

[ISBN 978-89-6211-265-8]

# OntoFrame과 Google-Scholar 사용자 테스트를 통한 기능적 효용성 평가

김도완(배재대)

정한민(KISTI)

이미경(KISTI)

성원경(KISTI)

한선화(KISTI)



한국과학기술정보연구원

## 목 차

제 1장 Software 품질 및 기능적 효용성에 대한 고찰

제 2장 OntoFrame과 Scholar.Google 기능적 효용성 테스트 프로세스

2.1 테스트 목표, 테스트 범위 정의

2.2 테스트 사용자 선정 및 사전 인터뷰

2.3 사전 인터뷰 분석

2.4 사용자 테스트 시나리오 (Test Task) 및 테스트 후 인터뷰

제 3장 OntoFrame과 Scholar.Google 기능적 효용성 평가를 위한 사용자  
테스트

3.1 사용자 테스트 환경

3.2 사용자 테스트 일정

3.3 사용자 테스트 결과 정리 견본

제4장 OntoFrame과 Scholar.Google 기능적 효용성 평가를 위한 사용자  
테스트를 통한 결과 분석

4.1 사용자 테스트 결과 분석 방법

4.2 사용자 테스트 결과 분석

제5장 OntoFrame과 Scholar.Google 사용자 테스트를 통한 기능적  
효용성 평가 결과 정리

## 제 1장 Software 품질 및 기능적 효용성에 대한 고찰

소프트웨어 품질은 사용자의 시각에서 결정된다. 사용자들이 원하는 작업을 쉽고 빠르게 수행할 수 있도록 소프트웨어가 작동될 때, 그리고 소프트웨어가 처리한 결과가 사용자가 기대하였던 바와 일치할 때, 우리는 “좋은 소프트웨어(Good Software)”라 한다. 소프트웨어는 사용자가 필요로 하는 기능성과, 성능을 가져야 하며, 신뢰성과 사용성이 있어야 한다. 따라서 좋은 소프트웨어를 만들기 위하여, 일련의 소프트웨어 프로세스가 필요하다. 예를 들어 소프트웨어 품질관리 활동 또한 이러한 소프트웨어 프로세스 Activities의 하나이다.

ISO/IEC 9126에 따르면, 프로세스 품질(Process Quality in ISO/IEC12207)은 제품품질(Product Quality)에 영향을 미치며, 제품의 품질은 상용상의 품질(Quality in Use)에 기여하는 것으로 서술하고 있다.

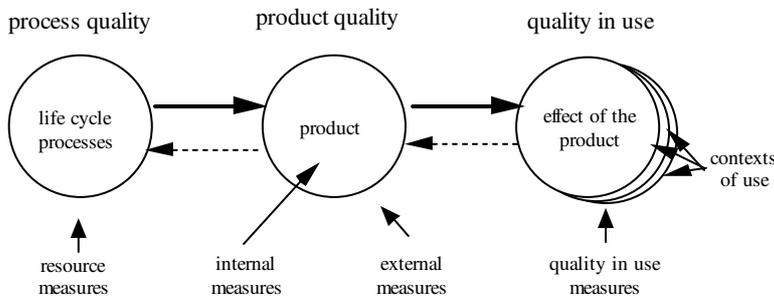
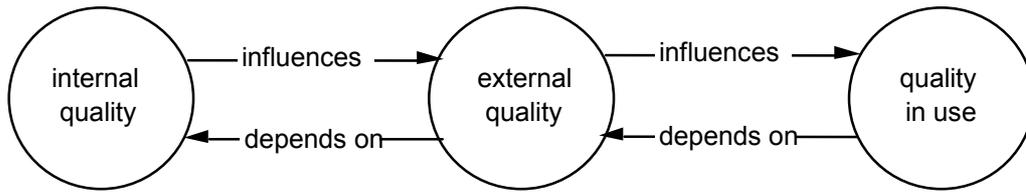


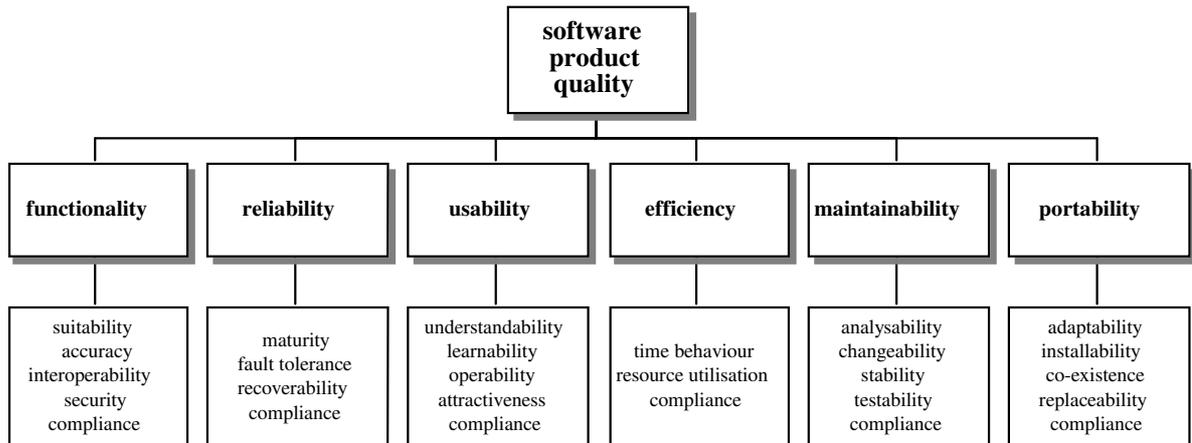
그림 1: Approaches to software quality (ISO 9126-1)

사용상의 품질(Quality in Use)은 소프트웨어 품질에 대한 사용자의 관점이다. 사용상의 품질은 소프트웨어를 사용하는 결과로 측정된다. 아래 [그림 2]는 소프트웨어 품질사이의 관계를 보여주고 있다.



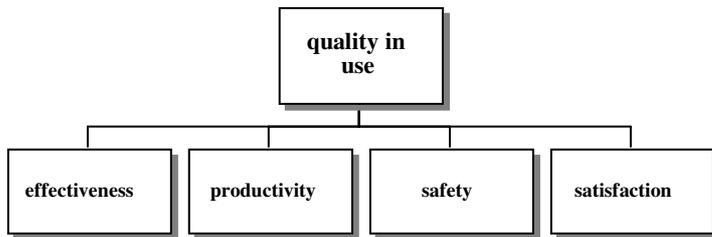
**그림 2: Relationship between different vies of quality**

소프트웨어 품질은 품질 모델을 사용하여 평가하여야 한다. 소프트웨어 품질속성은 [그림 3]에서 보는 바와 같이 기능성(Functionality), 신뢰성(Reliability), 사용성(Usability), 효율성(Efficiency), 유지보수성(Maintainability), 이식성(Portability) 6개 특성으로 범주화 될 수 있다. 이는 품질 모델은 소프트웨어 품질을 다루기 위한 체크리스트로 사용될 수 있다.



**그림 3: Software product quality**

[그림 4]는 사용자 관점에서 사용상의 품질이 효율성(Effectiveness), 생산성(Productivity), 안전성(Safety), 만족성(Satisfaction) 4개의 특성에 달려 있다는 것을 보여주고 있다. 사용상의 품질은 internal measures와 연결된 external measures에 종속된다[그림 2참조].



**그림 4: Software Quality in Use**

소프트웨어 사용상의 품질을 평가하기 위한 메트릭스(Metrics)는 “사용자가 해당 소프트웨어를 이용하여 수행하고자 하는 목적을 얼마나 효율적으로, 정확하며, 만족스럽게 할 수 있는가?” 계측하기 위한 지표로서, “사용자 작업 시나리오” (User-task scenarios)에 따라 소프트웨어 품질을 평가한다. 소프트웨어 사용상의 품질을 평가하기 위한 ISO/IEC 9126 매트릭스는 추상적 정의이다. 따라서 실제 소프트웨어 사용상의 품질을 평가하기 위한 매트릭스는 소프트웨어 유형에 따라 세부적인 측면에서 재 정의되어야 할 필요가 있다. 예를 들어 OntoFrame과 Scholar.Google와 같은 웹 기반 정보서비스 시스템의 경우, 웹 상에서 정보검색서비스를 제공하는 소프트웨어 유형의 관점에서, 테스트 프로세스에 따라 사용상의 품질 평가를 위한 매트릭스를 재 정의하여야 한다[2장 참조].

위에서 이미 살펴본 바와 같이, 소프트웨어의 품질에 있어서, 최종 사용자에게 직접적으로 관계되는 항목은 “사용상의 품질”이다. 사용상의 품질은 소프트웨어의 6개 품질 속성 중 기능성, 신뢰성, 사용성 그리고 효용성을 평가하여 살펴볼 수 있다. 따라서 소프트웨어 사용상의 품질을 평가하기 위한 메트릭스는 크게 아래 4개 항목을 측정하는 항목으로 구성된다.

- 사용자의 작업을 지원하는 기능을 소프트웨어가 제공하고 있는가?
- 소프트웨어 기능을 활용한 사용자 액션에 따른 결과의 신뢰성(정확성)이 얼마나 충족되고 있는가?
- 해당 소프트웨어의 기능을 사용자가 쉽고 간결하게 이용할 수 있는가?
- 소프트웨어를 이용한 사용자의 작업 과정 및 작업결과가 얼마나 효율적이며, 이에 대한 사용자의 만족도는 어떠한가?

결론적으로, 소프트웨어 사용상의 품질은 본질적으로 “소프트웨어 기능의 효용성”에 관계되는 것으로 귀추된다. 즉, “사용자가 요구하는 기능을 소프트웨어가 제공하며, 사용자가 해당 소프트웨어의 기능을 사용하여 원하는 작업 결과를 얼마나 쉽고 빠르고 정확하게 얻을 수 있는가”라는 소프트웨어 기능의 효용성이 소프트웨어 사용상의 품질을 결정한다고 할 수 있다.

"OntoFrame과 Scholar.Google 사용자 테스트를 통한 기능적 효용성 평가"는 정보서비스 시스템 최종사용자를 대상으로 다음과 같은 관점에서 이루어진다;

- OntoFrame과 Scholar.Google과 같은 정보서비스 시스템에 대하여 사용자들은 어떠한 기능을 기대하는가?
- 사용자는 원하는 작업 결과를 얻기 위하여 OntoFrame과 Scholar.Google 가 제공하는 기능을 적절히 활용할 수 있는가?
- OntoFrame과 Scholar.Google에서 얻어진 작업의 결과에 대한 사용자의 신뢰성(서비스 된 정보의 정확성)이 얼마나 높은가?

- 원하는 정보를 서비스 받기 위한 사용자의 작업 과정, 작업 시간, 작업 만족도는 어떠한가?

## 제2장 OntoFrame과 Scholar.Google 기능적 효용성 테스트 프로세스

소프트웨어 사용상의 품질 또는 기능적 효용성을 테스트(이하 소프트웨어 기능적 효용성 테스트)한다는 것은, 사용자가 소프트웨어를 이용하여 문제를 해결하는 과정을 직접적으로 관찰하는 것을 의미한다.

소프트웨어 테스트 프로세스는 다음과 같이 단계로 구성될 수 있다.

- 1단계: 테스트 목표, 테스트 범위 정의- 소프트웨어 사용자 테스트는 개발 중인 소프트웨어를 테스트하여, 개발 목표에 부합하는지 여부를 평가하거나, 또는 개발 완료된 소프트웨어의 완성도를 높일 목적으로 수행될 수 있다. 소프트웨어 사용자 테스트는 테스트 목표 정의에 따라 테스트 프로세스에 영향을 미치게 된다.
- 2단계: 테스트 목적에 따른 작업(Task) 정의- 테스트 수행 목적에 따라 사이트 목표와 사용자 목표를 평가 할 수 있는 테스트 과제, 인터뷰 주제 등이 결정된다.
- 3단계: 테스트 장소, 일정, 참가인원, 테스트 방법 정의 - 테스트 장소, 일정, 테스트 사용자, 테스트 방법, 테스트 진행 기록 등에 대한 테스트 최종 준비 단계에 해당된다.
- 4단계: 테스트 프로세스 정의 - 실제 테스트 진행 순서를 명시한다.
- 5단계: 테스트 실행 - 테스트 프로세스에 따라 오리엔테이션, 과제 시작, 테스트 진행 기록, 소요시간 측정, 테스트 후 인터뷰 등이 테스트 실행 단계에서 행하여 진다.
- 6단계: 테스트 결과 분석, 평가 - 테스트 과제 수행 결과 평가,

분석이 이루어 진다.

- 7단계: 테스트 결과에 따른 개선안 반영 - 사용자 테스트 목적에 따라 적용될 수 있는 테스트 결과를 반영하여, 소프트웨어 개선이 이루어 진다.

OntoFrame과 Scholar.Google의 기능적 효용성 테스트 프로세스는 다음과 같은 테스트 프로세스에 따라 진행된다.

- ㄱ. 웹 정보검색서비스 사용자 사전 인터뷰(Pre-Interview)를 통한 사용자 요구 분석 (2.3장 참조)
- ㄴ. 사전 인터뷰 결과에 따라 객관적이고 공정한 사용자 테스트를 위한 테스트 시나리오 작성(2.4장 참조)
- ㄷ. 사용자 테스트 실행(3장 참조)
- ㄹ. 테스트 결과 분석( 4장 참조)

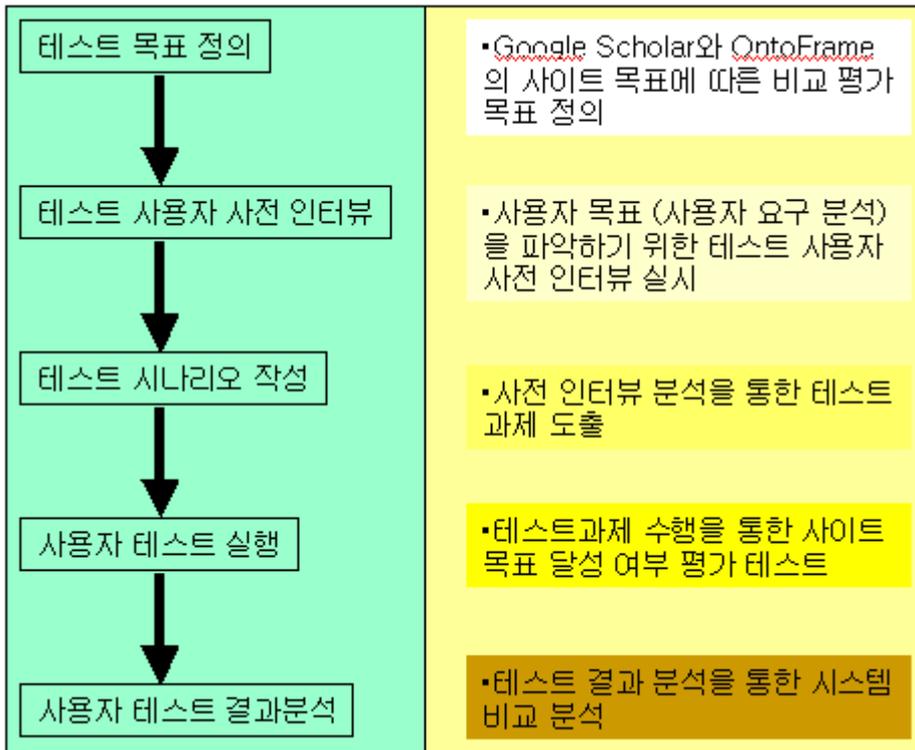


그림 5: OntoFrame과 Scholar.Google의 기능적 효용성 테스트 프로세스

## 2.1 테스트 목표, 테스트 범위 정의

Scholar.Google는 학술정보서비스를 선도하고 있다. Scholar.Google의 사이트 목표는 전문 학술정보를 서비스 하는 것이다. Scholar.Google의 사용자 목표는 특정 주제에 대한 연구정보를 검색하거나, 특정 주제를 연구하는 선도적 연구자 정보 등 사용자가 필요로 하는 정보 Needs를 해결하는 것이다.

OntoFrame의 사이트 목표 및 사용자 목표 또한 Scholar.Google와 크게 다르지 않다.

OntoFrame과 Scholar.Google 기능적 효용성 평가를 위한 사용자 테스트 목표는 “OntoFrame과 Scholar.Google의 기능적 효용성 비교 평가를 통하여, OntoFrame에 대한 시장성 및 기술성 제고, 향후 연구 개발 방향 설정”에 있다.

## 2.2 테스트 사용자 선정 및 사전 인터뷰

OntoFrame과 Scholar.Google 기능적 효용성 평가를 위한 테스트 사용자(Test Person, TP)는 테스트 목표에 부합하도록 선정되어야 할 필요가 있다. OntoFrame과 Scholar.Google의 이용자는 일반적 Web 유저들이 아니라, 특정 분야의 전문가들이라 할 수 있으며, 따라서 석사 이상의 전문지식을 지닌 연구개발 인력 중 특정 분야의 학술 논문을 수시로 검색하여야 하는 필요성을 가지고 있는 직업군에서 선정되어야 한다.

테스트 사용자 선정이 테스트 프로세스에서 우선적으로 수행되어야 하는 필요성은, 테스트 사용자들이 전문 학술정보를 검색하기 위하여 어떠한 정보서비스 시스템을 이용하고 있는지, 어떠한 종류의 정보를 주로 필요로 하고 있는지, 테스트 사용자들이 정보 Needs를 해결하기 위하여 이용하는 방법 등을 미리 파악하여, 객관적이고 공정한 테스트 시나리오를 작성하는데 반영하여야 하기 때문이다.

### 가. 테스트 사용자 선정

25~40연령대에서 석사학위 이상 학력을 가지고 3년 이상 소프트웨어 연구개발 업무에 종사하고 있는 전문 직업군에서 테스트 사용자를

선정하였다.

[표 1] 테스트 사용자의 직업 연령에 따른 구분

Test Person	직업 구분	연령구분	학위구분	성별구분
TP1(전)	대학 교수	45	박사	남
TP2(L)	대학 교수	28	석사	여
TP3(김)	연구소	48	박사	남
TP4(유)	연구소	43	박사	남
TP5(권)	기업	35	석사	여
TP6(이)	기업	43	박사	남

[표 2] 테스트 사용자의 업무 및 연구개발 업무 기간,

Test Person	전담 업무	업무 기간(년)
TP1	교육, 연구	15
TP2	교육, 연구	3
TP3	연구 개발	16
TP4	연구 개발	13
TP5	연구 개발	7
TP6	연구 개발	10

#### 나. 테스트 사용자 수

Jakob Nielsen의 연구결과에 따르면, 3명의 사용자 테스트를 통하여, 전체 문제점 중 70%를 발견할 수 있다고 한다.

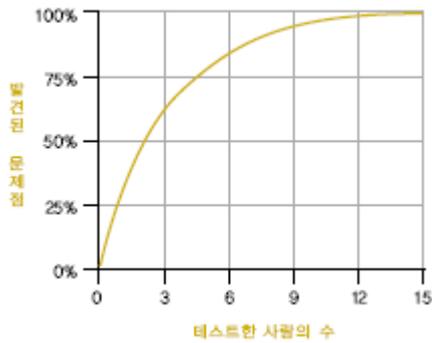


그림 6: 테스트 사용자와 테스트를 통하여 발견되는 문제의 관계 그래프 (J. Nielsen)

OntoFrame과 Scholar.Google 기능적 효용성 평가를 위한 테스트 사용자 수는 6명으로 하였다. 선정된 6명의 테스트 사용자는 풍부한 정보서비스 시스템 이용 경험을 가지고 있다. 사용자의 이러한 선지식(Pre-Knowledge)은 테스트를 통하여 달성하고자 하는 목표에 기여할 것으로 가정되며, 따라서 6명의 테스트 사용자 수는 충분하리라 판단된다.

#### 다. 사전 인터뷰(Pre-Interview) 질의문

사전 인터뷰는 정보서비스 시스템 사용자의 요구를 분석하고 파악하기 위하여 실시되었다. 사용자의 요구는 크게 다음과 같은 2가지 관점에서 파악 분석된다.

- 정보서비스 시스템 기능에 대한 사용자 요구
- 사용자의 Information Needs 측면에서의 사용자 요구

위 2가지 관점에서 사용자 요구분석을 위한 사전 인터뷰 질의문은 다음과 같다.

[표 3] 사전 인터뷰 질의문

번호	질의문
1	전문적 학술정보 검색을 위하여 어떠한 서비스를 이용하고 계십니까?
2	전문적 학술정보 검색을 위하여 주 평균 몇 회 정보서비스 시스템을 이용하십니까. 주 평균 총 이용시간은 얼마입니까?
3	주로 검색하시는 정보의 종류 (논문, 기술동향, 연구자 정보, 연구기관정보 등)는 어떠한 것이 있습니까?
4	현재 사용하시는 검색 서비스 시스템의 장단점을 나름대로 말씀하여 주십시오.
5	평소 정보를 검색하시면서, 필요하다고 느껴졌던 시스템 기능으로서 어떠한 것들이 있는지 말씀하여 주십시오.
6	특정한 부가적 정보(예를 들면, 핵심 연구자 그룹, 연구기관, 연구자간의 관계 등)를 찾으시는 경우가 있다면, 어떠한 방법으로 해결하시는지 보여주시거나 말씀하여 주십시오.
7	정보검색 시스템에서 가장 중요하게 여기는 특징(검색결과의 정확성, 검색 속도, 사용편의성, 검색결과 순위 등)이 있다면 어느것이 있습니까?
8	정보서비스 시스템이 제공하는 검색결과가 다수의 페이지에 걸쳐 주어졌을 때, 검색결과 내 몇 페이지 또는 몇 리스팅 순위까지 탐색하십니까?

사전 인터뷰의 응답 결과는 객관적이고 공정한 테스트를 위한 테스트 시나리오(Test Task) 작성 기초로 활용된다.

## 2.3 사전 인터뷰 분석

사전 인터뷰는 테스트 사용자를 직접 방문하여 질의에 대한 응답을 기록하는 방식으로 행하여 졌다(4명). 2명의 테스트 사용자는 실제 테스트가 행하여 질 공간을 방문하여 실행하였다. 사전 인터뷰 분석 결과는 다음과 같다.

[표 4] 사전 인터뷰 질의 1과 질의 2 결과: 이용하는 정보서비스 시스템, 주당 평균 이용 회수, 주당 평균 이용시간

TP	TP1	TP2	TP3	TP4	TP5	TP6
이용 시스템	society.kis ti..re.kr	Portal.acm .org	scholar.go ogle.co.kr	scholar.go ogle.co.kr	www.nd sl.kr	scholar.go ogle.co.kr
이용 회수	1	3	2	3	2	1
이용 시간 (분)	90	150	120	100	60	60

6명의 TP 중 3명이 전문학술정보 검색을 위하여 scholar.google.co.kr를 이용하고 있었으며, society.kisti.re.kr 이용자 1명, portal.acm.org 이용자 1명, www.ndsl.kr 이용자 1명으로 파악되었다.

TP가 전문 학술정보를 검색하기 위하여 정보서비스 시스템을 이용하는 주당 평균 횟수는 2회 정도이며, 소요하는 평균 시간은 1시간 40분 이다.

[표 5] 사전인터뷰 질의 3에 대한 결과: 검색하는 정보 종류

TP	응답
TP1	연구논문>연구보고서
TP2	연구논문
TP3	기술동향> 연구논문> 연구자 정보> 연구기관 정보
TP4	연구논문> 기술동향
TP5	연구논문> 기술동향
TP6	연구논문> 기술동향>연구기관

(‘>’는 검색하는 정보 종류 중 TP가 상대적으로 우선시 함을 표시)

TP는 주로 연구 논문을 검색하기 위하여, 정보서비스 시스템을 이용하고 있다. 연구소 또는 기업에 재직 중인 TP의 경우 기술 동향 파악 및 부가적 정보를 검색하기 위하여 전문학술정보 서비스 시스템을 이용하는 것으로 나타났다.

[표 6] 사전인터뷰 질의 4에 대한 결과: 시스템 장단점

TP	응답
TP1	국내 학술지 정보 검색에 장점이 있다
TP2	검색된 정보에 대한 요약정보를 제공하는 것이 좋다
TP3	최신 논문 및 정보의 량이 방대하면서도 속도가 빠르다. 검색되는 정보가 너무 방대하여, 검색된 자료들을 다시 스스로 탐색하는 것이 불편하다.
TP4	키워드 검색 정확성이 좋다. 검색결과 리스팅 순위를 믿을 수 있다. 초기화면의 단순성이 좋다. 검색결과가 너무 방대하다.

<b>TP5</b>	학술논문, 학위논문, 해외저널, e-book등 다양한 정보를 찾을 수 있어 좋다.
<b>TP6</b>	IEEE 논문 등 최신 정보와 최신 링크를 찾을 수 있고, 풍부한 정보를 제공하는 점이 좋다. 그러나 획일적인 키워드 검색에 치중되고, 개인화된 서비스를 받을 수 없는 점이 아쉽다.

TP는 대부분 키워드 검색과 같은 전통적 검색 기법을 이용하여 정보를 검색하고 있는 것으로 응답하였다. 각 시스템의 장단점으로는 TP에 따라 약간씩 차이는 있으나, 대체로 필요로 하는 정확한 정보 검색 관점에서 장 단점을 언급하였다. 특히, Scholar.Google 사용자의 경우, 방대하게 검색 제공되는 정보를, 스스로 탐색하여 필터링하여야 하는 점을 단점으로 지적하였다.

[표 7] 사전인터뷰 질의 5, 6에 대한 결과: 정보서비스 시스템에서 사용자가 필요로 하는 기능, 특정한 부가적 정보 획득 방법

<b>TP</b>	<b>응답</b>
<b>TP1</b>	주제 분류와 같은 방식의 유연성 있는 검색 방식이 있으면 좋겠고, 키워드 검색으로 논문간의 상호 밀접도를 분석하여 제시하는 기능이 있으면 좋겠다. 키워드 검색으로 논문을 찾고, 저자별로 소팅한 다음, 수동으로 저자와의 선지식으로부터 유추함
<b>TP2</b>	주제 분류를 통하여 검색 도움기능이 있으면 좋겠다 키워드 검색으로 논문을 찾은 다음, 선지식으로부터 유추한다
<b>TP3</b>	Full Text Retrieval기능이 있으면 좋겠다. 부가 정보 검색기능이 있으면 좋겠다.

	선지식으로부터 유추한다. 키워드로 연구논문을 찾고, 수동으로 살펴본다.
<b>TP4</b>	데이터 포맷(doc, ppt, pdf, hwp, avi, mpeg 등)에 따른 검색결과 소팅기능이 있으면 좋겠다
	부가적 정보를 찾기 위하여 특화된 사이트를 이용한다. 예를 들어 기관정보를 모아놓은 사이트를 방문하여 기관링크를 찾아 이동한다. 또는 선지식에 근거하여 추론하기도 한다.
<b>TP5</b>	
	웹 검색을 통하여 유추하거나, 인용자료를 찾아 참고한다.
<b>TP6</b>	
	Bottom-Up방식으로 먼저 키워드 검색으로 논문을 찾은 다음, 선지식에 기초하여 유추한다

TP가 필요로 하는 정보서비스 시스템의 기능으로는 주제에 따른 분류 기능을 언급하였으며, 학술정보와 연관된 부가정보 검색 기능이 필요하다고 응답하였다. 부가적 정보를 찾기 위한 방법으로, 대부분의 TP가 키워드 검색 후 스스로 선 지식에 근거한 추론을 통하여 해결하고 있었다.

[표 8] 사전인터뷰 질의 7에 대한 결과: 정보서비스 시스템에 사용자가 중요하게 여기는 특징 순위

TP	응답			
	1순위	2순위	3순위	4순위
순위				
TP1	검색결과정확성	리스팅순위	사용편의성	검색속도

TP2	검색결과정확성	리스팅순위	사용편의성	검색속도
TP3	검색속도	리스팅순위	검색결과정확성	사용편의성
TP4	사용편의성	검색결과정확성	검색속도	리스팅순위
TP5	검색결과정확성	사용편의성	리스팅순위	검색속도
TP6	검색결과정확성	리스팅순위	사용편의성	검색속도

정보서비스 시스템에 대하여 TP 대부분이 검색결과와 정확성을 중요하게 여기고 있었다. 또한 대부분의 시스템이 대량의 검색된 정보를 제공하고 있음에 따라, 검색된 정보의 리스팅 순위에도 높은 관심을 나타내었다.

[표 9] 사전 인터뷰 질의 8에 대한 결과: 검색결과 내에서 탐색페이지 수

TP	응답
TP1	3페이지 이내, 대략 검색결과 리스팅 순위 30위 이내에서 탐색한다.
TP2	평균적으로 3페이지 이내에서 탐색한다.
TP3	검색결과에서 3페이지 이상은 탐색하지 않는다.
TP4	검색결과가 대량으로 발견될 경우, 4~5페이지까지 검색결과 페이지를 탐색하는 경우도 있다.
TP5	검색결과 3페이지 이내를 넘지 않는다.
TP6	검색결과 내에서 2~5페이지까지 탐색한다. 그 이하는 정확성이 떨어진다.

TP는 대량으로 검색되어 리스팅 되는 정보 내에서 정확한 정보를 찾아내기 위하여, 제시된 정보를 하나씩 열어 확인 작업을 수행한다.

TP는 대부분 3페이지 (리스팅 순위 30위) 이내에서 탐색을 멈추는 것으로 보인다. 다시 말하여, TP 대부분은 검색된 정보의 리스팅 순위를 중요하게 여기며, 신뢰하려 한다는 것을 보여준다.

사전 인터뷰 결과를 종합하여 보면, TP의 User Goal은 정확한 정보를 검색 하는데 있다. 또한 TP가 시스템에 기대하는 것은 정확한 정보를 쉽고 간편하게 검색할 수 있도록 지원하는 시스템 기능이라 할 수 있다. 사전 인터뷰 결과는 사용자 테스트 태스크 (User Test Task, UTT) 작성을 위한 기반 자료로 활용되었다.

## 2.4 사용자 테스트 시나리오 및 테스트 사후 인터뷰

UTT를 위한 시나리오는 2.1장의 테스트 목표와, 2.3장의 사전 인터뷰 결과 등 사용자 요구분석에 기초하여, 객관성 및 공정성을 보장할 수 있도록 작성되었다. Scholar.Google와 OntoFrame은 전문 학술정보서비스 시스템이라는 공통 분모를 가지고 있으나, 세부 기술적인 면에서는 상호 다른 점들을 가지고 있다. 예를 들어 Scholar.Google와 OntoFrame의 데이터 집합(Data Sets)은 상호 상이하며, 데이터 색인-관리-검색-서비스 전반에 걸친 메커니즘도 다르다. 따라서 OntoFrame과 Scholar.Google 기능적 효용성 평가를 위한 시나리오 (Test Task)는 아래의 요점을 포괄하여야 한다.

- ㄱ. OntoFrame과 Scholar.Google의 공통되는 Site Goal
- ㄴ. 정보서비스 시스템 이용자가 가지는 다양한 User Goal

ㄷ. Pre-Interview분석 결과로 도출된 시스템 기능에 대한 Specific User Needs

UTT는 OntoFrame과 Scholar.Google 기능적 효용성을 평가할 수 있도록 3개 레벨의 난이도를 가지도록 하였다.

- A. 전통적 기술을 활용한 사용성 평가과제(Level I): 잠재적 사용자 대부분이 이미 알고 있다고 가정할 수 있는 전통적 방식으로 해결 가능한 과제.
- B. 추론을 필요로 하는 사용성 평가과제(Level II): 잠재적 사용자 대부분이 선지식으로부터 유추하여 해결 가능한 과제.
- C. 고급 기능을 필요로 하는 사용성 평가과제(Level III): 잠재적 사용자의 일부 만이 선지식으로부터 추론하여 해결 가능하거나, 시스템에 적용된 고급기능을 활용하여야 해결가능한 과제.

#### 가. 사용자 테스트 태스크(User Test Task)

UTT는 사전 인터뷰 분석 결과를 참조하여 위 언급된 요점을 포괄할 수 있도록 작성되었다. 총 5가 테스트 태스크가 다음과 같이 준비되었다.

- UTT1: neural network에 관한 연구 논문 중 sufficient condition을 다루고 있는 5개의 논문을 검색하여 요약문과 함께 리스팅하십시오. 검색된 논문 중, 논문의 제목에서 두개의 개념이 포함되거나, 또는 요약문에 neural network과 sufficient condition에 관한 서술이

있어야 합니다.

- UTT2: neural network 연구 분야 중 최근 연구 동향을 주도하고 있다고 판단되는 5개의 주제 토픽(Topic)을 찾아 리스팅하십시오(예: hierarchical neural network). 각각 상세분야의 논문 1편을 찾아 요약문과 함께 리스팅하십시오.
- UTT3: Jinde Cao씨의 주요 연구 토픽 5개를 찾아 리스팅하시고, 각각 연구토픽을 대표하는 논문 1편을 찾아 요약문과 함께 리스팅하십시오.
- UTT4: Jinde Cao씨와 Fengli Ren씨의 공통 연구 토픽 5개를 찾아 리스팅하십시오.
- UTT5: neural network분야 연구를 선도하고 있는 기관 5개를 찾아 리스팅하십시오.

[표 10] 테스트 태스크의 요점 부합 여부

UTT	Site Goal	User Goal	Specific User Needs
UTT 1	●	●	○
UTT 2	●	●	◐
UTT 3	●	●	◐
UTT 4	◐	◐	●
UTT 5	○	◐	●

(● 요점에 강하게 부합됨, ◐ 부합됨, ○약하게 부합됨)

UTT 1은 OntoFrame과 Scholar.Google가 가지는 Site Goal인 “정보

제공”, UTT의 User Goal인 “정확한 정보의 발견”이라는 요점에 강하게 부합된다. 그러나 UTT1은 사용자의 특별한 정보 Needs를 해결하기 위한 과제라기 보다는 일반적인 키워드 검색을 요구하는 과제이다. UTT 2와 UTT 3는 Site Goal인 “정보 제공”, User Goal인 “정확한 정보의 발견”이라는 요점에 강하게 부합되고 있으며, 약간은 특별한 사용자의 정보 Needs를 해결하여 주어야 하는 과제이다. UTT 5는 일반적인 정보서비스 시스템이 가지고 있는 Site Goal에 약하게 부합되며, 그 보다는 특정 사용자 그룹이 가지는 User Needs에 강하게 부합된다.

[표 11] UTT 난이도 Level

UTT	Level I	Level II	Level III
UTT 1	●	○	○
UTT 2	◐	●	◐
UTT 3	◐	●	◐
UTT 4	○	●	◐
UTT 5	○	◐	●

(● Level 수준에 강하게 부합됨, ◐ 부합됨, ○약하게 부합됨)

UTT 난이도 측면에서 UTT 1은 TP가 익숙하게 사용하고 있는 키워드 검색을 통하여 해결할 수 있는 과제이다. 이와 달리 UTT 4와 UTT 5의 경우 전문적 도메인 지식을 가진 사용자가 추론을 통하여 해결하거나, 특별한 시스템 고급기능을 요구하는 과제이다.

위와 같이, UTT는 테스트 목표에 따라 OntoFrame과

Scholar.Google의 Site Goal 달성 정도, User Goal 지원 수준을 평가하기 위한 다면적 기준에 따라 구성되었다.

#### **나. 테스트 사용자를 대상으로 테스트 사후 인터뷰 질의**

사용자 테스트 종료 후 각각의 사용자를 대상으로 다음과 같은 사후 인터뷰가 실시되었다.

- ① OntoFrame과 Scholar.Google의 기능성에 대한 총평을 하여주시기 바랍니다. 가능하면 구체적인 사례를 들어 주십시오.
- ② OntoFrame과 Scholar.Google의 사용성에 대한 총평을 하여주시기 바랍니다. 가능하면 구체적인 사례를 들어 주십시오.
- ③ OntoFrame과 Scholar.Google의 기능성 및 사용성 측면에서 향후 기대하고 싶은 점이 있다면 말씀하여 주십시오.

사후 인터뷰는 사용자로부터 피드백을 얻기 위한 목적으로 주어졌다. 인터뷰를 통하여 얻어진 결과는 향후 시스템 성능 개선에 도움을 줄 수 있을 것으로 기대된다.

## 제3장      OntoFrame과      Scholar.Google      기능적

### 효용성 평가를 위한 사용자 테스트

#### 3.1 사용자 테스트 환경

사용자 테스트를 실행하는데 있어서, 다음과 같은 휴리스틱(Heuristic)이 적용되었다.

- ① Visibility of system status: 시스템의 상태를 TP와 관찰자가 가시적으로 같은 화면을 볼 수 있도록 구성한다.
- ② Match the system to the real world: 시스템 상에서 TP의 실제 작업 환경과 최대한 일치하도록 한다.
- ③ User control and freedom: TP는 테스트 환경 및 진행에 있어서 자유를 가진다.
- ④ Help and assistance: TP가 요청할 경우 테스트 결과에 영향을 미치지 않는 범위 내에서 테스트 진행과 관련된 TP 질문에 도움을 제시한다.

#### 가. 테스트 장소 및 테스트 실험실 배치도

테스트 장소는 배재대학교 정보통신과학관 508호에서 실행되었다.

테스트 실험실 배치도는 [그림 7]과 같다.

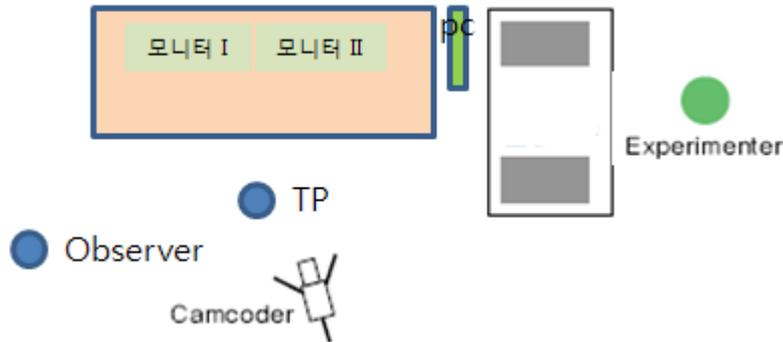


그림 7: 사용자 테스트 실험실 배치도

Observer는 사용자 테스트에 영향을 미치지 않는 범위 내에서, 테스트 장면을 관찰하며, 특이한 사용자의 액션을 찾아 기록하는 역할을 수행한다. TP의 모니터 화면은 비디오 녹화되었다. 평가자인 Experimenter는 별도로 설치된 모니터를 통하여, 사용자와 동일한 화면을 보며, 사용자의 작업 내용을 관찰 기록한다.

#### 나. 실험 시스템 사양 및 구성

실험에 이용된 실험 시스템 사양은 다음과 같다.

- PC: Intel Dual Core Q9400, 2Gb Ram, 19'inch Monitor 4대
- Network 속도: 100Mbps
- Browser: MS-Explorer v. 7.0.5730.13
- Camcorder: Sony Handycam

모니터 I에 OntoFrame과 MS-Word processor, 모니터 II에 Scholar.Google와 MS-Word processor를 활성화 하여, TP는 테스트 결과를 Copy & Paste하도록 하였다. 즉 TP는 모니터 I에서

OntoFrame을 이용하여 얻은 결과를 모니터 II의 MS-Word를 이용 기록하고, 모니터 II에서 Scholar.Google를 이용하여 얻은 결과를 모니터 I의 MS-Word에 Copy&Paste하도록 하였다.

### 3.2 사용자 테스트 일정

사용자 테스트 일정은 2008년 11월 11일부터 11월 14일까지 다음과 같이 4일간 실행되었다.

[표 12] 사용자 테스트 일정

TP	테스트 일시
TP 1	11월 13일 14:00~
TP 2	11월 11일 11:00~
TP 3	11월 12일 11:00~
TP 4	11월 13일 10:30~
TP 5	11월 13일 15:00~
TP 6	11월 14일 10:39~

### 3.3 사용자 테스트 정리 견본

TP가 UTT를 수행하며 얻은 정보를 Copy&Paste 방식으로 기록하여

Word파일로 제시한 결과는 평가자(Experimenter)에 의하여 다음과 같이 정리되었다.

가. TP4가 Scholar.Google에서 과제 수행 결과로서 제시한 Word파일을 정리한 견본

문제: 1            답: 1            검색페이지: 1            페이지 순위: 7

**Necessary and sufficient condition for absolute stability of neural networks**

Forti, M.    Manetti, S.    Marini, M.  
Dept. of Electron. Eng., Florence Univ.;

**Abstract**

The main result in this paper is that for a neural circuit of the Hopfield type with a symmetric connection matrix  $T$ , the negative semidefiniteness of  $T$  is a necessary and sufficient condition for Absolute Stability. The most significant theoretical implication is that the class of neural circuits with a negative semidefinite  $T$  is the largest class of circuits that can be employed for embedding and solving optimization problems without the risk of spurious responses

정확성판단: 1

문제: 1            답: 2            검색페이지: 2            페이지 순위: 1

**Stability analysis of delayed neural networks**

Arik, S.  
Dept. of Electron., Istanbul Univ.;

**Abstract**

In this paper, the author derives several sufficient conditions for the asymptotic stability of delayed neural networks. These conditions ensure the asymptotic stability of the equilibrium point of a delayed neural network independently of the delay parameter. The results obtained impose constraint conditions on the interconnection matrix of the neural system. These results are also compared

with the earlier results derived in the literature

정확성판단: 1

TP 는 위 예에서처럼 각각의 UTT 에 대하여 5 개의 정확하다고 가정되는 정보를 찾아 제시하며, 평가자는 제시된 결과에 대하여 정확성 판단을 하게 된다.

**문제: 1                      답: 1                      검색페이지: 1                      페이지 순위: 7** 에서

문제:1 은 UTT 번호이며, 답:1 은 UTT1 에 대하여 제시된 5 개의 정보 중 첫번째 임을 의미한다. 검색페이지:1 은 UTT1 검색을 수행하였을 때, 답 1 이 첫번째 검색페이지에서 탐색되어 제시되었음을 나타낸다. 페이지 순위: 7 은 제시된 답:1 정보가 첫번째 검색페이지의 일곱번째 순위로 리스팅 되어 있었음이 평가자에 의하여 확인된 것이다.

정확성판단: 1 은 평가자에 의하여 TP 가 제시한 정보가 UTT 에 결과로 제시된 정확한 정보라고 판단되었음을 의미한다. 정확성판단: 0 경우 제시된 정보는 UTT 의 결과로 부정확한 정보로서 판단될 경우이다.

#### 나. PT 3 이 OntoFrame 에서 과제 수행 결과로 제시한 Word 파일을 정리한 견본

**문제: 2                      답: 1                      검색페이지: 1                      페이지 순위: 1**

Topic : **backpropagation neural network**

**Classification of smoking cessation status with a backpropagation neural network**

This study examined the ability of a **backpropagation neural network** (BPNN) classifier to distinguish between current and former smokers in the 2000 National Health Interview Survey (NHIS) sample adult file. The BPNN classifier performance exceeded that of random chance, with asymmetric 95% confidence intervals for  $A_z$  (area under receiver operating characteristic curve) = (0.7532, 0.7790). Separation of current and former smokers was imperfect, as illustrated by the receiver operating characteristic (ROC) curve. Additionally, performance did not exceed that of a comparison classifier created using logistic regression. Attribute subset selection identified three novel attributes related to **smoking cessation status**. This study establishes the ability of **backpropagation neural networks** to classify a complex health behavior, **smoking cessation**. It also illustrates the hypothesis-generating capacity of data mining methods when applied to large population-based health survey data. Ultimately, BPNN classifiers of **smoking cessation status** may be useful in decision support systems for **smoking cessation** interventions.

정확성 판단: 1

문제: 2      답: 2      검색페이지: 1      페이지 순위: 1

Topic : **fire prediction**

**Neural network and GA approaches for dwelling fire occurrence prediction**

(2006) [Abstract](#)

This paper describes three **approaches** for the **prediction** of **dwelling fire occurrences** in Derbyshire, a region in the United Kingdom. The system has been designed to calculate the number of **fire occurrences** for each of the 189 wards in the Derbyshire. In terms of the results from statistical analysis, eight factors are initially selected as the inputs of the **neural network**. Principal Component Analysis (PCA) is employed for pre-processing the input data set to reduce the number of the inputs. The first three principal components of the available data set are chosen as the inputs, the number of the **fires** as the output. The first **approach** is a logistic regression model, which has been widely used in the forest **fire prediction**. The **prediction** results of the logistic regression model are not acceptable. The second **approach** uses a feed-forward **neural network** to model the relationship between the number of **fires** and the factors that influence **fire occurrence**.

The model of the **neural network** gives a **prediction** with an acceptable accuracy for the **fires** in **dwelling** areas. Genetic algorithms (**GAs**) are the third **approach** discussed in this study. The first three principle components of the available data set are classified into the different groups according to their number of **fires**. An iterative **GA** is proposed and applied to extract features for each data group. Once the features for all the groups have been identified the test data set can be easily clustered into one of the groups based on the group features. The number of **fires** for the group, which the test data belongs to, is the **prediction** of the **fire occurrence** for the test data. The three **approaches** have been compared. Our results indicate that the **neural network** based and the **GA** based **approaches** perform satisfactorily, with MSEs of 2.375 and 2.875, respectively, but the **GA approach** is much better understood and more transparent.

정확성 판단: 1

문제: 2

답: 3

검색페이지: 10

페이지 순위: 3

Topic : mathematical formular

**Predictions** of vapor pressures of aqueous desiccants for cooling applications by using artificial **neural networks**

This paper presents a new **approach** based on artificial **neural networks** (ANNs) to determine the vapor pressure of three widely used inorganic desiccant solutions, namely, calcium chloride, lithium chloride, and lithium bromide. The vapor pressure of liquid desiccants depends on temperature and concentration. Empirical expressions generally provide vapor pressure with limited accuracy. Further, the expressions currently in use are tedious and valid for narrow ranges and must be adjusted constantly. In this paper **neural networks** were trained to predict vapor pressure of desiccant solutions with a reasonable accuracy without mathematical formulae. Trained **neural network** models provided wide ranges of vapor pressure for desiccant solutions without the need to cross reference several tables or charts. Results showed potential of using ANNs for the **prediction** of vapor pressure of desiccant solution for cooling applications.

정확성 판단: 1

6명 TP의 UTT를 수행한 결과는 [별첨 A]에 주어졌다.

## 제4장      OntoFrame과      Scholar.Google      기능적 효용성 평가를 위한 사용자 테스트 결과 분석

### 4.1 사용자 테스트 결과 분석 방법

사용자 테스트 결과는 아래와 같은 방식으로 평가된다.

- ① UTT별      소요시간:      각각      과제별로      OntoFrame과  
Scholar.Google를      이용하여      주어진      과제를      해결하는데      소요된  
시간
- ② OntoFrame과 Scholar.Google에서      테스트      사용자가      UTT      수행  
결과로      제시한      정보      개수:      과제별      5개의      정보를      제시하라고  
하였을      때,      사용자가      과제      수행      결과로      제시한      정보의      개수
- ③ 리스팅      순위      :      과제      수행을      통하여      테스트      사용자가      제시한  
정보가      각각      시스템에서      리스팅      된      순위
- ④ 제시된      정보의      정확성:      테스트      사용자가      제시한      결과      정보      중  
과제      수행      결과로서      정확하다고      판단되는      정보의      비율
- ⑤ 제시된      정보의      일치성:      각      UTT      별      양      시스템에서      과제      수행      후  
제시된      정보      중      중복되는      정보를      제거한      후      남은      정보의      수.

#### 가. UTT별 소요시간

OntoFrame과 Scholar.Google      시스템      검색속도는      UTT      소요시간에  
영향을      줄      수      없을      정도의      미미한      차이를      보인다.      UTT별      소요시간은  
TP가      각각      UTT를      수행함에      있어서      5개의      결과값을      제시하거나,  
최종적으로      포기할      때까지      걸린      시간을      의미한다.      TP는      대부분의

소요시간을 시스템이 검색하여 제공한 정보들을 탐색하며, 정확한 정보를 찾아내는데 소요한다. 따라서 UTT별 소요시간이 짧다는 것은 시스템이 검색하여 제공한 정보 및 리스팅 순위가 정확하다는 것을 의미한다. 다시 말하여, 소요시간이 짧을 경우 검색 엔진 기능의 효용성이 상대적으로 뛰어나다는 것을 의미할 수 있다. UTT별 소요시간은 상대적 비율로 평가된다.

#### **나. 검색 결과로 제시된 정보의 개수**

검색결과로 제시된 정보의 개수는 Site Goal과 User Goal의 매칭 여부를 평가하는데 이용된다. 정확한 정보를 제공하려는 시스템의 Site Goal과 필요로 하는 정보를 찾기를 원하는 User Goal이 일치될 때 시스템 기능의 효용성이 극대화 될 수 있기 때문이다. 따라서 검색 결과로 제시된 정보의 개수가 요구된 바를 충족시키지 못하였을 때, UTT를 수행하려던 User Goal에 달성하지 못하였음을 의미한다. 요구된 정보 개수 vs. 제시된 정보 개수는 비율로 평가된다.

#### **다. 리스팅 순위**

시스템의 Site Goal은 정확한 정보를 사용자에게 제공하는데 있다. 따라서 대부분의 정보서비스 시스템은 사용자 질의에 정확하다고 평가되는 정보를 높은 순위로 하여 사용자에게 제공한다. 사용자는 경험으로부터 리스팅 순위가 높을수록 검색 질의에 정확한 정보라고 가정한다. 제시되는 정보의 리스팅 순위가 정확할수록, 사용자의 검색 소요시간은 짧아지고, 사용자 만족도는 높아지며, 검색 결과 또한 정확한 것으로 나타난다. 네이버에서 조사한 통계자료에 따르면, 웹 사용자는

68%가 첫 번째 검색결과페이지만을 참고하며, 세 번째 페이지까지 탐색하는 사용자는 92%로 조사되고 있다. 즉, 정확한 정보를 검색하여 제공한다 하더라도, 리스팅 순위가 부정확 할 경우, 사용자는 해당 정보를 발견하지 못한다는 것을 의미한다.

리스팅 순위는 소요시간 및 제시된 정보의 정확성, 제시된 정보의 일치성에 크게 영향을 미치는 것으로 나타난다. UTT 1~3번에 대하여, TP가 제시한 정보가 차지하는 리스팅 순위를 평가하고, 평균 순위를 구하여 상대적 비율로 평가한다.

#### **라. 제시된 정보의 정확성**

TP는 각각 UTT에 대하여 5개의 정확하다고 여겨지는 정보를 찾아 제시하여야 한다. 위 제시된 정보의 개수와 마찬가지로, 검색결과로 제시된 정보의 정확성은 Site Goal과 User Goal의 매칭 여부를 평가하는데 이용된다. 정확한 정보를 제공하려는 시스템의 Site Goal과 필요로 하는 정보를 찾기를 원하는 User Goal이 일치될 때 시스템 기능의 효용성이 극대화 될 수 있기 때문이다. 즉, 시스템 기능의 효용성이 뛰어들수록 제시된 정보의 정확성이 높다. 요구된 정보의 개수 vs. 정확하다고 판단된 정보의 개수를 비율로 평가한다.

#### **마. 제시된 정보의 일치성**

‘검색엔진 별 제시된 정보의 일치성’은 6명의 테스트 사용자가 UTT N에 대하여 OntoFrame과 Scholar.Google에서 작업 결과로 제시한 검색결과가 얼마나 일치하는지를 평가한다. 예로 UTT 1(N)에 대하여 6명의 테스트 사용자가 OntoFrame을 이용하여 5개씩 총 30개의

검색결과(SI)를 제시하였을 때, 중복되는 검색결과를 제외하게 되면, 남게 되는 검색정보(RI)의 개수는 검색결과의 정확성을 의미한다고 할 수 있다. 즉, 남게 되는 정보의 개수가 적을수록 사용자가 원하는 정보를 빠르게 검색할 수 있어, 정보를 탐색(Navigation)하는 시간이 작다는 것을 의미한다. 반대로 남게 되는 정보의 개수가 많을수록, 사용자가 여러 페이지에 걸쳐 정보를 탐색하고, 이에 따라 정보의 정확성 결여 및 소요 시간이 길어짐을 의미할 수 있다. 제시된 정보의 일치성 비율을 상대적으로 평가한다.

## 4.2 사용자 테스트 결과 분석

### 가. UTT 수행 소요시간

[표 13] UTT 별 소요 시간 (분)

UTT	UTT1		UTT2		UTT3		UTT4		UTT5		총계	
	O	G	O	G	O	G	O	G	O	G	O	G
TP 1	4	7	3	6	7	8	5	13	1	16	20	50
TP 2	10	9	3	7	2	3	3	4	1	9	19	32
TP 3	6	4	6	9	3	7	5	4	1	14	21	38
TP 4	6	7	17	13	7	10	6	5	2	6	38	41
TP 5	6	11	4	11	12	9	6	16	8	18	36	65
TP 6	7	7	3	12	3	6	4	6	1	11	18	42
합계	39	45	36	58	34	43	29	48	14	74	152	268





TP 5	5.04	62	4.02	3.08	6.2	4.75	15.26	69.83
TP 6	15	3.08	1	25.75	3.6	5.8	19.6	34.63
합계	58.64	125.67	42.22	89.5	23.6	42.2	124.46	257.67
평가(%)	9.77	20.95	7.04	14.92	3.93	7.03	257.67/124.46X100=207.03%	

정보서비스 시스템에서 검색된 정보의 리스팅 순위는 중요한 의미를 갖는다. 검색 포털인 네이버에서 조사한 자료에 따르면 사용자의 68%는 첫번째 검색결과 페이지 만을 참고하여 정보를 검색하며, 세 번째 페이지까지 탐색하는 사용자는 92%로 조사되었다. 이러한 연구 결과는 본 사용자 테스트를 위한 사전 인터뷰에서도 같은 결과를 보였다.

UTT 1 경우 수행결과로 제시된 5 개 정보가 OntoFrame 과 Scholar.Google 에서 리스팅된 순위를 검사하여 평균을 구하였다. UTT2 와 UTT3 의 경우 사용자가 제시한 주제(토픽)와 주어진 검색어(Jinde Cao)를 키워드로 하여 재검색하였을 때, 각 문항별 검색된 정보의 리스팅 순위를 찾아 평균 값을 구하였다. 리스팅 순위 값이 높을수록 정보의 위치가 뒤에 있음을 의미하며, 검색결과 페이지 중 10 페이지를 넘어갈 경우 리스팅 순위를 100 으로 처리하였다.

OntoFrame 에서 TP 는 시스템이 검색결과로 제시한 첫번째 페이지에서 대부분의 정보를 발견하는 것으로 나타나고 있으며, Scholar.Google 에서는 평균적으로 두 번째 페이지까지 탐색을 통하여 정보를 얻는 것으로 평가된다. 즉 Scholar.Google 에서 정보를 찾아 제시하기 위하여, TP 는 OntoFrame 에서 검색 대비 2 배 이상의 정보를 탐색하여야 함을 의미 한다.

OntoFrame 에서 사용자가 검색 제시한 정보의 리스팅 순위가 Scholar.Google 대비 207.03% 우수한 것으로 평가되었다.

### 라. 제시된 정보의 정확성

[표 16] 제시된 정보의 정확성

UTT	UTT1		UTT2		UTT3		UTT4		UTT5		총계	
	O	G	O	G	O	G	O	G	O	G	O	G
TP 1	5/5	4/5	5/5	5/5	5/5	4/5	5/5	5/5	5/5	5/5	25/25	23/25
TP 2	5/5	4/5	5/5	5/5	5/5	2/2	5/5	4/4	5/5	5/5	25/25	20/21(25)
TP 3	5/5	3/5	5/5	5/5	5/5	5/5	5/5	5/5	5/5	0	25/25	18/20(25)
TP 4	5/5	3/5	5/5	3/3	5/5	5/5	5/5	5/5	5/5	5/5	25/25	21/23(25)
TP 5	5/5	4/5	5/5	5/5	5/5	4/4	5/5	2/2	5/5	5/5	25/25	20/21(25)
TP 6	4/5	4/5	5/5	4/5	5/5	5/5	5/5	5/5	5/5	0	24/25	18/20(25)
합계	29/30	22/30	30/30	27/30	30/30	25/25	30/30	26/26	30/30	20/20	149/150	120/130 (150)
평균(%)	99.33/80X100=124.16											80

TP 는 5 개 UTT 에 대하여 각각 5 개씩, OntoFrame 과 Scholar.Google 에서 정보를 찾아 각 25 개씩 정보를 제시하여야 한다. 따라서 6 명의 TP 는 양 시스템에 대하여 각각 총 150 개의 정보를 대상으로 정확성 여부를 평가하였다.

OntoFrame 의 경우 총 제시된 정보 개수는 150 개이며, 그 중 149 개의 정보가 정확한 것으로 판단되었다. Scholar.Google 에서 검색되어 제시된

정보는 총 130 개이며 정확한 것으로 평가된 정보는 120 개 이다. 따라서 OntoFrame 에서 제시된 정보의 정확률은 99.33%이며, Scholar.Google 에서 제시된 정보의 정확률은 80%이다. OntoFrame 에서 사용자가 검색 제시한 정보의 정확성이 Scholar.Google 대비 124.16% 우수한 것으로 평가되었다.

### 마. 제시된 정보의 일치성

[표 17] 제시된 정보의 일치성

UTT	UTT1		UTT2		UTT3		UTT4		UTT5		총계	
	O	G	O	G	O	G	O	G	O	G	O	G
중복된 정보를 제외 정보 수	15	14	24	26	12	20	8	13	5	19	64	92
검색된 정보 중 정확한 정보 수	29	22	30	27	30	25	30	26	30	20	149	120
불일치율(%)	51.72	63.63	80.96	30.40	80	26.67	50	16.67	95	42.95	76.67	
검색된 정보의 일치율(%)											57.05	23.33
	$57.05/23.33 \times 100 = 244.53$											

TP 가 UTT 에 대하여 제시한 정보들의 중복성이 높을수록, TP 는 검색된 페이지에서 정보 탐색을 효율적으로 수행하였다는 것을 의미한다. 다시 말하여, 제시된 정보의 일치성이 높을수록 UTT 수행시간이 짧고, 제시된 정보의 정확성이 높아짐을 의미할 수 있다. 다른 한편으로 제시된 정보의 일치성이 높다는 것은 정보서비스 시스템의 검색결과 리스팅 순위가 정확하다는 것을 의미할 수 있다.

OntoFrame 에서 6 명의 TP 가 5 개 UTT 를 수행한 결과로 제시한 정확하다고 판단되는 정보 중 중복되는 정보를 삭제하였을 때 남은

정보의 개수는 64 개 이다. 즉 총 149 개 제시된 정보 중 57.07%의 정보가 중복되었다. Scholar.Google 에서는 총 120 개의 정보 중 중복된 정보를 제외한 정보수가 92 개로, 검색된 정보의 일치율은 23.33%에 불과하였다. 위 결과는 평가항목 1 의 소요시간 및 평가항목 3 의 제시된 정보의 리스팅 순위와 비교할 때 일치되는 결과 임을 알 수 있다.

검색된 정보의 일치율은 OntoFrame 이 Scholar.Google 대비 244.53% 높게 평가되었다.

## 제5장 OntoFrame과 Scholar.Google 사용자 테스트를 통한 기능적 효용성 평가 결과 정리

OntoFrame과 Scholar.Google 사용자 테스트를 통한 기능적 효용성 비교 평가 결과는 Semantic functionality와 같은 고급 기능을 가진 OntoFrame이 시스템을 이용하여 얻고자 하는 User Goal을 보다 효율적으로 지원한다는 것을 보여주었다.

OntoFrame은 Scholar.Google에 비하여, 넓은 범주의 Site Goal을 가지고 있는 것으로 나타났다. OntoFrame의 Site Goal은 Semantic functionality와 같은 기능에 의하여 확장된 것으로 나타났으며, 이를 통하여 정보를 필요로 하는 User Goal과의 간격을 메우는 것으로 평가된다.

소요 시간 평가지표 분석 결과를 살펴보면, UTT 1 처럼, Site Goal 과 User Goal 이 일치하고, UTT 의 Level I 에 해당하는 과제 수행에서, 소요시간은 OntoFrame 과 Scholar.Google 양 시스템에서 큰 차이를 보이지 않는다. 그러나 Site Goal 과 User Goal 이 일치하지 않고, UTT 수준이 Level III 으로 높은 경우 OntoFrame 과 Scholar.Google 에서 UTT 수행시간은 TP 에 따라 최대 16 배 차이를 보이며, 평균 6 배의 소요시간 차이를 보이고 있다. 이러한 결과는 시스템의 기능 효용성이 OntoFrame 에서 우수하게 나타남을 의미한다. OntoFrame 의 소요시간 효용성이 Scholar.Google 대비 176.31% 우수한 것으로 평가되었다.

제시된 정보의 개수 평가지표 분석 결과를 보면, 테스트 사용자는 OntoFrame 에서 요구된 정보를 모두 제시하고 있는데 반하여,

Scholar.Google 에서 사용자는 목표대비 86%의 정보 만을 제시하고 있다. 이러한 결과는, User Goal 을 시스템 기능이 달성할 수 있도록 지원하지 못하고, 또는 사용자의 선지식으로부터 정보를 생성하지 못할 때 인 것으로 평가된다. OntoFrame 을 이용한 정보검색의 효용성이 제시된 정보의 개수 측면에서 Scholar.Google 대비 116.28% 우수한 것으로 평가되며, 100% vs. 86%로 계산할 경우 OntoFrame 이 Scholar.Google 대비  $100/86 \times 100 = 116.28\%$  높은 효용성을 보이는 것으로 평가되었다.

정보서비스 시스템에서 검색된 정보의 리스팅 순위는 중요한 의미를 갖는다. 검색 포털인 네이버에서 조사한 자료에 따르면 사용자의 68%는 첫번째 검색결과 페이지 만을 참고하여 정보를 검색하며, 세 번째 페이지까지 탐색하는 사용자는 92%로 조사되었다. 이러한 연구 결과는 본 사용자 테스트를 위한 사전 인터뷰에서도 같은 결과를 보였다. OntoFrame 에서 TP 는 시스템이 검색결과로 제시한 첫번째 페이지에서 대부분의 정보를 발견하는 것으로 나타나고 있으며, Scholar.Google 에서는 평균적으로 두 번째 페이지까지 탐색을 통하여 정보를 얻는 것으로 평가된다. 즉 Scholar.Google 에서 정보를 찾아 제시하기 위하여, TP 는 OntoFrame 에서 검색 대비 2 배 이상의 정보를 탐색하여야 함을 의미 한다. OntoFrame 에서 사용자가 검색 제시한 정보의 리스팅 순위가 Scholar.Google 대비 207.03% 우수한 것으로 평가되었다.

OntoFrame 의 경우 총 제시된 정보 개수는 150 개이며, 그 중 149 개의 정보가 정확한 것으로 판단되었다. Scholar.Google 에서 검색되어 제시된 정보는 총 130 개이며 정확한 것으로 평가된 정보는 120 개 이다. 따라서 OntoFrame 에서 제시된 정보의 정확률은 99.33%이며, Scholar.Google 에서 제시된 정보의 정확률은 80%이다. OntoFrame 에서 사용자가 검색

제시한 정보의 정확성이 Scholar.Google 대비 124.16% 우수한 것으로 평가되었다.

OntoFrame 에서 6 명의 TP 가 5 개 UTT 를 수행한 결과로 제시한 정확하다고 판단되는 정보 중 중복되는 정보를 삭제하였을 때 남은 정보의 개수는 64 개 이다. 즉 총 149 개 제시된 정보 중 57.07%의 정보가 중복되었다. Scholar.Google 에서는 총 120 개의 정보 중 중복된 정보를 제외한 정보수가 92 개로, 검색된 정보의 일치율은 23.33%에 불과하였다. 제시된 정보의 일치성이 높을수록 UTT 수행시간이 짧고, 제시된 정보의 정확성이 높아짐을 의미할 수 있다. 다른 한편으로 제시된 정보의 일치성이 높다는 것은 정보서비스 시스템의 검색결과 리스팅 순위가 정확하다는 것을 의미할 수 있다. 검색된 정보의 일치율은 OntoFrame 이 Scholar.Google 대비 244.53% 높게 평가되었다.

위 테스트 분석 결과는 소프트웨어의 기능이 “Software quality in use”를 결정할 수 있음을 충분히 증명하고 있다. 소프트웨어의 고급 기능은 사용자의 작업 부담을 줄여주며, 작업의 효율성을 높여주는 것으로 연구결과는 증명하고 있다. 특히 시스템이 최초에 제공하는 검색 결과 순위의 정확성이, 소요시간, 검색 제시된 정보의 개수, 제시된 정보의 정확성, 제시된 정보의 일치성 등과 밀접하게 연관되어 있음이 밝혀졌다.