

# 한국 기술 현황에 적합한 신규 Heuristics 생성을 위한 기존 Usability Heuristics 평가

## Evaluating Existing Usability Heuristics to Create a New Set of Heuristics for the Current State of Korean Technologies

정 영 주\*, 정 구 철\*\*

Young-Joo Jeong\*, Goo-Cheol Jeong\*\*

### 요 약

유용성 평가는 컴퓨터를 사용함에 있어 사용자 인터페이스를 설계하는데 가장 고려되어야 할 중요한 측면이다. 유용성 평가 방법은 Human Computer Interaction(HCI) 전문가들이 주로 사용하는데, 가장 보편적으로 많이 사용되는 Nielsen이 제시한 평가 기준은 한국 실정에서 실제로 검증된 바가 없다. 지난 연구에서 Nielsen의 heuristics은 일반적인 부분들이 많아 적용하기가 너무 보편적일 수 있고, 또 평가 기준이 영어로 되어 있어 문화적 차이 등의 이유로 몇 가지 기준들이 이해하고 적용되기 어려웠다는 것을 발견하였다. 이에 본 연구에서는 유용성 평가에서 대표적으로 사용되는 Jakob Nielsen의 10가지 usability heuristics를 컴퓨터 전문가 집단을 대상으로 설문조사하여 유용성 정도와 삭제 및 추가 평가 기준의 논리적 근거를 마련하였다. 본 연구에서 분석된 결과물은 추후 새로운 평가 기준안을 도출하는데 많은 도움이 될 것이다.

**Key Words** : Human computer interaction, usability evaluation, software design, user-centric design, heuristics

### ABSTRACT

Usability heuristics(“heuristics”) are general principles for usability evaluation during user interface design. This method is commonly used by Human Computer Interaction(HCI) professionals; however, the most widely used set of heuristics which were originally created by Nielsen has not yet been found practical in the current state of Korean technologies. In our prior research, we found that some of Nielsen’s heuristics are difficult for some evaluators to understand and insufficient to fully evaluate Korean applications, due to the broad-applicability of these heuristics and differences in cultural context. Therefore, in this study, professionals in computer science and related fields evaluated Nielsen’s ten usability heuristics in order to gather logical bases for finding areas for improvement. The results of this study will help creating a new set of heuristics that will better evaluate more recently developed applications.

---

\* Bank of America Merrill Lynch, Market Risk Strategic Business Analyst(jyj@cmu.edu)

\*\* 한국기술교육대학교 컴퓨터공학부(jeong@kut.ac.kr)

제1저자 (First Author) : 정영주

교신저자 : 정구철

접수일자 : 2011년 5월 02일

수정일자 : 2011년 6월 02일

확정일자 : 2011년 6월 14일

## 1. 서론

컴퓨터를 이용한 첨단기기들은 갈수록 라이프 스타일이 짧아지고 있으며 무수히 많은 신제품들이 경쟁적으로 출시되고 있다. 더욱이 최근에 발표되는 컴퓨터 응용 기기들은 휴대용 컴퓨터인지, 이동통신 기기인지 구분이 모호할 뿐 아니라 한 명의 사용자가 다수 기기를 복합적으로 사용하고 있기 때문에 서로 다른 프로그램과 운영체제를 배우고 익숙해져야 하는 부담을 안고 있다. 따라서 프로그램과 플랫폼을 개발하는 제작사의 입장에서는 각자의 고유함을 유지하며 독창성을 주장하는 것도 중요하지만, 사용자 측면에서 어플리케이션간의 상호호환성을 유지할 수 있는 방안을 고려하지 않으면 안 될 것이다. 이러한 측면에서 어떤 프로그램이 유용할 것인가? 라는 평가기준은 또 하나의 우선적인 연구 과제라 생각된다. 지난 연구에서 HCI 비전문가들을 대상으로 설문하였을 때 Jakob Nielsen의 평가 기준은 몇 가지가 이해하기 어려웠고 이해하기 어려운 이유 중에 다른 기준들과 중복되거나 혼동되고, 우리말로 번역되었을 때 문화적 또 언어적 차이가 있다는 점을 발견하였다[5]. 이는 평가 기준이 영어로 작성됨에 따른 문화적 차이와 Whorfian effect (i.e. Sapir-Whorf hypothesis)[8]의 개인의 세계관은 모국어에 의해서 결정된다고 하는 설, 그리고 좀 더 응용해서는 한국인의 단어적 이해와 문화적 배경이 기존의 영어로 표현된 평가기준을 이해하는데 영향을 미친다는 것(그래서 오해의 소지가 생길 수 있다는 것) 때문에 기존 평가 기준이 이해되기 어려웠을 수도 있다.

또한 기존의 사용성 평가 기준은 다양한 시스템과 어플리케이션을 평가할 수 있도록 일반적인 평가 준거를 제시하는데 초점이 두었으나[1][6], 한국의 기술 현황과 우리나라 사용자들의 특성, 또 우리나라에서 개발된 프로그램들이 많다는 점, 그래서 해외의 프로그램들과 상호 호환성을 유지해야 하는 부분이 많다는 점 등을 염두 했을 때, 기존의 사용성 평가 기준이 다 충족시키지 못하는 점이 있어서, 우리나라 상황에 좀 더 알맞은 사용성 평가도구가 만들어져야 한다는 필요성이 제기되었다. 또한 오늘날 사용자들은 급속도로 발전하고 개발되어 출시되는 다양한 플랫폼들을 함께 사용한다는 점에서도 프로그램들에 대한 평가기준은 달라져야 한다.

기존의 사용성 평가 기준이 평가 준거를 제시하는데 초점을 두어왔다는 한계를 염두하여 Bertini[2],

Chattrachart[3] 등이 이와 관련한 논문을 발표하고 우리나라에서도 입찰일 등에 의하여 평가하고자 하는 프로그램에 상세한 평가 기준을 만들고자 하는 시도가 있었지만 이 논문에서 제시한 평가 기준은 모바일 컴퓨팅[2]이나 디지털 교과서 플랫폼의 사용성[1]에만 주안점을 두었다. 따라서 본 논문에서는 보편적이면서도 한국 사정에 좀 더 적합한 평가기준을 제안하고자 한다.

### 1. 연구의 배경

2010년 연구 진행된 Jakob Nielsen의 10 Heuristic의 이해 정도 조사에서 얻은 결론에 의하면 조사대상자들은 대체로 Nielsen의 열 가지 기준을 좋은 컴퓨터 프로그램의 개발과 평가의 기준으로 잘 이해하고 있었으나, 몇몇 항목은 단어나 내용이 수정되어야 한다고 생각하고 있으며, 다른 항목의 기준과 중복되거나 혼동되므로 기준으로 적합한지 재고되어야 한다는 결론을 얻었다. 또한 급속히 발전해 나가는 컴퓨터 관련 업계의 첨단 개발 제품들은 단지 컴퓨터뿐만이 아니라 다양한 유사 운영체제와 소프트웨어를 사용하고, 사용자가 여러 가지 다양한 플랫폼을 사용하며 작업하는 상황에 이르렀다. 그러므로 좋은 프로그램의 평가 기준도 시대적 현황에 적합하게 그 기준이 수정되어야만 한다.

### 2. 연구의 목적

우리나라는 세계 여러 나라에서 개발되는 제품의 첨단 사양들을 빨리 도입하여 사용하고 평가하는 평가시장으로서의 기능뿐만이 아니라 이들 첨단 제품들을 다양한 측면에서 개발하여 컴퓨터관련 제품의 첨단 개발국으로서의 선두적인 입장에 있으므로 연구 개발자들의 관점에서 프로그램 개발의 기준으로 삼고자 하는 heuristic이 존재할 것으로 판단된다.

그러나 HCI 분야의 학문적 발전은 아직 초보적인 수준이고 우리나라 사람들에 의해 정리된 heuristic은 아직 공식적으로 발표되지 않는 상황이다. 이에 본 연구에서는 먼저 Nielsen이 제시한 기준을 현 시점에서 전문가 집단의 평가를 토대로 개정안을 도출하고, 이를 바탕으로 새로운 기준을 개발할 목적으로 의견 수렴을 위한 설문도구를 작성하고자 한다.

### 3. 연구기준이 되는 Nielsen의 10가지 Usability Heuristics

Nielsen의 10가지 usability heuristics는 Human Computer Interaction에서 보편적으로 사용되는 유용성 원칙들 중에서 대표적인 것으로 보통 몇 명의 평

가자 즉 전문가나 사용자들이 지침이 되는 10가지 heuristics를 가지고 사용자 인터페이스를 평가하는 방법이다. 이 10가지 usability heuristic는 Nielsen이 고안한 Enhancing the Explanatory Power of Usability Heuristics 연구 등 여러 연구들을 거쳐 개정된 것이다.

다음은 Nielsen이 제안한 10가지 Heuristics이다.

1. Visibility of system status
2. Match between system and the real world
3. User control and freedom
4. Consistency and standards
5. Error prevention
6. Recognition rather than recall
7. Flexibility and efficiency of use
8. Aesthetic and minimalist design
9. Help users recognize, diagnose, and recover from errors
10. Help and documentation

## II. 연구 방법

새로운 기준을 개발하기 위해서는 먼저 기존 Nielsen이 제시한 10가지 기준이 컴퓨터 프로그램의 평가 기준으로 유용하다고 생각하는지를 조사해야 하며 만약 문제점이 있다고 판단한다면 어떤 이유로 또 어떻게 수정되어야 하는지를 우선적으로 파악해야 한다. 따라서 현재 가장 많이 사용하는 Nielsen의 기준에서 제외시킬 항목은 무엇인지 또 이 기준에 새롭게 추가할 기준을 발굴하기 위해서 설문조사서를 작성하였다. 설문조사서는 객관식 문항과 자유롭게 의견을 피력할 수 있는 주관식 문항으로 작성하였고, 설문조사 대상으로는 컴퓨터 프로그램을 다양하게 사용해본 경험자와 컴퓨터 프로그램 개발 분야에서 장기간 종사하는 전문인에게 응답하게 하였다. 그리고 설문 응답자가 제시한 의견과 2010년도 저자가 발표한의 닐슨의 열 가지 기준에 대한 이해 정도 조사[5]에서 수집된 의견들을 고려하여 선택형 설문지를 작성하였다.

1) 1단계: 2010년 11월의 연구조사서를 활용하여 Nielsen의 10가지 기준을 “전혀 동의하지 않음” 1에서 “절대 동의” 10까지의 열 단계로 비율 평가하도록 요구하였고, 평가한 이유를 주관식으로 기록하도록 하였다. 이 객관식, 주관식 설문 조사서를 2명의 프로그램 개발자와 컴퓨터 관련 비전문가이나 프로그램 사용자로서 오랜 경험이 있는 교수를 대상으로

응답지를 작성하게 하였다.

2) 2단계: 설문 응답한 내용과 2010년의 이해도 조사 연구의 결과에서 얻은 견해를 기초로 하여 유용성 정도를 비율 평가하는 척도를 10단계에서 5단계로 축소하여 평가하게 하였다.

- 절대 동의 한다 : 5
- 대체로 동의 한다 : 4
- 필요한 항목이라고는 생각 한다 : 3
- 대체로 동의하지 않음 : 2
- 전혀 동의하지 않음 : 1

그리고 응답자가 유용성 정도를 평가한 이유를 선택 항목을 열거해 선택하도록 작성하였고, 선택 항목에 그 이유가 없는 경우는 추가로 응답자의 의견을 직접 기입할 수 있는 항목을 넣었으며, 조사대상자가 프로그램을 평가할 때 Nielsen의 10가지 기준을 모두 사용할 것인지, 제외한다면 어느 항목의 기준을 제외 할 것인지 직접 기입하고 또한 설문 조사에 대한 방해 요인이 있었는지에 대해 직접 서술할 수 있도록 설문지를 작성하였다.

3) 3단계: 2단계에서 작성된 설문 조사서를 컴퓨터 관련 전문 분야 종사자로서 컴퓨터 프로그램들을 다양하게 사용한다고 판단되는 전문가들에게 배포하여 응답지의 결과를 정리 분석하였다.

이를 기준으로 무응답의 배제와 정확한 의미 파악을 위해 설문지의 형태를 좀 더 선택형으로 재조정하였고, 제안된 의견을 바탕으로 연구자들의 의견을 고려하여 추가하기 원하는 기준에 대한 선택형 질문을 작성하였다.

## III. 연구 결과 분석

### 1. 프로그램 평가도구 개발 기초 설문조사 분석

조사대상자의 일반적인 특성은 다음과 같다.

연령은 20대:3명, 30대:3명, 40대:3명, 50대:2명, 성별은 남 10명, 여 1명이다. 학력은 학사:2, 석사과정 이수:3, 석사:2, 박사과정 수료:1, 박사:3 이었다. 전공학위는 경영학 전공자 1명, 디자인 전공자 1명이며 나머지는 모두 전기 전자 컴퓨터 관련 공학계열 전공자였다. 컴퓨터 사용 년 수는 7년~10년:4명, 11년~20년:2명, 21년~30년:5명으로 장기간 컴퓨터 프로그램을 사용해 온 전문가 집단이었다.

컴퓨터 프로그램 평가도구 개발 기초조사서를 이용하여 Nielsen의 10가지 기준의 유용성을 평가한 설문 조사를 실시하였다. 구체적으로 13명의 설문조

사 응답지에서 자료 분석에 유용하지 않은 2명의 설문 조사서를 제외하고 11명의 응답을 정리한 결과는 다음과 같다.

1) Visibility of system status

“시스템의 상황을 사용자에게 보이기” 기준에서 유용성은 절대 유용하다(5)에 8명, 대체로 유용(4)에 3명이 응답하였는데, 대체로 유용으로 평한 이유는 (4)항 “사용자가 프로그램을 사용할 때, 오히려 번거롭다 느끼게 한다”에 1명, (6)항 “익숙하지 않은 사용자에게는 프로그램 사용을 어렵게 만들어 줄 수 있다”에 1명, (9)항 “이 기준 항목이 프로그램 진행과 처리 자체를 더디게 한다”에 1명으로 나타났다.

2) Match between system and the real world

“시스템과 사용자가 실생활에서 익숙하게 사용 혹은 연상 하는 것들과 연결시키기” 기준은 응답 결과가 절대 유용(5)에 9명, 대체로 유용(4)에 1명, 대체로 동의 안함(2)에 1명으로 대부분 긍정적인 평가를 하였으나 극단적인 반대의 의견을 표명한 경우도 극소수 있었다.

대체로 유용하다를 선택한 이유는 (9)항 “이 기준 항목이 프로그램 진행과 처리 자체를 더디게 한다”에 1명, 대체로 동의하지 않음(2)을 선택한 1명도 (9)항과 같은 이유를 들었다.

3) User control and freedom

“사용자가 쉽게 컨트롤과 취소를 할 수 있는 자유를 주어야 한다.” 기준은 대부분 절대 유용(5)하다고 평가하고, 1명만이 필요한 항목이라고는 생각(3)에 응답하였는데 (3)으로 평가한 이유로는 프로그램의 사용이 익숙하지 않은 사용자는 어렵게 느껴져 사용을 잘 못하게 할 수 있기 때문이라 설명하였다.

4) Consistency and standards

“언어나 습관이 일률적이고 표준에 맞아야 한다” 기준은 절대 유용(5)에 4명, 대체로 유용(4)에 6명, 평가기준으로는 인정(3)에 1명으로 평가되어 평가기준으로 많은 동의를 얻지 못하였다. 그 이유도 다양하게 지적되었는데, 유용성이 대체로 유용하다(4)를 선택한 이유로서는 (1)항 “이 기준 항목은 평가기준으로 필요하나 반대되는 경우가 있다”에 1명, (6)항 “익숙하지 않은 사용자는 프로그램 사용을 어렵게 만들어 준다”에 1명, (9)항 “이 기준 항목이 프로그램의 질을 저하시킨다.”에 1명, (11)항 “이 기준 항목

은 다른 항목과 중복된다.”에 1명, (12)항 “이 기준 항목은 다른 기준 항목과 혼돈된다.”에 1명이 각각 선택하여 다양한 지적을 받은 기준 이었다.

기타 의견으로 직접 서술한 내용으로는 지속적이란 단어의 의미 제한이 필요하고, 새 개념 도입을 방해한다는 견해를 주었다. 또한 유용성 평가가 (3)인 이유로는 (8)항 “이 기준 항목이 좋은 프로그램 개발을 방해한다.”에 1명이 선택하였다. 그러나 이 기준 항목은 평가기준에서 제외시키고 싶은 항목을 열거 하라는 문항에서 제외시키고 싶다고 지적되지는 않았다.

5) Error prevention

“실수 예방의 측면을 고려해야 한다” 기준은 절대 유용(5)에 5명, 대체로 유용(4)에 5명, 평가기준으로는 인정(3)에 1명으로 평가되었고, 대체로 유용(4)를 선택한 이유로서는 (1)항 “이 기준 항목은 평가기준으로 필요하나 반대되는 경우가 있다.”에 1명, (4)항 “이 기준 항목이 오히려 사용자가 번거롭다 느끼게 한다.”에 2명, (6)항 “익숙하지 않은 사용자는 프로그램 사용을 어렵게 만들어 준다.”에 3명이 선택하였고, 필요한 항목이라고는 생각함(3)으로 평가한 응답자는 (4)항 “이 기준 항목이 오히려 사용자가 번거롭다 느끼게 한다.”는 항목을 선택하였다.

6) Recognition rather than recall

“기억보다는 인지하게 하기” 기준은 절대 유용(5)에 7명, 대체로 유용(4)에 3명, 평가기준으로는 인정(3)에 1명이며 그 이유로서는 대체로 유용(4)인 경우 (4)항 “오히려 사용자가 번거롭다 느끼게 한다.”에 2명, (6)항 “익숙하지 않은 사용자는 프로그램 사용을 어렵게 만들어 준다.”에 1명, (9)항 “프로그램 진행과 처리 자체를 더디게 한다.” 1명으로 답하였고, 유용성을 중간 수준으로 동의한 경우의 이유로는 (7)항 “프로그램 평가기준으로 필요치 않다.”에 1명으로 극소수이지만 부정적인 지적이 나왔다.

7) Flexibility and efficiency of use

“사용의 효율성과 신속성” 기준에 대해서는 절대 유용(5)에 8명, 대체로 유용(4)에 3명으로 평가되었고, 대체로 유용(4)를 선택한 이유로서는 (4)항 “오히려 사용자가 번거롭다 느끼게 한다.”에 1명, (6)항 “익숙하지 않은 사용자는 프로그램 사용을 어렵게 만들어 준다.”에 1명, (8)항 “좋은 프로그램 개발을 방해한다.”에 1명으로 각각 지적되었다.

8) Aesthetic and minimalist design

“보기 좋고 단순하게 고안하기” 기준에 대해서는 절대 유용(5)에 9명, 대체로 유용(4)에 2명으로 대부분 유용성을 인정하였다. 여기서 (4)로 평가한 2명의 응답자는 (1)항 “평가기준으로 필요하나 반대되는 경우가 있다.”와 (11)항 “다른 항목과 중복된다.”를 각각 들었다.

9) Help users recognize, diagnose, and recover from errors

“실수 발생에 대한 인식, 진단, 복구를 할 수 있게 해 주어야 한다.” 기준은 는 절대 유용(5)에 9명, 평가기준으로는 인정(3)에 2명으로 나타났는데 (3)으로 평가한 이유는 (1)항 “평가기준으로 필요하나 반대되는 경우가 있다.”에 1명, (4)항 “사용자가 번거롭다 느끼게 한다.”에 1명이 선택하였다.

10) Help and documentation

“도움말이나 설명이 있어야 한다.” 기준은 절대 유용(5)에 10명, 대체로 유용(4)에 1명으로 거의 대부분 찬성하였고, 이유로서는 오히려 사용자가 (5)항 “치명적인 실수를 할 수 있다.”로 지적하였으나 이는 응답자가 항목을 잘못 체크한 것으로 보인다. 도움말이나 설명이 사용자의 치명적인 프로그램 사용 오류를 만들 가능성은 별로 없다고 판단되기 때문이다.

그림 1은 10가지 Usability Heuristics 기준에 대한 유용성 평가 결과를 그래프로 도시한 것이다.

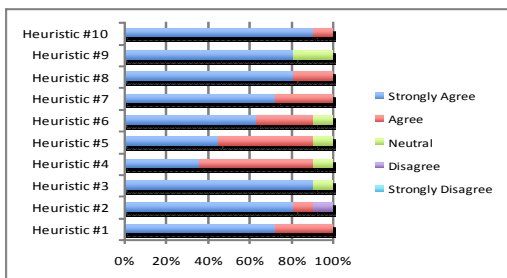


그림 1. 10가지 Usability Heuristics 기준에 대한 유용성 평가 결과  
Fig 1. Survey Responses for Evaluating the Applicability of Nielsen's 10 Usability Heuristics

결과적으로 Nielsen의 10가지 기준에 대한 평가도구로서의 유용성 조사에서는 기준 항목 1,2,3,7,8,9,10 기준항목은 높은 유용성을 나타내고, 기준항목 4,5,6은 사용자에게 혼돈이나 어려움을 줄 수 있는 가능성이 높아 좋은 프로그램 평가기준으로 그리 좋은 기준이 아니라는 평가를 받았으며 수정의 여지를 가지고 있음을 말 할 수 있다.

그림 2는 10가지 Usability Heuristics 기준에 대한 유용성을 평가할 때 평균 가중치를 그래프로 도시한 것이다.

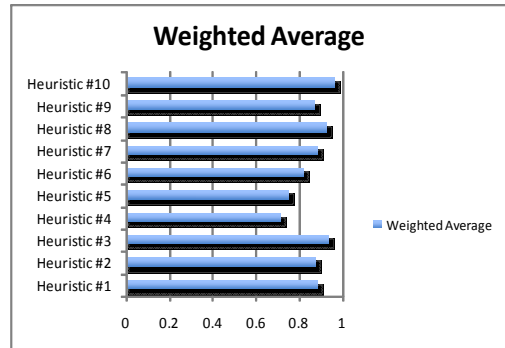


그림 2. 10가지 Usability Heuristics 기준에 대한 유용성 평가 평균 가중치  
Fig 2. Weighted Average of the Survey Responses for Evaluating the Applicability of Nielsen's 10 Usability Heuristics

이 경우 평균 가중치는 평가 스케일 5, 4, 3, 2, 1을 가장 많이 필요하다고 판단할 경우 가중치가 가장 높게 반영될 수 있도록 5\*4=20, 4\*3=12, 3\*2=6, 2\*2=2로 하고 전혀 동의하지 않으면 필요성이 없다고 판단하는 경우이므로 가중치를 0 즉 1\*0=0으로 처리하였다. 이 결과 1~3 문항과 7~10 문항은 기준이 “합격” 되고 4~6번은 개선되어야 한다는 것을 알게 되었다.

2. 프로그램 평가기준 신설 여부 설문조사 분석

1) 1항 - “닐슨의 평가기준 10가지를 모두 사용하시겠습니까?” 질문에는 그렇다가 10명, 아니다가 1명으로 10가지 기준항목을 대부분 사용하겠다고 응답하였다.

2) 2항 - “닐슨의 평가기준 중 제외시키고 싶은 항목은 어떤 것인가?” 질문에는 제외시키고 싶은 것이 없다고 응답한 경우가 4명이었고, 무응답이 4명, 그리고 제외하고 싶은 기준항목을 제시한 경우가 3명이었다. 제외하고 싶은 항목으로는 기준항목 2번과 6번, 기준항목 5번, 그리고 기준항목 5번과 9번을 각각 한명씩 지적하였다. 기준항목 2번과 9번은 유용성 평가에선 높은 점수를 보였음에도 제외시키고 싶은 항목에 거론 되었고, 4번은 비교적 낮은 유용성 점수를 보였음에도 제외시키고 싶은 기준항목에서는 제외되어 응답자들이 선택형과 객관식으로 응답할 때 약간 다른 반응을 보였다고 판단되었다.

3) 3항 - “다른 기준 항목을 만들기 원하신다면 어떤 기준 항목을 추가하시겠습니까?” 질문에는 거의 대부분(7명) 무응답이었고, 한 명만 분명하게 “없다”는 의사를 표시하였다. 모든 항목에 다른 의견이 있다고 응답한 경우가 있었으나 구체적인 의견은 기술하지 않은 경우가 있었고, 적은 의견일지라도 좋은 지적을 해 준 경우도 있었다. 이 의견은 정리하여 최종 설문지를 만드는 데 반영하였다.

그밖에 추가할 다른 의견을 제시한 응답자 3명이 기술한 내용은 다음과 같다.

① 커스터마이징이 가능해야한다. 즉 환경설정이나 절차의 변경 등 사용자의 환경에 맞도록 스스로 프로그램 운영을 컨트롤할 수 있어야 한다.

② 프로그램을 설치하고 사용함에 있어 사용자 수준을 설정하여 설치, 사용할 수 있게 조정하는 기능이 요구된다.

③ 다양한 플랫폼을 지원해야 한다. 즉 프로그램을 사용하는 사용자가 사용하는 플랫폼(Windows, Linux, MAC, 모바일 통신기기 등)이 어떤 것이던지 다양하게 지원이 가능해야 한다.

④ 완벽한 프로그램은 존재하지 않으며 문제가 발생할 때 문제해결을 용이하게 도와주는 패치나 디버그에 대한 준비가 필요하다는 지적들이 있었다.

이 중에서 위의 세 가지 의견에 대해서는 완성 설문지에 조사대상자들의 의견을 묻기 위해 삽입하였으나, ④번 항목은 일반 프로그램 사용자들에게는 실제로 절실히 필요하고 많은 프로그램들이 이를 준비하는 부분이라 할지라도 프로그램 개발자와 평가기준으로는 적합하지 않을 것이라는 판단과 다른 전문가들의 의견을 고려해 제외시켰다.

그 밖에 추가할 다른 의견을 제시한 내용은 다른 설문지들과 마찬가지로 설문지가 눈에 잘 들어오지 않으니 시각적인 측면을 을 고려하여 작성되면 더 좋겠다는 의견이 있었고 이를 고려하여 최종 설문지를 작성하는데 최대한 반영하고자 노력하였다.

#### IV. 결론

설문결과에 의하면 Nielsen의 4, 5, 6번 heuristics 즉 Consistency and standards, Error prevention, recognition rather than recall 항목은 개선점이 필요하다는 것을 알게 되었다. 그 이유는 Nielsen의 기준들이 너무 보편적이고(예: consistency가 어떤 consistency를 의미하는지, 플랫폼에 맞는 비주얼을 써야 한다는 것에 국한되는 것인지 아니면 어플리

이션간의 상호관계 까지도 포함하는 것인지), 사용자를 오히려 더 귀찮게 할 수 있으며(예: 만약 에러를 막기 위해서 사용자가 어떤 스텝들을 더 해야 한다면) 또 이 기준이 정확하게 무엇을 의미하는지 설문 응답자들이 이해하기 어려운 것(예: recognition과 recall의 차이를 이해하지 못함) 일수도 있다. 따라서 다음 연구에서는 좀 더 이해하기 쉽고 사용하기 편한 평가 기준을 제안할 것이다.

Heuristics 평가기준은 사용성을 평가하는데 있어 신속하고도 쉽게 사용될 수 있는 방법이다. 또한 보편적으로 Heuristics는 각각의 어플리케이션을 그 어플리케이션에서만 사용성을 평가하지 어플리케이션 상호간의 호환성이나 융화성은 평가하지 않는다.

따라서 우리나라 실정에 적합한 평가 기준을 작성하는 것은 의미 있는 일이지만, 이 기준을 프로그램 유용성을 향상시키는데 이용되는 유일한 방법으로 사용되면 안 될 것이다. 즉 우리나라에서 독자적으로 제작한 어플리케이션이 한국뿐 아니라 세계 각국에서 널리 사용되기 위해서는 해외 유사 프로그램들과 호환성 문제점을 해결할 수 있도록 프로그램 개발단계에서부터 유의성 있는 평가 기준을 제시하여 보다 개선된 어플리케이션을 개발하여야 할 것이다.

이처럼 현재 기술 발전의 현황과 각 나라의 상황에 따라서 좀 더 알맞은 heuristics을 개발하는 것은 의미 있는 일이지만, 또한 사용성 평가 기준 외에도 usability를 평가하는 다른 방법들(예 : 프로토타입이나 완성된 제품을 사용자가 직접 사용하는 것을 관찰하는 것, eye-tracking device와 녹음/녹화 기능 등을 사용하여 사용자가 어플리케이션을 사용하는 것을 분석하는 것 등)을 병행하여 사용함으로써 다양한 문제점들을 보다 쉽게 도출할 수 있을 뿐 아니라 프로그램을 개선하는 측면에도 큰 기여를 하게 될 것이다.

#### 감사의 글

본 논문은 2011년도 한국기술교육대학교 교수 연구 진흥비 지원에 의하여 수행되었습니다.

#### 참고 문헌

- [1] 임철일, 송해덕, 이예경, 이영태, (2009). 디지털 교과서 플랫폼 사용성 평가도구 개발 및 적용에 관한 연구. 교육공학연구. 25권 4호. 125-155.

