26	52.0	13	26.0	11	22.0	21	42.0	12	24.0	17	34.0	16	32.0	12	24.0	22	44.0
평가자3	16	38.0	13	31.0	13	31.0	16	38.1	7	16.7	19	45.2	13	31.0	9	21.4	20	47.6
평가자4	27	54.0	17	34.0	6	12.0	26	52.0	16	32.0	8	16.0	21	42.0	20	40.0	9	18.0
전체	104	54.2	52	27.1	36	18.7	86	44.8	48	25.0	58	30.2	69	35.9	45	23.4	78	40.7

상기의 <표 4>에서 처럼 연구과정에 따른 적합여부의 평가는 상당한 차이가 있는 것으로 나타나고 있다. 연구가 진행됨에 따라 점차로 그 적합도는 감소하고 있으며, 부적합도는 증가하는 것으로 나타났다. 이러한 현상은 연구과정에 따라 다양한 연구지식, 정보이용과 관련 있는 다양한 외적, 상황적인 요인이 적합판단에 상당한 영향을 미치고 있는 것으로 생각된다. 즉, 연구영역의 지식이 없는 초기에는 정보요구를 표현한 질문이나 질의어와 검색된 문헌요소에서 나타나는 용어와의 일치의 관점에서 적합문헌을 평가하는 경향이 있는 반면에, 연구가 어느정도 진행되면 용어상의 일치보다는 이용자의 정보요구와 검색문헌의 관련성을 고려하여 적합문헌을 평가하고 있는 것으로 짐작된다. 또한 평가자 개인들도 대체로 이러한 현상을 보이고 있는데, 평가자의 개인적인 판단환경이 달라도 적합성의 평가는 대체로 일치하고 있음을 알 수 있다. 한편, 부분적합으로 평가되는 문헌은 연구과정에 따라 약간 감소하는 경향을 보이고 있으나, 이는 큰 의미는 없다. 왜냐하면, 적합성 평가에서 부분적합은 적합으로 판단될 수도 있고 부적합으로 판단될 수 있는 개연성이 있기 때문이다.

### 3) 적합성 평가기준의 분포현황

#### (1) 문헌요소의 분포현황

연구과정에 따라 적합여부를 평가하기 위한 최초의 단서로서 미리 설정된 11개의 문헌요소 가운데 어떠한 문헌요소를 선정하였는지에 대한 분포현황을 연구과정별로 빈도분석을 통하여 살펴보았는데, 그 결과는 다음 <표 5>와 같다.

&lt;표 5&gt; 연구과정에 따른 문헌요소 선정분포표

문헌요소	시작과정				진행과정				완성과정			
	적합	부분적합	부적합	전체	적합	부분적합	부적합	전체	적합	부분적합	부적합	전체
표제	90	26	27	143	44	13	12	69	26	5	19	50
저자명	6	1	0	7	4	4	0	8	4	2	3	9
색인어	27	26	3	56	18	16	4	38	5	5	2	12
초록	13	11	10	34	33	16	28	77	37	22	28	87
잡지명	0	7	3	10	10	10	4	24	6	5	13	24
출판일자	3	2	0	5	3	2	6	11	9	2	4	15
자료구분	1	2	0	3	3	4	0	7	2	1	1	4
문헌길이	1	0	0	1	1	0	4	5	5	8	8	21
주제분류	17	8	8	33	12	2	7	21	5	2	7	14
사용언어	1	1	1	3	0	6	19	25	4	5	31	40
기사구분	1	3	0	4	2	0	2	4	0	0	2	2

위의 <표 5>에서 표제, 색인어, 주제분류 등은 점점 그 참고 빈도가 감소하는 문헌요소인 반면, 잡지명, 출판일자, 문헌길이, 사용언어 등은 그 참고빈도가 증가하는 문헌요소임을 알 수 있다. 이러한 현상은 질문과 검색문헌간 용어의 일치를 고려하는 연구초기에는 표제, 색인어, 주제분류 등을 주로 참고하거나 연구가 진행되면서 점점 용어의 일치는 감소하는 동시에 다양한 외적인 즉, 잡지명 인식, 가독능력, 연구시간제약 등이 적합성 판단에 영향을 미치는 것으로 생각된다. 또한 선정빈도가 10개 이상인 문헌요소를 비교했을 때 연구의 초기과정에는 표제, 색인어 초록, 주제분류 등 몇몇 문헌요소만이 적합여부의 주요 단서로 활용되고 있지만, 완성과정에서는 초록, 표제, 사용언어, 잡지명 등 다양한 문헌요소를 활용하여 적합여부를 평가하고 있음을 알 수 있다. 이러한 현상은 연구가 진행되면서 적합성 판단에는 문헌의 이용과 관련된 많은 요인들이 적합성 판단에 영향을 미치며 그에 따라 참고하는 문헌요소도 다양해지는 것으로 짐작된다.

## (2) 적합여부 평가기준의 분포현황

적합성 평가기준 분포현황을 연구의 시작, 진행, 완성순으로 분석해본 결과는 <표 6>과 같다. 아래 <표 6>에서 처럼 전반적으로 정보내용기준군들이 연구과정과 관계없이 적합여부평가시 가장 많이 언급되는 기준군으로 확인되고 있다. 그러나, '주제적 적합', '내용의 새로움', '개인의 감정/정서', '과거연구경험/배경' 등은 연구가 진행됨에 따라 사용빈도가 점점 감소하고 있다. 반면에 내용의 명확성, 해석성, 정보소스의 신뢰성, 유형, 내용의 이해능력, 입수가능성, 입수소요시간, 비용성, 연구시간압박감 등의 평가기준들은 상대적으로 증가하고 있는 것으로 나타났다.

&lt;표 6&gt; 연구과정에 따른 적합평가기준의 분포표

적합평가기준		시작과정				진행과정				완성과정			
		적합	부분적합	부적합	전체	적합	부분적합	부적합	전체	적합	부분적합	부적합	전체
정보내용기준	주제적 적합	55	17	14	86	17	10	2	29	10	6	7	23
	내용범위/깊이	34	23	21	78	42	10	23	75	30	9	24	63
	객관적 타당성	16	1	3	20	8	10	7	25	12	0	7	19
	내용의 최신성	8	4	2	14	3	4	8	15	9	2	4	15
	내용의 명확성	6	5	5	16	26	6	10	42	18	13	12	43
	내용의 해석성	0	1	4	5	0	5	11	16	4	3	18	25
이용자특성기준	내용의 새로움	17	2	1	20	14	2	3	19	1	1	3	5
	정보소스 신뢰성	0	7	0	7	6	5	0	11	6	3	3	12
	정보소스의 명성	5	0	1	6	1	6	2	9	1	1	7	9
	정보소스의 유형	2	2	1	5	3	5	1	9	6	3	5	14
	개인감정/정서	5	2	2	9	3	1	1	5	3	0	0	3
	과거경험/배경	16	3	0	19	7	0	1	8	2	0	0	2
	내용 이해능력	2	6	0	8	2	4	11	17	2	3	12	17
개인적 활용성	3	7	0	10	3	0	0	3	5	1	0	6	
물리적기준	입수가능성	1	1	3	5	3	2	7	12	9	6	7	2
	입수소요시간	0	1	0	1	1	0	0	1	2	1	5	8
	비용성	0	0	0	0	1	0	1	2	1	2	2	5
	연구시간압박감	1	0	0	1	3	0	6	9	4	6	16	26

이러한 현상은 연구분야 지식이 거의 없는 연구의 초기에는 과거경험이나 배경등에 의존하여 주관적으로 적합성을 평가하거나 용어상의 일치를 고려하는 주제적 적합으로 적합성을 평가하는 경향이 강한 것으로 짐작된다. 반면에 연구가 진행되면 이용자는 보다 명확한 내용을 원하거나, 이해할 수 있고, 입수가능한 문헌을 원하는 경향이 강한 것으로 판단된다.

## V. 실험결과에 따른 가설 검증

본 장에서는 제 4장의 연구환경에서 실제 사용자들의 적합성 판단기준들을 밝히기 위한 실험을 통하여 도출된 결과를 토대로 본 연구에서 설정된 연구가설들을 검증하고자 하였다. 가설의 검증을 위하여 상관관계분석, 일원분산분석이 활용되었다.

### 1. 적합여부와 평가기준과의 상관관계분석

연구과정에 따라 검색문헌의 적합여부, 즉 적합으로 평가할 때, 부분적합으로 평가할 때, 부적합으로 평가할 때와 평가기준사이에는 유의적인 상관관계를 보일 것이라는 가설을 검증하기 위하여 상관관계 분석을 <표 7>과 같이 실시하였다.

이중에서 문헌요소는 적합성 평가를 유발시키는 단서이자 적합성 평가의 참고요소로서 직접적으로 적합여부의 평가와는 관련이 없다고 판단되었기 때문에 상관관계분석에서 제외하였다. 상관관계분석에서 나타나는 상관계수의 값은 음(-)의 값은 적으면 적을수록 비교대상 요소의 상관관계는 그 값에 비례하여 적고, 정(+의 값이 크면 클수록 상관관계는 그 값에 비례하여 큼을 의미한다.

<표 7> 적합여부와 평가기준의 상관관계

평가기준	적합여부평가		
	적합	부분적합	부적합
내용의 주제적 적합	.109(0.009)**	.008(0.844)	-.127(0.002)
내용의 범위/깊이	.053(0.202)	-.108(0.009)	.046(0.275)
내용의 객관적인타당성	.091(0.029)*	-.049(0.239)	-.052(0.212)
내용의 최신성	.010(0.818)	-.008(0.847)	-.003(0.947)
내용의 명확성	.033(0.429)	-.025(0.547)	-.012(0.775)
내용의 해석성	-.201(0.000)	-.022(0.592)	.242(0.000)***
내용의 새로움	.172(0.000)***	-.100(0.017)	-.092(0.028)
정보소스의 신뢰성	-.015(0.723)	.124(0.003)**	-.103(0.013)
정보소스의 명성	-.056(0.181)	.011(0.784)	.050(0.231)
정보소스의 유형	-.042(0.311)	.034(0.421)	.014(0.737)
개인적인 감정/취향	.098(0.018)*	.037(0.376)	-.072(0.085)
이용자의 과거경험/배경	.212(0.000)***	-.084(0.044)	-.151(0.000)
내용의 이해능력	-.138(0.001)	.030(0.476)	.122(0.003)**
개인적 활용가능성	.069(0.098)	.039(0.356)	-.113(0.007)
입수가능성	-.069(0.099)	.012(0.782)	.064(0.124)
문헌입수 소요시간	-.024(0.561)	-.024(0.567)	.050(0.236)
문헌입수 소요비용	-.037(0.380)	.007(0.859)	.033(0.429)
연구시간압박감	-.119(0.004)	-.053(0.203)	.181(0.000)***

\*P<0.05 \*\*P<0.01 \*\*\*P<0.001

위의 <표 7>에서 처럼 전체연구과정에서 연구자들의 적합여부평가는 몇몇 개별평가기준과 상관관계가 있는 것으로 나타났다. 구체적으로 적합으로 평가와는 상관계수값  $\gamma$ 이  $-.201 \leq \gamma \leq .212$  사이에 분포하고 있었으며, 내용의 주제적 적합, 내용의 객관적인 타당성, 내용의 새로움, 개인적인 감정 및 정서, 이용자의 과거경험과 연구배경에서 유의적인 정(+의 상관관계가 있는 것으로 나타났다. 부분적합으로 평가와는 상관계수값  $\gamma$ 은  $-.108 \leq \gamma \leq .124$  사이에 분포하고 있으며, 정보소스의 신뢰성에서 유의적인 정(+의 상관관계가 있는 것으로 나타났다. 부적합으로 평가와는 상관계수값  $\gamma$ 은  $-.151 \leq \gamma \leq .242$  사이에 분포하고 있으며, 내용의

해석성, 내용의 이해능력, 연구시간 압박감에서 유의한 정(+)의 상관관계가 있는 것으로 확인되었다.

## 2. 적합여부평가의 인식검증

검색문헌의 적합여부평가에 대한 이용자들의 인식도는 연구과정에 따라 차이가 있을 것이라는 가설을 검증하기 위하여 이용자들의 인식 정도를 관찰하고 이에 따라 인식도를 검증하였다. 이는 적합여부 평가결과의 분포현황에서 정확하게 이용자들의 적합여부평가에 대한 인식정도를 관찰할 수 없었기 때문에 인식도를 검증하기에 앞서 인식의 정도를 먼저 살펴보았다. 아래 <표 8>에서 처럼 적합으로 평가되는 문헌의 경우는 시작과정에서는 104건으로 나타나 전체 평가문헌의 54.2%를 차지하고 있으나 연구과정이 진행될수록 적합으로 평가되는 문헌은 점차 감소하여 완성과정에 이르러서는 적합으로 평가되는 문헌은 69건으로 전체의 35.9%로 낮아지는 것으로 나타났다.

<표 8> 적합여부평가의 인식 정도

구 분	시작과정		진행과정		완성과정	
	빈도(건)	비율(%)	빈도(건)	비율(%)	빈도(건)	비율(%)
적 합	104	54.2	86	44.8	69	35.9
부분적합	52	27.1	48	25.0	45	23.4
부 적 합	36	18.7	58	30.2	78	40.7
전 체	192	100.0	192	100.0	192	100.0

이것은 연구를 시작하는 과정에서는 질문과 검색문헌사이의 적합성을 평가하는 적합성을 지향하는 경향이 이용자들 사이에서 나타나고 있다는 증거이다. 반면에 부적합으로 평가되는 문헌은 시작과정에서는 36건으로 전체의 18.7%를 차지하고 있으나 연구과정이 진행될수록 그 수는 증가하여 완성과정에 이르면 전체의 40.7%로 높아지는 것으로 나타났다. 이러한 현상은 연구과정이 높으면 높을수록 이용자의 정보요구와 검색문헌사이의 적합성을 평가하는 적절성을 지향하는 경향이 이용자들 사이에 나타나고 있는 것으로 해석할 수 있다. 이러한 결과데이터를 토대로 연구과정에 따른 적합여부평가의 인식도를 검증하였다.

다음 <표 9>에서 처럼 연구과정에 따른 적합여부의 인식도를 검증한 결과 적합과 부적합에서 인식에 유의한 차이가 있었다. 즉 적합은 연구의 시작과정에서 보다 높게 인식하고 있는 반면 부적합은 시작과정보다 연구의 진행, 완성과정에서 보다 높게 인식되는 것으로 나타났다.

<표 9> 적합여부평가의 인식 검증

적합여부	시작	진행	완성	F	P
적합	1.54	1.44	1.36	6.011	0.003**
부분적합	1.27	1.26	1.24	0.270	0.763
부적합	1.19	1.30	1.40	10.230	0.000***

\*P<0.05 \*\*P<0.01 \*\*\*P<0.001

### 3. 적합성 평가기준의 인식검증

연구자들이 연구과정에 따라 검색문헌의 적합성을 평가할 때 검색문헌의 평가기준에 대한 인식도는 차이가 있을 것이라는 가설을 분산분석을 통하여 검증하였다.

#### 1) 적합여부에 사용된 평가기준의 인식검증

##### (1) 문헌요소의 인식검증

다음 <표 10>에서 처럼 연구과정에 따라 이용자들이 검색문헌의 적합여부를 평가할 때 참고하는 문헌요소에 대한 인식차이를 검증하였다.

<표 10> 적합여부평가에 사용된 문헌요소 인식검증

문헌요소	시작과정	진행과정	완성과정	F	P
표제	1.74	1.37	1.26	59.490	0.000***
저자명	1.04	1.04	1.05	0.166	0.847
색인어	1.29	1.15	1.06	19.510	0.000***
초록	1.19	1.46	1.47	22.708	0.000***
잡지명	1.05	1.13	1.13	4.716	0.009**
출판일자	1.03	1.06	1.08	2.574	0.077
자료구분	1.02	1.03	1.02	0.311	0.733
문헌길이	1.01	1.04	1.11	11.979	0.000***
주제분류	1.17	1.11	1.07	4.847	0.008**
사용언어	1.02	1.12	1.21	16.686	0.000***
기사구분	1.03	1.03	1.02	0.389	0.678

\*P<0.05 \*\*P<0.01 \*\*\*P<0.001

그 결과 연구과정에 따라 적합여부를 평가할 때 표제, 색인어, 초록, 잡지명, 문헌길이, 주제분류, 사용언어에서 p값이 0.05보다 작게 나타나 인식에 유의적인 차이가 있는 것으로 나타났다. 즉 표제, 색인어, 주제분류 등은 적합여부를 평가할 때 진행, 완성과정보다 시작과정에서 더 많이 인식하고 있었다. 반면에 초록, 잡지명, 문헌길이, 사용언어등은 시작과정보다

진행, 완성과정에서 더 높게 인식하는 것으로 나타났다.

## (2) 평가기준의 인식

연구과정에 따라 문헌의 적합여부를 평가할 때 사용하는 평가기준에 대한 이용자들의 유의적인 인식차이를 검증하였다<표 11>.

<표 11> 적합여부평가에 사용된 평가기준의 인식검증

평가기준	시작	진행	완성	F	P
내용의 주제적 적합	142	1.16	1.13	29.798	0.000***
내용의 깊이와 범위	1.41	1.41	1.33	1.843	0.159
내용의 객관적 타당성	1.12	1.14	1.09	0.894	0.410
내용의 최신성	1.07	1.07	1.09	0.194	0.824
내용의 명확성	1.09	1.23	1.22	7.325	0.001**
내용의 해석성	1.03	1.08	1.14	7.444	0.001**
내용의 새로움	1.12	1.10	1.03	6.009	0.003**
정보소스의 신뢰성	1.04	1.06	1.06	0.860	0.424
정보소스의 명성	1.04	1.05	1.05	0.539	0.583
정보소스의 유형	1.03	1.05	1.06	1.040	0.354
개인적인 취향	1.05	1.03	1.02	2.241	0.107
이용자 과거경험/배경	1.10	1.04	1.01	9.092	0.000***
내용의 이해능력	1.04	1.08	1.09	2.159	0.116
개인적인 활용가능성	1.05	1.02	1.03	1.677	0.188
문헌입수가능성	1.03	1.06	1.13	7.648	0.001**
문헌입수 소요시간	1.01	1.01	1.05	5.946	0.003**
문헌입수비용	1.00	1.01	1.03	2.730	0.066
연구시간압박감	1.01	1.05	1.14	14.994	0.000***

\*P<0.05 \*\*P<0.01 \*\*\*P<0.001

그 결과 위의 <표 11>에서 보는 바와 같이 연구과정에 따라 적합여부를 평가할 때 내용의 주제적 적합, 내용의 명확성, 내용의 해석성, 내용의 새로움, 이용자의 과거경험 및 배경, 입수가능성, 문헌입수 소요시간, 연구시간압박감 등에서 유의적인 인식 차이를 보였다. 즉, 내용의 주제적 적합, 내용의 새로움, 이용자의 과거경험 및 배경 등은 연구과정에 따라 적합여부를 평가할 때 시작과정에서 더 많이 인식하고 있는 반면, 내용의 명확성은 진행과정에서, 그리고 내용의 해석성, 입수가능성, 문헌입수 소요시간, 연구시간압박감 등은 완성과정에서 더 많이 인식하고 있었다.

## V. 결 론

본 연구는 연구과정에 따른 이용자 적합평가기준의 변화와 개별 평가기준에 대한 인식 차이를 검증하기 위하여 실제 연구환경에서 이용자의 검색문헌에 대한 적합성 평가 기준의 변화를 관찰 연구하였다. 적합성 판단기준 변화를 규명하기 위한 실험에는 문헌요소군, 정보내용기준군, 이용자특성기준군, 물리적 환경기준군에 속하는 29개 평가기준을 도출하여 사용하였다.

본 연구의 결과를 요약하면 다음과 같다.

첫째, 연구과정에 따른 문헌의 적합여부에 영향을 미치는 평가기준들을 식별하기 위하여, 적합여부결과와 평가기준들의 상관관계분석을 통해 가설1을 다음과 같이 검증하였다.

연구의 전과정을 통하여 ① 적합으로 평가되는 내용의 주제적 적합, 내용의 객관적인 타당성, 내용의 새로움, 개인적인 감정 및 정서, 이용자의 과거경험과 연구배경에서, ② 부분적합으로 평가되는 정보소스의 신뢰성에서, ③ 부적합으로 평가되는 내용의 해석성, 내용의 이해능력, 연구시간 압박감에서 유의한 정(+의 상관관계가 확인되었다.

둘째, 연구과정에 따른 검색문헌의 적합여부에 영향을 미치는 평가기준들의 인식도 차이를 검증하기 위하여 가설2에서 3을 분석한 결과 다음과 같은 사실이 밝혀졌다.

1. 연구과정에 따른 적합여부평가에 대한 인식검증에서 적합과 부적합에서 인식에 유의한 차이가 있었다. 즉 적합은 연구의 시작과정에서 보다 높게 인식하고 있는 반면 부적합은 시작과정보다 연구의 진행, 완성과정에서 보다 높게 인식되는 것으로 나타났다.

2. 연구과정에 따른 검색문헌의 평가기준에 대한 인식검증중에서 ① 문헌요소에서는 표제, 색인어, 초록, 잡지명, 문헌길이, 주제분류, 사용언어에서 인식차이가 나타났으며, ② 평가기준에서는 내용의 주제적 적합, 내용의 명확성, 내용의 해석성, 내용의 새로움, 이용자의 과거경험 및 배경, 입수가능성, 문헌입수 소요시간, 연구시간압박감 등에서 인식차이가 나타났다.

이상의 연구결과에서 보는 바와 같이 이용자의 연구과정을 통하여 문헌요소와 평가기준의 판단행위가 역동적으로 변화되고 있다는 사실이 규명되었다. 이것은 이용자의 지식과 상황적 요인이 변화함에 따라 적합성 평가에 대한 인식이 수정될수 있음을 의미하며, 이러한 이유로 인하여 이용자의 적합성 판단행위가 변화되는 것으로 해석할 수 있다. 그러나 이용자의 적합성 판단행위를 변화시키는 요인들을 정보검색시스템이 통제하는 것은 어려운 일이다. 다만 이용자 중심의 정보탐색환경에서 검색시스템들은 다양한 연구상황에서 이용하고 수정할 수



있도록 이전의 검색에 활용한 탐색전략을 저장하고 활용할 수 있는 기능을 마련해야 한다. 또한 인터넷을 통한 웹의 접근이나 온라인 망을 통한 정보접근시 사용언어, 출판일자, 입수 가능성, 입수비용 등의 요소를 검색할 수 있는 시스템 설계가 필요하다. 온라인탐색서비스 환경에서도 탐색중개자들은 기존의 적합문헌 평가요소이외에 이용자의 연구상황에 따른 다양한 상황적 요인들을 정보검색과 서비스에 활용함으로써 온라인중개서비스의 효율성을 높여야 할 것이다.

### 참 고 문 헌

- 유재욱. "적합 판단 영향 요인에 관한 이론적 고찰", 《정보관리학회지》 제13권, 2호(1996). pp. 143-163.
- Barry, C. L. "User-defined relevance criteria: an exploratory study." *Journal of the American Society for Information Science*, vol.45, no.3(1994). pp. 149-159.
- Belkin, N. "The Problem of matching in information retrieval." *Proceedings of the Second International Research Forum on Information Science*, 1980. pp. 187-197.
- Burgin, R. "Variations in relevance judgments and the evaluation of retrieval Performance." *Information Processing & Management*, vol.28, no.5(1992). pp. 619-627.
- Cooper, W. "A Definition of relevance for information science." *Information Storage and Retrieval*, vol.7, no.1(1971). pp. 19-37.
- Cuadra C. and Katter, R. "Experimental studies of relevance judgments." NSF Final reports, Vol. I, Santa Monica: System Deveopment Corporation, 1967.
- Cuadra, C., Katter, R. "Opening the black box of relevance." *Journal of Documentation*, vol.23, no.3(1967). pp. 291-303.
- Eisenberg M., Schamber, L. "Relevance: the search for a definition." *ASIS Mid-Year Proceedings*, 1988. pp. 164-168.
- Foskett, D. "A Note on the concept of relevance." *Information Storage and Retrieval*, vol.8, no.1(1972). pp. 77-78.
- Harter, S. P. "Psychological relevance and inforamtion science." *Journal of the American Society for Information Science*, vol.43, no.14(1992), pp. 602-615.

- Rees, A., Schultz, D. *A Field experimental approach to the study of relevance assessments in relation to document searching*, 2vols (1967). NSF report.
- Saracevic, T. *The Concept of relevance in information science: a historical view*, New York: Bowker, 1970.
- Saracevic, T. "Relevance: a review of and a framework for the thinking on the notion in information science." *Journal of the American Society for Information Science*, vol.26, no.6(1975). pp. 321-343.
- Schamber, L., Eisenberg, M. B. and Nilan, M. S. "A Re-examination of relevance: toward a dynamic, situational definition." *Information Processing & Management*, vol.26, no.6(1990). pp. 755-776.
- Schamber, L. "Relevance and information behavior." *Annual Review for Information Science and Technology*, vol.29(1994). pp. 3-48.
- Swanson, D. "Subjective versus objective relevance in bibliographic retrieval systems." *Library Quarterly*, vol.56, no.4(1986). pp. 389-398.
- Vickery, B. "Subject analysis for information retrieval." *Proceedings of the International Conference on Scientific Information*, vol.2(1959). pp. 855-865.
- Wilson, P. "Situational relevance." *Information Storage and Retrieval*, vol.9, no.9(1973). pp. 457-471.