

# 비즈니스용 웹 사이트 개발에서의 실시간 발성사고법 사용성 테스트 사례 연구

## Research on Usability Test for Business Web Site Development using Concurrent Think Aloud Protocol

김재훈, 신재민, 김정환, 임석중  
한국과학기술정보연구원

Jayhoon Kim(jay.kim@kisti.re.kr), Jaemin Shin(jmshin@kisti.re.kr),  
Jeong Hwan Kim(kimjh@kisti.re.kr), Seok Jong Lim(seoklim@kisti.re.kr)

### 요약

본 연구는 비즈니스용 웹 사이트 개발에 있어 비용효과적인 사용성 테스트 모델을 제시할 목적으로 수행되었다. 본 연구에서는 실시간 발성사고법(Concurrent Think Aloud), 성능 측정, 코칭 방법, 원격 테스트 방법을 혼합한 사용성 테스트 모델을 제시하였다. 테스트 대상은 연구진이 운영하는 전자정보 공동구매 컨소시엄 웹 사이트로 하였다. 테스트 참가자는 실제 사용자로 하였으며 업무 역할과 숙련도별로 차이를 두어 구성하였다. 사용성 테스트 분석 결과를 반영하여 시스템을 개선하였고, 동일한 태스크를 수행하여 소요시간과 마우스 궤적 복잡도를 비교한 결과 모든 참가자의 사용성이 개선된 것을 확인하였다. 본 연구에서 제시한 사용성 테스트 모델을 활용하여 비즈니스용 웹 사이트를 비용 효과적으로 평가하여 분석하는데 도움이 될 것으로 기대한다. 웹 사이트들은 목적별, 분야별로 특성이 있으므로 향후에는 특성화된 사용성 테스트 모델 연구가 필요하다고 본다.

■ 중심어 : | 사용성 테스트 | 실시간 발성사고법 | 마우스 트래킹 |

### Abstract

The purpose of this study is to provide a cost effective usability test model for business web site development. In this study, we present a usability test model that combines Concurrent Think Aloud, performance measurement, coaching method, and remote test method. Testing target was the consortium website of electronic information joint licensing. Test participants were actual users selected by their task roles and proficiency in different levels. We improved the system by reflecting the results of user test analysis and re-tested usability test again. The result showed that usability of all participants was improved in the aspect of the time required and the complexity of mouse trajectory. It is expected that the usability test model presented in this study can be used to cost-effectively evaluate and analyze business web sites. We suggest that research various usability test model considering the characteristics of websites.

■ keyword : | Usability | User Test | Think Aloud | Mouse Tracking |

\* 본 연구는 2017년도 한국과학기술정보연구원(KISTI)의 주요사업 과제로 수행한 것입니다.

접수일자 : 2018년 03월 08일

심사완료일 : 2018년 04월 09일

수정일자 : 2018년 04월 09일

교신저자 : 신재민, e-mail : jmshin@kisti.re.kr

## I. 서론

### 1. 배경 및 목적

비즈니스 관점에서 웹 사용성의 효과는 수익 증가나 비용 절감 등으로 설명될 수 있어야 한다. 이러한 관점에서 웹 사용성의 수익성(ROI, Return On Investment) 분석 연구 사례를 몇 가지 들어본다. 2003년 Nielsen[1]이 수행한 웹 사용성의 수익성 연구에 따르면 웹 사용성을 개선한 42개 사례에서 평균 판매량은 100% 증가, 사용자 효율/생산성은 160% 증가한 것으로 나타났다. Weinschenk[2]는 웹 사용성 기법을 통해 개발기간을 33~50%까지 단축할 수 있으며, 고객지원, 직원훈련, 매뉴얼 관리, 시스템 유지관리 등 비용 절감 관점에서 웹 사용성의 효용성을 주장하였다. 일례로 아메리칸 에어라인 항공사가 개발 단계에서 사용성 문제를 해결함으로써 비용을 약 60~90% 절감한 사례를 들었다. 하지만 아직도 웹 사용성 평가는 웹 개발의 필수 절차라는 자리 잡지 못한 듯하다. 전문 업체를 활용한 사용성 평가에는 많은 비용이 소요되고 활용되는 장비나 소프트웨어도 대부분 고가이다. 사용성 평가에 소요되는 시간 또한 상당하다. 사용성 개선으로 인한 수익 발생 또는 효용 가치를 경영진에게 명확히 인식시키지 못한다면 웹 사이트 개발 비용에 사용성 평가비용을 포함하기는 어렵다.

본 연구는 웹 사용성 평가에 관한 한 논문 “Is the Mouse a ‘Poor Man’s Eye Tracker?’”으로 부터 아이디어를 얻어 전문적인 웹 사용성 전문가나 전문 업체를 활용하지 않고서도 비용효과적인 사용성 평가를 해보고자 시도되었다. 웹 사이트 운영진이 직접 사용성 평가를 진행해 봄으로써 최소한의 비용과 시간을 투입하여 얻을 수 있는 사용성 테스트 결과와 수준, 한계를 알아보고자 하였다.

### 2. 대상 및 방법

본 연구에서는 전자정보 국가 컨소시엄 KESLI 시스템을 대상으로 사용성 테스트를 수행하였다. KESLI는 학술연구 분야 전자정보 콘텐츠를 출판사와 도서관이 거래할 수 있도록 중계하는 B2B 마켓플레이스 시스템

이다. KESLI와 같은 도서관 전자정보 컨소시엄 전 세계적으로 100개 이상 존재하지만 표준화된 컨소시엄 시스템 모델이 제시되지 않고 있으며 상용 시스템도 존재하지 않아 KESLI 시스템을 자체 개발하여 운영하고 있다. 상품 구성이 다양하고 가격 모델이 모두 달라 시스템 표준화가 어려워 시장에 참여하는 출판사와 도서관의 담당자들도 시스템 사용이 어려움을 겪고 있다.

업무 처리량도 상당하다. 판매자 그룹에서는 KESLI 시스템을 통해 컨소시엄으로 판매할 품목의 제안서를 매년 등록하고 참가신청을 접수하여 계약한다. 큰 규모의 업체의 경우 33개 상품 제안서를 등록하는데 전자정보 컨소시엄 관련 전문 용어를 이해하여야 하고 상품설명, 제공콘텐츠리스트, 협상자료, 계약서류 등 기입할 항목이 최소 78개에 달한다. 제안서 등록완료까지 10회 이상 오류 수정을 반복하는 경우도 있다. 제안서 접수와 조건 협상이 완료된 이후에는 판매자는 구매자로부터 품목별로 참가신청을 받는데 2016년에는 업체 평균 21개의 구매 계약을 하였으며 큰 규모의 업체는 약 1,300여의 계약을 체결하였다.

이와 같은 사용성 문제를 개선하고 서비스를 확장하기 위해 KESLI 시스템을 재개발하게 되었으나 비용 문제로 웹 사용성 평가 전문가나 전문 업체를 활용할 수 없는 상황이었고 이로 인해 내부 운영진이 무료로 활용할 수 있는 장비와 소프트웨어를 활용하여 직접 사용자 테스트를 수행하고 시스템 개발에 반영한 후 그 결과를 검증하는 시도를 하게 되었다.

사용성 평가의 방법으로는 사용자 테스트 방법을 선택하여 사용자들이 시스템 사용 중에 겪는 문제점과 어려움을 찾아내고자 하였다. 이중에서도 실시간 발성사고법(Concurrent Think Aloud)을 활용하였다. 실험 대상자는 실제 KESLI 시스템을 사용하는 출판사와 도서관 2개 부문에 담당자 각 2명씩 선정하였으며 부문별로 고빈도 사용자와 저빈도 사용자로 구성하였다. 개발 진행 단계 및 개발 완료 단계에서 사용자 인터페이스의 개선 정도를 확인하기 위해 소요 시간과 마우스 궤적을 측정하여 비교하였다.

## II. 이론적 배경

사용성 평가 방법은 평가 방식, 주체, 도구 등에 따라 다양한 방법이 존재한다. 전문가들도 각자의 관점에 따라 여러 기준으로 사용성 평가 방법들을 분류하고 있다. Jacob Neilson은 1995년 그의 저서 "Usability Inspection Methods"[3]에서 Heuristic evaluation, Cognitive walkthrough, Pluralistic walkthrough 등 8가지 방법으로 구분하였다. Hom[4]은 1996년 "The Usability Methods Toolbox" 웹 사이트에 사용성 평가 방법을 질의(Inquiry), 검사(Inspection), 테스트(Testing)의 세 가지로 구분하여 각 방법을 설명하였다. 국내의 연구에도 2005년 김정룡[5] 등이 사용성 평가 방법을 평가 주체에 따라 평가자, 사용자, 개발자 또는 둘 이상의 평가 주체들을 조합하는 방식으로 구분한 연구가 있다. Fernandez[6] 등은 웹 사용성 평가에 관한 206편의 논문을 분석한 연구에서 사용성 평가에 대한 원천분야, 사용성 평가에 대한 정의, 채택한 사용성 평가 방법, 평가방법 유형, 사용성 평가방법을 적용한 웹 산출물 및 시기, 사용성 평가 피드백 수준, 사용성 평가 실제 검증 여부 등으로 구분하였다.

본 연구에서 사용한 발생사고법(Think Aloud) 방법은 실험 대상자를 관찰하는 사용성 테스트(usability testing) 방법에 해당한다. 발생사고법은 실시간 발생사고법과 회고적 발생사고법으로 구분된다. Haak[7]은 도서관 사용자를 대상으로 하는 사용성 평가에서 실시간 발생사고법으로는 주로 관찰에 의해 문제를 검출하였다고 한다. 반면 회고적 발생사고법을 통해서 관찰로는 검출할 수 없는 말로 표현되는 문제를 더욱 검출할 수 있었다고 한다. 실시간 발생사고법은 실시간 태스크 수행에 영향을 주기 때문에 중간 정도 난이도로 태스크를 개발하여야 한다고 연구방법을 제시하였다. Teague[8] 등도 실시간 발생사고법과 회고적 발생사고법 비교연구를 하였는데 실시간 발생사고법 실험 속도 참가자들의 테스트 속도가 15% 정도 느린 것으로 나타났다. 대신 실시간 발생사고법 참가자들이 테스트 중간 점검 시간에 평가에 대한 설명을 더 많이 하고 웹 사이트에 대한 평가의견을 더 많이 제시하였다고 한다. 실

시간 발생사고법을 성공적으로 수행하려면 중간 점검 시간을 적절히 배치하는 것이 중요하다고 하였다.

Cooke[9]는 실시간 발생사고법의 정보 손실로 인한 한계를 지적하였다. 실제 생각을 말로 표현하는 비율은 5% 정도로 많은 정보가 누락되어 사용성 평가에 부족함이 있으며 이를 보완하기 위해서는 안구 운동 등 부가 정보를 활용하여 피실험자를 관찰해야 한다고 하였다. 이에 대해 고가의 아이트래킹 실험 장비를 대체하여 마우스 움직임으로 안구 운동 관찰을 대체할 수 있는지 모색한 여러 연구가 있다. Chen[10] 등은 2001년에 웹 사이트를 크게 7개의 영역으로 나누고 시선과 마우스가 체류하는 시간을 비교하여 상관관계가 있음을 밝혔다. Cooke[11]는 2006년 실험을 통해 시선과 마우스의 체류이 69% 일치함을 밝혀냈다. Yahoo의 연구원 Navalpakkam과 Churchill[12]는 광고로 인한 웹 사이트 읽기 방해 정도를 측정함으로써 시선과 마우스 체류 시간이 상관관계가 있음을 밝혔다.

사용성 평가에 있어 특정 연구방법으로 인한 편향성을 제거하기 위하여 두 가지 이상의 연구방법을 채택하는 여러 사례가 있다. 2003년 Yang[13]은 개인적 편향, 가능성, 해석의 문제 인지 등 연구과정에서 고의적 조정으로 인한 잠재적 문제 발생을 방지하기 위해 삼각측량 방법을 제안하였으며 구체적으로는 발생사고법과 마우스 추적 방법론을 병행한 사용성 평가 실험을 하였다. 국내에서는 김병주와 이건표[14]가 2007년 사용자의 시선과 마우스의 움직임, 생각을 기록하는 웹 사용성 통합 평가모형을 제안하였다. 김영준과 김영진[15]은 2013년 안구 운동을 관측하여 웹 사이트 탐색경로와 호감도를 측정하고 사후 인터뷰를 통해 여러 측정치들 간의 관련성을 분석하였다.

운영자로서 웹 사이트의 사용성을 평가하여 개선한 사례도 있다. 2001년 McMullen[16] 도서관 웹사이트 재개발 과정에서 단계별로 인터뷰, 발생사고법, 인지적 시찰법, 발견적 평가 법을 적용하여 사용성 평가를 수행하고 개선한 사례가 있다. 국내에서는 유진환 등[17]이 2007년 국내 R&D 정보 네비게이션의 사용자 경험 개선을 위해 NTIS-NDSL 연계서비스인 R&D Navigator 사용성 평가를 수행한 연구가 있다.

선행 연구와 비교할 때 본 연구는 다음과 같은 차별성이 있다. 첫째, 사용성 테스트 방법은 다양하지만 비즈니스 분야라는 도메인 중심의 사용성 테스트 사례로서 본 연구는 차별화 된 연구이다. 둘째, 웹 사이트 개발 및 운영자 관점에서 태스크 시나리오를 개발하고 세부적으로 분석하였다.

### III. 연구 방법

#### 1. 연구 설계

본 연구에서는 KESLI 시스템 개발 과정에 실제 사용자를 참여시켜 태스크 수행과정을 실시간 발생사고법으로 사용성 문제점을 파악하고 사용성을 개선하였다. 사용성 개선 정도를 확인하고자 시스템 개발 단계와 시스템 개발 완료 후 동일한 태스크로 테스트를 실시하였으며 각각의 마우스 궤적과 소요 시간을 기록하여 비교하였다.

본 연구의 테스트 대상 웹 사이트는 비즈니스 중심인 웹 사이트로 분류할 수 있다. 본 연구방법 설계에 있어서도 비즈니스 중심 웹 사이트의 특성을 반영하는데 중점을 두었다. 시스템에서 사용자가 처리할 업무를 명확히 인지하고 정해진 프로세스에 따라 오류 없이 신속히 업무 처리할 수 있는지 테스트하도록 태스크를 개발하고 적합한 참가자를 선정하였다.

본 연구에서는 시스템 개발 단계와 개발 완료 후 두 번의 테스트를 실시하였다. 개발 완료 후 테스트는 사용성 개선 정도에 대한 양적 측정을 목적으로 하였고 2명은 현장 테스트로 현장 참가가 불가한 2명은 원격 테스트로 진행하였다.

사용성 테스트는 발생사고법을 중심으로 하였다. 발생사고법에는 실시간, 회고적 두 가지 방법이 있는데 본 연구에서는 실시간 발생사고법을 수행하였다. 선행 연구들에 따르면 실시간 발생사고법은 태스크 수행과 병행이 어렵고 시간이 더 많이 소요되며 즉각적인 반응 중심이라서 단독으로 사용성 테스트하기에는 어려움이 있다는 견해가 있다. 그럼에도 연구진은 실시간 발생사고법을 채택하였다. 기존 시스템에 대한 개선 요구 내

용은 시스템 운영 과정에서 진화, 이메일 및 설문 등으로 이미 수집하여 파악된 상태였기 때문이다. 연구진은 실시간 발생사고법이 참가자들이 신규 시스템에서 태스크를 수행하며 맞닥뜨리는 즉각적 반응을 확인함으로써 그간 사용자들의 개선 요구사항이 정확히 파악되었는지 신규 시스템에 반영되고 사용성 문제를 개선하였는지 확인하는데 더욱 효과적일 것으로 판단하였다.

본 연구에서 마우스 궤적을 기록하는 무료 소프트웨어 IOGraph를 사용하여 개발 단계와 개발 완료 후 시스템의 사용성 개선 정도를 비교하였다. 태스크 수행 과정의 마우스 궤적을 기록하고 도식화하여 비교하는 수준으로는 유용하였다.

본 연구의 대상 시스템은 비즈니스 중심 웹 사이트로 사용자와 기능의 특수성으로 인하여 설명 없이 테스트 진행이 어려운 점이 있고 연구의 목적 또한 참가자의 사용성 의견 수집에 중점을 둔 실증적 연구였기 때문에 연구진이 개입하는 코칭 방식으로 테스트를 수행하였다.

표 1. 테스트에 사용된 웹 사이트 및 방법론 구분

구분	유형
웹 사이트 유형	<ul style="list-style-type: none"> <li>•정보 중심 <input type="checkbox"/></li> <li>•서비스 중심 <input type="checkbox"/></li> <li>•비즈니스 중심 <input checked="" type="checkbox"/></li> </ul>
제품 개발 단계	<ul style="list-style-type: none"> <li>•제품 개발 전 단계 <input type="checkbox"/></li> <li>•제품 개발 단계 <input checked="" type="checkbox"/></li> <li>•제품 개발 완료 단계 <input checked="" type="checkbox"/></li> </ul>
사용성 테스트 방법	<ul style="list-style-type: none"> <li>•발생사고법 <input checked="" type="checkbox"/></li> <li>•성능 측정 <input checked="" type="checkbox"/></li> <li>•코칭 방법 <input checked="" type="checkbox"/></li> <li>•원격 테스트 <input checked="" type="checkbox"/></li> <li>•공동 발견 방법 <input type="checkbox"/></li> </ul>

Aamir와 Mansoor[18]의 웹 애플리케이션 사용성 테스트 방법론에 따라 본 연구에서 채택한 사용성 테스트 방법론을 정리하면 [표 1]과 같다.

#### 2. 테스트 참가자 선정

테스트 참가자는 전자정보를 공급하는 업체 담당자 2명(판매자 그룹), 전자정보를 구매하는 도서관 사서 2명(구매자 그룹)으로 총 4명으로 구성하였다. 각 그룹 내에서도 고빈도 사용자와 저빈도 사용자를 구분하여

구성하였다. 테스트 참가자들에게는 연구의 목적, 절차, 방법, 신분과 비밀보장, 동영상 촬영 및 녹취함을 알리고 서면으로 동의를 받았다.

### 3. 태스크 개발

태스크는 KESLI 시스템의 핵심 기능 중심으로 구성하였다. 공급 업체에는 전자정보 품목을 제안하고 접수하는 기능 중심으로, 도서관측에는 제안된 전자정보를 확인하고 구매 신청하는 기능을 중심으로 구성하였다. 태스크의 개수는 전자정보를 공급 업체 측에는 3개, 전자정보를 구매하는 도서관측에는 2개로 하였다.

### 4. 실험 절차

사용자 테스트 전에 내부 직원을 통해 파일럿 테스트를 실시하였다. 연구진이 사용성 테스트 전문가는 아니고 처음 진행하는 상황이므로 테스트 환경 설정 및 진행 방법을 충분히 숙련할 필요가 있었다. 또한 태스크를 잘 이해할 수 있도록 태스크 지시서의 설명을 보완하고 쉬운 표현으로 정비하였다.

사용자 테스트는 공급 업체 참가자를 “판매자 그룹”, 도서관 사서 참가자를 “구매자 그룹”으로 명명하여 구분하고 실험하였다. 조용한 실험 공간에 노트북 컴퓨터를 설치하고 KESLI 시스템에 접속하여 1명씩 순차적으로 테스트를 진행하였다. 테스트 참가자에게는 과제를 수행하면서 웹 사이트 작동에 대한 예상과 기대, 실제 작동 결과에 대한 반응을 최대한 말로 표현하도록 지시하였다. 실험과정에 IOGraph 소프트웨어를 구동시켜 마우스 궤적을 기록하였다. 실험과정의 노트북 컴퓨터 화면 및 테스트 참가자의 음성을 동영상으로 캡처하였으며 동시에 DSLR 카메라로 현장의 동영상을 촬영하여 기록하였다. 태스크에 대해 이해가 부족하거나 진행에 어려움이 있는 경우 연구진이 테스트 참가자에게 최소한의 설명을 하였다. 테스트 종료 후에는 테스트 참가자와 연구진이 함께 테스트 동영상을 재확인 하면서 의도가 명확치 않거나 의도하지 않은 작동에 대한 내용을 확인을 하였다.

### 5. 데이터 처리

실험 종료 후에는 소요시간을 기록하고 테스트 참가자의 구술 내용을 엑셀 파일에 녹취하였다. 양적 분석을 위하여 단위 기능 수행 중 발생한 구술 내용을 단락으로 정하였다. 녹취록 완성 후 확인한 결과 하나의 단락은 최소 1개에서 최대 14개까지의 문장으로 구성되었다. 녹취 내용만으로 실험 상황을 알 수 없는 부분에는 괄호 표시를 하고 상황 설명을 기재하였다.

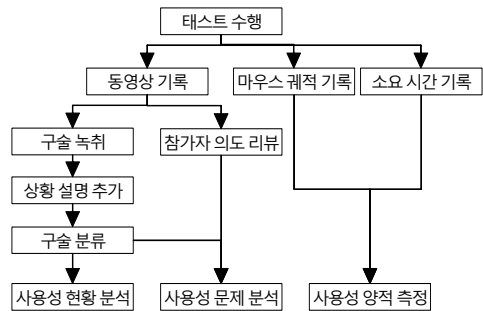


그림 1. 사용성 분석을 위한 데이터 준비 절차

정량적 분석을 위해 구술 내용을 7가지 유형으로 구분하여 코딩하였다. 선행 연구에서는 Cooke(2010)[7]가 코딩 분류를 5개 유형으로 구분한 구체적인 사례가 있어 참고하고자 하였다. 그러나 본 연구에는 사용성 문제점을 구체적으로 구분할 필요가 있어 그대로 적용할 수는 없었다. 선행연구에서의 “읽기”, “절차”, “관찰” 세 개 유형에 해당하는 내용은 “진행 과정” 하나로 단순화하였다. 대신 선행연구의 “설명”에 해당 하는 부분을 피 실험자의 관점에서 6개의 유형으로 세분화하였다. 이로써 웹 사이트의 각 요소(레이블, 기능, 도움말 등)중 어떤 관점에서 사용성을 개선할 것인지 방향을 수립하고 사용성 개선의 역할 분담(개발자, 디자이너, 운영자)에 활용하고자 하였다.

개발 중 시스템에 대한 테스트이지만 기존 시스템에서의 문제점이 해소된 부분도 있어서 “만족 의견” 구술 내용 유형을 포함하였다.

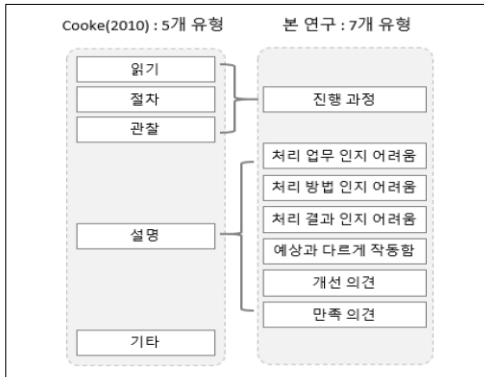


그림 2. 기술내용 유형 : 선행연구와의 비교

#### IV. 연구 결과

연구 설계에 따라 KESLI 시스템의 사용성 테스트를 시스템 개발 단계와 완성 단계에 2회 실시하였고 발생 사고법 테스트는 개발 단계에서만 수행하였다. 사용성 테스트 분석결과 확인된 문제점은 연구진과 시스템 개발자가 함께 논의하여 시스템 구현 방안을 결정하였다. 시스템 완성 후에는 사용성 개선 정도를 확인하기 위해 동일한 참가자를 대상으로 동일한 태스크를 수행하여 소요 시간과 마우스 궤적을 측정하여 비교하였다. 본 연구결과에는 연구방법에서 정한 사용성 테스트 및 사용성 개선 정도 비교 분석 결과만을 기술하였다.

##### 1. 발생사고법 테스트 결과 분석

###### 가. 참가자 개인 간의 차이

테스트 시작 단계에서 모든 참가자들이 발생사고법을 낯설어하며 잘 적용하지 못했다. 이로 인해 연구진은 태스크 단계별로 참가자의 반응을 묻거나 참가자들의 질문에 답변하는 등 약간의 코칭 방법을 사용하였다. 하지만 개인적 성향 차이를 극복하지는 못하였다. 구매자 그룹 2번 참가자의 경우 업무 경험이 많아 업무 프로세스를 잘 이해하고 있었기 때문에 새로 개발하는 시스템에서도 별다른 반응 없이 태스크를 잘 수행하였다. 하지만 해당 참가자의 기술 건수가 다른 참가자들의 절반 이하로 적어서 어떤 부분에서 사용성의 어려움

을 느끼는지 거의 확인할 수 없었다.

표 2. 개인별 발생사고 건수 비교

구분	구매자 그룹		판매자 그룹	
	참가자1	참가자2	참가자1	참가자2
진행과정 표현	10	6	2	5
처리 업무 인지 어려움	0	0	3	2
처리 방법 인지 어려움	4	3	1	4
처리 결과 인지 어려움	4	0	4	3
예상과 다르게 작동함	2	0	6	3
개선 의견	2	0	6	8
만족 의견	0	0	4	2
<b>합계</b>	<b>22</b>	<b>9</b>	<b>26</b>	<b>27</b>

###### 나. 참가자 그룹 간의 차이

구매자 그룹과 판매자 그룹간의 발생사고법 결과의 차이가 있는지 알아보고자 그룹별 기술 내용 단락 수를 비교하였다.

구매자 그룹과 판매자 그룹의 반응을 비교하면 전반적으로는 판매자 그룹이 기술 건수가 많았고 다양한 관점에서 문제점과 개선의견을 표현하였다. 이는 판매자 그룹이 컨소시엄 업무 처리 양이 많고 상품 특성상 업무처리 난이도가 높기 때문으로 보인다. 판매자 그룹에서는 공동구매 컨소시엄 일정에 맞추어 많은 트랜잭션을 처리하여야 하므로 상품 제안과 참가신청 업무를 오류 없이 최소한의 기간 내에 처리하기를 원한다. 이로 인해 업무 처리를 위한 시스템 지원 최소화, 사용 편의성 증대를 지속적으로 요구하고 있다.

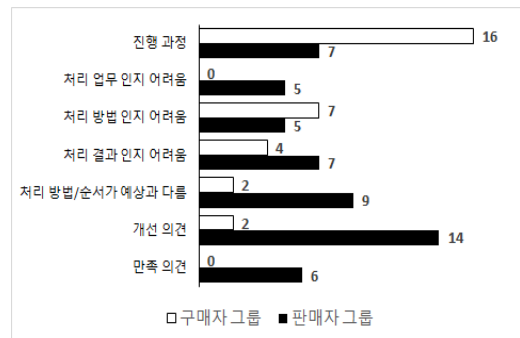


그림 3. 그룹간 발생사고 건수 비교

이러한 그룹간의 특징은 개선의견과 만족 의견에 있어서도 명확하게 비교되었다. 구매자 그룹은 개선의견을 거의 제시하지 않았고 만족 의견은 전혀 없었으나 판매자 그룹에서는 적극적으로 표현하였다. 만족 의견을 통해 기존 시스템 대비 개선 효과를 확인하였다.

다. 주요 사용성 분석 사례

[처리 업무 인지 어려움]

처리 업무 인지 어려움 유형은 구매자 그룹에서는 발생하지 않았다. 판매자 그룹 참가자도 업무 유형험자들이어서 업무 내용에 대한 이해가 부족한 것은 아니었으며 대부분은 실제계의 업무와 시스템에서의 업무 프로세스간의 차이로 발생하는 어려움이였다. 대부분의 사용성 문제는 시스템 기능보다는 업무매뉴얼을 보완하고 기능별로 도움말 추가가 필요한 것으로 분석되었다.

[처리 방법 인지 어려움]

처리 방법 인지 어려움 유형에서 대부분 사용성 문제가 구매자 그룹에서는 컨소시엄 참가 신청시, 판매자 그룹에서는 컨소시엄 제안시 발생하였다. 이는 공동구매 운영정책상 다양한 상품 구성을 허용하기 때문인데, 이로 인하여 어떠한 조건의 상품 구성도 가능한 대신 판매자와 구매자에게는 상품별로 선택 옵션이 많고 복잡도가 높아지는 사용성의 문제가 존재하였다.

특히 판매자 측에서는 온라인으로 작성하는 제안서에 최소 78개 항목을 기입하게 되는데 작업자에게 심리적 부담은 물론, 작업 중 웹 브라우저가 다운되는 경우 데이터 유실문제가 있었다. 이를 해소하고자 개발 단계에서는 제안서를 단계별로 탭 항목으로 분리한 인터페이스를 개발하여 테스트를 실시하였다. 하지만 사용성 테스트에서 참가자들은 제안서 전체를 한 눈에 확인하는 것이 인식성이 좋고 작업이 편리하다는 의견을 제시하였으며, 완료 단계에서 제안서를 하나의 웹 페이지로 통합하여 다시 테스트한 결과 작업 시간과 마우스 움직임이 매우 감소한 것으로 확인되었다.

표 3. 주요 구술내용

구술내용 유형	주요 내용
처리 업무 인지 어려움	"주관기관 참고사항에는 뭐 가격이 안 나왔을 경우 쓰면 되나요?" [판매자 그룹 참가자1, 태스크1] "재공주기라는 개념이 한 번 뽑았을 때를 말하는 건가요?" [판매자 그룹 참가자2, 태스크2]
처리 방법 인지 어려움	"아, 옵션2번인가요? 그럼 패키지1로 갈 필요가 없는 건가요?" [구매자 그룹 참가자2, 태스크2] "이건 어떤거예요? (옵션 입력 팝업창을 열어 놓은 상태인) 이 상태에서 옵션을 추가하는 거예요? 잘 이해 안 가는데요." [판매자 그룹 참가자2, 태스크1] "타이틀, 콘텐츠에 대한 수정은? 여기서 이렇게 추가? 안 쓰는 건가요? (콘텐츠 등록 팝업창에서) 응. 안 떠서." [판매자 그룹 참가자1, 태스크1]
처리 결과 인지 어려움	"제출된 것 맞나? 접수대기" [구매자 그룹 참가자1, 태스크2] "자장을 꼭 하고 수정버튼을 눌러야 하나요?" [판매자 그룹 참가자1, 태스크2] "되었죠? 수정하면? 수정을 누르는 게 아니라, 이렇게 되었다고 저장할. 이걸 하면 되나?" [판매자 그룹 참가자2, 태스크1]
예상과 다르게 작동함	"옛날 같은 경우 업로드 하면 자동적으로 이용범위가 지정이 되었는데요. 저날 같은 경우는 그랬던 것 같은데. 그렇지 않았어요?" [판매자 그룹 참가자1, 태스크1] "(키보드로 "자동" 이라고 입력할. 이후 콤보박스 선택하면서) 아 여기 있는데" [판매자 그룹 참가자2, 태스크1] "참가신청, 약관 동의, 눌러야 나오나?" [구매자 그룹 참가자1, 태스크2]
개선 의견	"(계약 관련 정보화면) 이거를 일일이 쓰는 것은 좀 괴로울 것 같은데요. 대행사 아이디로 로그인 했으니까 그 정보를 불러서 쓸 수는 없을까요?" [판매자 그룹 참가자1, 태스크1]
만족 의견	"(계약서류 파일이 미리 입력되어) 아, 네 좋은 것 같아요" [판매자 그룹 참가자1, 태스크1] "아, 이걸 되네. 이런 것도 없었거든요. (웹 에디터 되돌리기 가능 클릭) 그래서 불편해가지고." [판매자 그룹 참가자2, 태스크1]

[처리 결과 인지 어려움]

구술 건수로 비교할 때 처리 결과 인지의 어려움은 처리 방법 인지의 어려움만큼이나 많은 사용성 문제가 있는 것으로 나타났다. 주요 원인은 테스트 당시 개발 중인 시스템이라 작업 결과 안내 창이나 알림 메시지가 부족한 때문인 것으로 나타났지만, 테스트 참가자 입장에서는 태스크를 정상적으로 수행했는지 확인받고 싶은 심리적 영향도 있는 것으로 보인다.

업무 처리를 해야 하는 웹 사이트에서 업무 프로세스가 완료 상태인지, 미완료 상태인지, 혹은 어떤 처리를

해야 완료가 되는지 사용자에게 인지시키는 것은 매우 중요한 문제이다. 이를 위해서는 명확하고 일관성 있는 용어 사용, 업무 처리 결과 여부에 대한 정확한 메시지 제공이 중요하다.

## 2. 사용성 개선 결과 비교

사용성 개선 후 개선 정도를 측정하고 비교하기 위해 시스템 개선 전후 각각 동일한 참가자에게 동일한 태스크를 수행시킨 후 태스크 수행시간과 마우스 궤적을 비교하였다.

태스크 수행시간은 전반적으로 단축된 것으로 나타났다. 단, 개발 완료 단계 테스트에서는 발생사고법을 수행하지 않았으므로 이로 인한 시간 단축을 감안하여야 한다. 선행연구에 따르면 실시간 발생사고법은 약 15%의 추가 시간이 소요된다는 연구결과가 있다. 본 연구에서 태스크 수행시간은 참가자와 태스크별로 차이가 있었지만 태스크에 따라서는 절반 수준의 시간 단축이 발생한 경우도 있어 사용성은 개선된 것으로 보인다. 다만, 사용성 테스트 참가자의 규모가 작아 통계적 의미를 분석할 수는 없었다. 개발 완료 후 테스트에서도서관측 1번 참가자와 공급 업체 측 1번 참가자의 수행시간을 대략적인 시간으로 기록한 것은 두 참가자가 현장 참석이 불가하여 원격지에서 사용자가 스스로 테스트를 실시하면서 소요시간을 정확히 측정하지 못했기 때문이다.

표 4. 태스크 수행 시간 비교

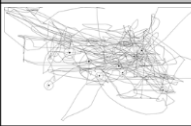
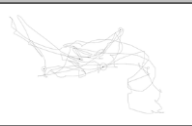
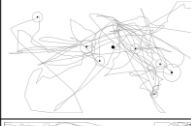
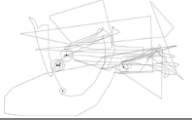
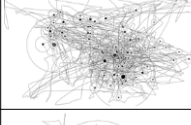
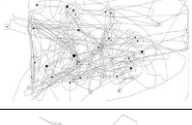

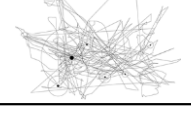
참가자 구분	태스크	소요시간	
		개발 진행 단계 (1차 테스트)	개발 완료 단계 (2차 테스트)
구매자 그룹 참가자1	1	4분 25초	약 1분
	2	2분 50초	약 3분
구매자 그룹 참가자2	1	2분 36초	60초
	2	1분 30초	49초
판매자 그룹 참가자1	1	18분 50초	약 9분
	2	2분 25초	약 1분
	3	1분 40초	약 1분
판매자 그룹 참가자2	1	21분 45초	13분 54초
	2	2분 20초	1분 45초
	3	1분 30초	1분 47초

마우스 궤적을 측정하여 1, 2차 테스트 결과를 시각

적으로 비교하였다. 모든 태스크에 대해 마우스 궤적을 기록하였으나 본 연구결과에서는 각 그룹의 1번 태스크 결과만 수록하였다. 측정 결과 모든 태스크에서 개발 완료 단계인 2차 테스트에서 마우스 궤적이 감소한 것으로 나타났다. 수행 시간과 마우스 궤적을 비교해 보아도 수행 시간이 길수록 궤적의 복잡도가 증가한다. 이와 같이 마우스 궤적만으로도 사용성 개선 정도를 시각적으로 제시할 수 있음이 분명하다.

하지만 구매자 그룹 참가자1과 참가자2의 2차 테스트 수행시간이 약 1분으로 비슷함에도 궤적의 복잡도에는 차이가 있는 것으로 보인다. 또한 판매자 그룹의 2차 테스트에서 참가자2의 수행 시간이 참가자 1보다 약 1.5배 소요되었음에도 불구하고 마우스 궤적은 상대적으로 단순해 보인다. 마우스 궤적만으로는 태스크 수행에 관한 개인 간의 차이를 파악하는데 한계가 있다.

표 5. 각 그룹 태스크1 마우스 궤적 비교

참가자 구분	개발 진행 단계 (1차 테스트)	개발 완료 단계 (2차 테스트)
구매자 그룹 참가자1		
구매자 그룹 참가자2		
판매자 그룹 참가자1		
판매자 그룹 참가자2		

## V. 결론

본 연구는 비즈니스용 웹 사이트 개발 단계에서 내부 직원이 직접 웹 사용성을 테스트하고 사용성 개선 정도를 체감할 수 있는 데이터로 제시한 사례 연구이다. 연



구진은 웹 사이트 운영 및 개발 현업에 비교적 용이하게 적용가능한 발생사고법, 성능 측정, 코칭 방법, 원격 테스트를 혼합한 테스트 방법을 마련하였다. 사용성 테스트 과정에서 발생하는 참가자들의 구술 내용, 촬영 동영상, 마우스 궤적, 소요 시간 데이터 처리 과정에 대해서도 하나의 모델로 제시하였다. 특히 웹 개발 시 정보설계, 웹개발, 디자인 등 분야별로 사용성 개선에 대한 역할 분담이 가능하도록 발생사고 내용 코딩 유형을 제시하였다. 또한 사용성 개선 정도를 손쉽게 측정하면서 서도 직관적으로 체감할 수 있는 방안으로 사용성 테스트 소요시간 및 마우스 궤적 비교를 시도하였다. 이러한 연구방법과 사용성 테스트 데이터를 통해 비용효과적인 웹 사용성 테스트 방법과 사용성 개선 성과 측정 방법을 제시하고자 하였다.

최소의 비용으로 사용성 테스트를 수행함에 따라 테스트 데이터의 정확도나 풍부함에 다소 부족함도 있었다. 마우스 궤적 측정은 단순 궤적의 모양만 기록되었고 거리나 좌표 등의 전문 소프트웨어의 기능은 없었다. 개발 완료 후 2차 테스트에서는 참가자의 사정상 원격 테스트가 불가피하였으며 이로 인해 소요 시간 측정에도 오차가 발생하였다. 전체 프로세스 개선을 위한 사용성 테스트라면 본 연구의 수준에서도필요한 데이터 수집은 가능할 것으로 사료되지만, 특정 상품이나 메뉴에서 수익 측정을 해야 하는 민감한 경우라면 보다 전문적인 툴의 사용이 필요할 것으로 보인다.

본 연구 수행결과 비즈니스용 웹 사이트 사용성 테스트와 사용성 개선을 위한 다음과 같은 후속 연구가 필요하다고 본다. 우선 실시간 발생사고법과 더불어 회고적 발생사고법 활용에 대한 연구가 필요하다. 웹 사용성 테스트 참가자로부터 표현되는 즉각적인 반응만으로는 사용성 문제를 파악하는데 한계가 있다. 후속으로 현업에서 쉽게 따라할 수 있는 사용성 테스트 분석 가이드 개발이 필요하다고 본다. 본 연구에서 발생사고 내용을 7개 유형으로 구분하여 웹 개발 과정에 있어 업무 분담을 시도하였다. 이를 보다 정교하게 구분하여 정보설계, 개발, 디자인의 각 역할 자들에게 업무를 구분하여 할당할 수 있는 지침이나 모델이 제공되어야 시스템 전체에서 일관성 있는 사용성 개선이 가능할 것이다.

본 연구에서는 비즈니스용 웹 사이트의 업무 효율성 증대에 초점을 맞추어 현업에서 즉시 활용 가능한 사용성 테스트 모델과 분석방법을 제시하고자 하였다. 사용성 테스트는 하나의 방법론을 범용적으로 활용할 수도 있겠으나 웹 사이트별로 목적, 분야, 사용대상 등에 따른 특성이 존재하므로 향후 도메인별로 차별화 된 사용성 테스트 모델을 연구개발한다면 현업에서 활용하는데 도움이 될 것으로 생각한다.

**참 고 문 헌**

- [1] <https://www.nngroup.com/articles/return-on-investment-for-usability/>, 2018.1.31.
- [2] [http://www.humanfactors.com/whitepapers/ROI\\_of\\_user\\_experience.asp](http://www.humanfactors.com/whitepapers/ROI_of_user_experience.asp), 2018.1.31.
- [3] <https://www.nngroup.com/articles/summary-of-usability-inspection-methods/>, 2017.11.26.
- [4] <http://usability.jameshom.com>, 2017.11.26.
- [5] 김정룡, 이하연, 최영철, “다자 중심적 사용성 평가 방법의 구조화 과정,” 대한인간공학회지, 제24권, 제2호, pp.25-33, 2005.
- [6] A. Fernandez, E. Insfran, and S. Abrahao, “Usability evaluation methods for the web: A systematic mapping study,” *Information and Software Technology*, Vol.53, No.8, pp.789-817, 2011.
- [7] M. Haak, M. Jong, and P. J. Schellens, “Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue,” *Behaviour & Information Technology*, Vol.22, No.5, pp.339-351, 2003.
- [8] R. Teague, K. Jusus, and M. Nunes-Ueno, “Concurrent Vs. Post-Task Usability Test Ratings,” *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, pp.289-290, 2001.
- [9] L. Cooke, “Assessing Concurrent Think-Aloud Protocol as a Usability Test Method: A

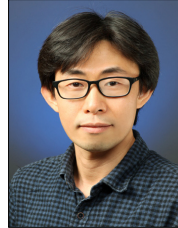
Technical Communication Approach,” IEEE Transactions on Professional Communication, Vol.53, No.3, pp.202-215, 2010.

- [10] M. Chen, J. Anderson, and M. H. Shon, “What can a mouse cursor tell us more?: correlation of eye/mouse movements on web browsing,” CHI '01 Extended Abstracts on Human Factors in Computing Systems, pp.281-282, 2001.
- [11] L. Cooke, “Is the Mouse a ‘Poor Man’s Eye Tracker’?,” Society for Technical Communication Conference, pp.252-255, 2006.
- [12] V. Navalpakkam and E. Churchill, “Mouse tracking: measuring and predicting users’ experience of web-based content,” Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp.2963-2972, 2012.
- [13] S. C. Yang, “Reconceptualizing think-aloud methodology : refining the encoding and categorization techniques via contextualized perspectives,” Computers in Human Behavior, Vol.19, No.1, pp.95-115, 2003.
- [14] 김병주, 이건표, “웹 사용성 평가를 위한 통합평가모형 제안 및 도구 개발,” 디자인학연구, 제20권, 제5호, pp.39-50, 2007.
- [15] 김영준, 김영진, “탐색경로 일치도 분석을 이용한 웹사이트 사용성 평가,” 한국산학기술학회논문지, 제14권, 제2호, pp.793-803, 2013.
- [16] S. McMullen, “Usability testing in a library Web site redesign project,” Reference Services Review, Vol.29, No.29, pp.7-22, 2001.
- [17] 유진환, 김승인, 김희연, 이소희, 최용수, 이병희, “NTIS-NDSL 연계서비스 R&D Navigator 사용성 평가,” 디지털디자인학연구, 제13권, 제1호, pp.35-46, 2012.
- [18] M. J. Aamir and A. Mansoor, “Testing Web Application from Usability Perspective,” 2013 3rd IEEE International Conference on Computer, Control and Communication (IC4), pp.1-7, 2013.

저 자 소 개

김 재 훈(Jayhoon Kim)

정회원



- 1999년 2월 : 연세대학교 문헌정보학과(학사)
- 2008년 2월 : 성균관대학교 경영학과(석사)
- 2006년 ~ 현재 : 한국과학기술정보연구원 선임연구원

<관심분야> : 디지털 콘텐츠, UX

신 재 민(Jaemin Shin)

정회원



- 2006년 2월 : 충남대학교 문헌정보학과(학사)
- 2008년 2월 : 충남대학교 문헌정보학과(석사)
- 2014년 ~ 현재 : 한국과학기술정보연구원 선임기술원

<관심분야> : 정보검색, 오픈엑세스, 미디어터

김 정 환(Jeong Hwan Kim)

정회원



- 2013년 : 충남대학교 문헌정보학(박사)
- 2014년 : 유럽핵입자물리연구소 CERN 방문연구원
- 2006년 ~ 현재 : 한국과학기술정보연구원 책임연구원

<관심분야> : 데이터베이스

임 석 중(Seok Jong Lim)

정회원



- 1998년 2월 : 중앙대학교 문헌정보학과(석사)
- 2009년 8월 : 중앙대학교 문헌정보학과(박사)
- 2005년 ~ 현재 : 한국과학기술정보연구원 선임연구원

<관심분야> : 오픈엑세스, 출판 플랫폼, 리포지터리