

---

# 무선인터넷 서비스 사용성 지수 평가 체계



## Usability index evaluation system for mobile WAP service



박환수, Hwansu Park

SK telecom

hspark@sktelecom.com

---



**요약** ~ ~ 모바일 무선인터넷 서비스(WAP)는 휴대폰이라는 제한된 화면 크기와 조작부, 브라우저 및 OS 에서의 UI 구현상의 제약 등으로 인해 사용성이 고객만족도에 더욱 중요한 영향을 미치는 요소가 되고 있다. 특히 여러 다양한 콘텐츠 제공사에 의해 각각의 서비스가 개발되어 제공되는 현 상황에서, 이러한 서비스들의 UI 품질수준을 일관된 기준과 방법으로 효과적인 관리를 할 필요가 있다. 본 연구에서는 다양한 무선 인터넷 서비스들에 대한 일관된 UI 수준관리를 위한 사용성 지수 평가 체계를 제안한다. 제안된 평가 체계는 사용성 관련 사용성 지표와 UI 원칙으로부터 무선 인터넷 서비스에서의 UI 설계 요소와 평가 항목들을 도출하는 top-down 방식과, 기존 UI 설계 지침으로부터 사용성과 관계된 평가 항목들을 도출하고, 이들을 사용성 원칙과 지표 관점에서 그룹핑하는 bottom-up 방식을 통해 평가 체계 및 항목들이 구축되었다. 이러한 양방향 평가 체계 구축 방법은 사용성 문제를 야기할 수 있는 다양한 측면을 사용성 지표 관점에서 빠짐없이 고려할 수 있을 뿐 아니라, 실제 서비스 환경과 관련된 UI component 관점에서 구체적인 평가 항목들을 도출할 수 있다는 장점이 있다. 이러한 방법을 통해 구조화된 평가 체계는 사용성 지표와 각 지표와 관계된 사용성 원칙인 UI 가이드라인, 각 가이드라인 별로 구체적인 평가를 위한 UI component 별 평가항목들간 연결을 통해 계층적으로 구성되어 있다. 특히 각 평가 항목들을 O/X 로 판정할 수 있는 구체적인 내용과 형태로 구성하고, 전체 평가 항목 대비 만족된 평가항목의 비율로써 해당 서비스의 사용성 점수가 도출될 수 있게 하여 무선 인터넷 서비스의 사용성 수준을 정량적인 값으로 파악할 수 있도록 하였다. 제안된 평가 체계는 사용자 테스트를 통한 실제 사용성 문제와 비교 분석되어 그 효과를 검증하였으며, 평가 항목별로 평가 대상 및 기준, 사례를 안내해주고 지표별 점수를 자동으로 계산해 주는 S/W 로 구현되어 실제 무선 인터넷 서비스의 평가 및 개선작업에 적용되었다.

**Abstract** Abstract The customer satisfaction of WAP service greatly relies on the usability of the service due to the limited display size of a mobile phone and limitation in realizing UI (User Interface) for function keys, browser, and OS (operating system). Currently, a number of contents providers develop and deliver varying services, and thus, it is critical to control quality level of UI in consistent standards and manner. This study suggests usability index evaluation system to achieve consistent UI quality control of various WAP services. The system adopts both top-down and bottom-up approaches.

The former concerns deriving UI design components and evaluation checklists for the WAP, based on the usability attributes and UI principles. The latter concerns deriving usability-related evaluation checklists from the established UI design features, and then grouping them from the viewpoint of usability principles and attributes. This bi-directional approach has two outstanding advantages: it allows thorough examination of potential elements that can cause usability problems from the standpoint of usability attributes, and also derives specific evaluation elements from the perspective of UI design components that are relevant to the real service environment. The evaluation system constitutes a hierarchical structure by networking usability attributes, UI guideline which indicates usability principles for each attribute, and usability evaluation checklist for each UI component that enables specific evaluation. Especially, each evaluation checklist contains concrete contents and format so that it can be readily marked in O/X. The score is based on the ratio of number of items that received positive answer to the number of total items. This enables a quantitative evaluation of the usability of mobile WAP service. The validity of the proposed evaluation system has been proved through comparative analysis with the real usability problems based on the user test. A software was developed that provides guideline for evaluation objects, criteria and examples for each checklist, and automatically calculates a score. The software was applied to evaluating and improving the real mobile WAP service.

**핵심어:** *Usability evaluation, Usability Index, mobile WAP service*

## 1. 서론

휴대폰을 통한 무선인터넷 서비스의 기능과 종류가 다양하고 복잡해짐에 따라 사용자에게 인지적 부담이나 실수의 가능성이 적은 User Interface 의 중요성이 더욱 강조되고 있다. 특히 개별 서비스 마다 다른 개발자들이 서비스를 개발함에 따라 개발자 개인의 역량에 따라 UI 수준이 변화하는 등 일관된 UI 품질 수준을 유지하기가 더욱 어려워지고 있는 실정이다. 이러한 문제점을 해결하기 위해서 많은 이동통신사에서는 자사의 UI 표준을 정의한 UI 설계 지침과 가이드라인들을 자체적으로 개발하여 서비스들의 UI 설계에 적용하고 있다. 이들은 사용편의성을 고려한 설계 원칙들을 포함하고 있어, 제품 설계 단계에서부터 사용편의성을 고려한 설계가 이루어지도록 지원하는 역할을 수행해 오고 있다. 그러나 이러한 UI 설계 지침들은 규격화 가능한 측면만을 주로 다루고 있어, 잠재적인 사용성 문제점들을 진단, 평가하고 개선하는 역할을 수행하는 데는 한계가 있으며, 근본적으로는 사용성 수준의 평가가 아닌 설계 지원 방법으로 활용되고 있다.

사용성 평가하기 위한 방법들로는 사용자가 실험실 환경이나 실제 이용환경에서 주어진 태스크를 수행하는 동안 소요된 시간이나 발생한 에러수를 측정하는 사용자 테스트 방법[1], 사용성과 관련한 다양한 속성에 대하여 Likert 척도 등을 기반으로 사용자의 주관적 만족도를 평가하는 SUMI, QUIS 와 같은 설문조사 방법[2, 3], 사용자가 태스크를 학습하는데 소요되는 시간을 모형화하여 예측 하는 GOMS, KLM 와 같은 모형 기반 평가[4], heuristic evaluation 이나 cognitive walkthrough 와 같이 UI 전문가가 UI 원칙이나 가이드라인을 바탕으로 사용성 문제점들을 파악하는 전문가 평가 방법[5]들이 있다. 가장 일반적으로 활용되는 사용자 테스트와 같은 방법들은 개별적인 서비스의 문제점들을 파악 하는 데는 용이하나, 많은 시간과 비용이 들 뿐만 아니라, 다양한 서비스들의 사용성 수준을 정량적으로 비교하기에는 적합하지 않다. 설문조사 방법은 다양한 서비스들에 대한 정량적인 비교 평가가 가능하나 사용자의 주관적인 척도에 의지하므로 객관성이 떨어지며, 구체적인 개선 요건을 설계 지침 관점에서 명확하게 파악하기는 힘들다는 단점이 있다. 모형 기반 평가의 경우 설계 초기 단계에 적용할 수 있다는 장점이 있으나 분석에 많은 시간이 소요될 뿐만 아니라 초보 사용자가 아닌 전문가 수준의 사용자를 가정하고 있어 에러가 발생하는 경우에는 적합하지 않다[6]. 전문가 평가의 경우 짧은 시간에 효율적으로 평가를 수행할 수 있어, 실제 현장에서 널리 이용되고 있다. 그러나, 전문가 평가에서 활용되는 평가 기준들은 일반적인 UI 원칙으로만 기술되어 있어, 평가자의 수준에 따라서 평가 결과의 변동이 존재할 수 있으며, 설계 요건화 하여 개발에 반영하기 위해서는 평가 결과에 대한 구체화 작업이 추가로 필요하다. 또한 평가 결과를 정량화 할 경우 평가자간 평가 결과의 변동이 존재할 수 있어 정량적 수치에 대한 객관성이 떨어진다는 단점이 있다.

그러므로 본 연구에서는 전문가 평가 시, 평가자간 평가 결과의 변동을 최소화 할 수 있도록 평가 항목들을 구체화 하고, 이를 통해 정량적인 사용성 수준을 지수화 하여 파악할 수 있는 사용성 지수 평가 체계를 제안한다. 특히 제안된 방법은 무선인터넷 서비스들의 UI 수준을 평가하고 개선방향을 도출하는데 초점을 맞추어 적용되었다.

## 2. 연구 방법 및 절차

본 연구에서는 무선인터넷 서비스의 사용성 수준을 정량적으로 파악하기 위하여 사용성 지표들을 구분하고 각 지표 별로 구체적인 평가 항목들을 구성하였다. 평가자 별 평가 결과의 변동을 최소화 하고, 평가자와 개발자간의 의사소통을 명확하게 할 수 있도록 하기 위해, 평가 항목들은 가능한 평가 대상과 기준이 구체적인 수준으로 정의될 필요가 있었다. 이를 위해 본 연구에서는 사용성 지표로부터 평가 항목을 도출한 후 관련된 UI 구성요소들을 파악하는 top-down 방식과 UI 설계 지침들로부터 평가 항목들을 도출하고 이를 사용성 지표에 대응시키는 bottom-up 방식을 함께 적용하였다(그림 1). Top-down 방식에서는 먼저 문헌조사를 통하여 기존 여러 다양한 사용성 평가 지표들을 수집하고 이들간의 상관관계 및 포함관계 등을 분석하여 주요한 사용성 지표들을 선정하였다. 선정된 지표 별로 관련된 일반적인 UI 원칙들을 수집 정리하였다. 각 UI 원칙에 대하여 무선 인터넷 서비스에서 발생하였거나 발생 가능한 실제 사용성 문제점들을 찾아 대응시키고 이들을 파악하기 위한 평가 항목 및 UI 설계 요소들을 도출하였다. Bottom-up 방식에서는 먼저 UI 설계 규격들이 정의된 설계 지침서들을 분석하여 위반 시 주요한 사용성 문제를 야기시킬 수 있는 UI 설계 규격 및 설계 요소들을 추출하였다. 이러한 설계 요소들을 평가 항목화하여 앞서 top-down 방식에서 도출된 UI 설계 원칙들과

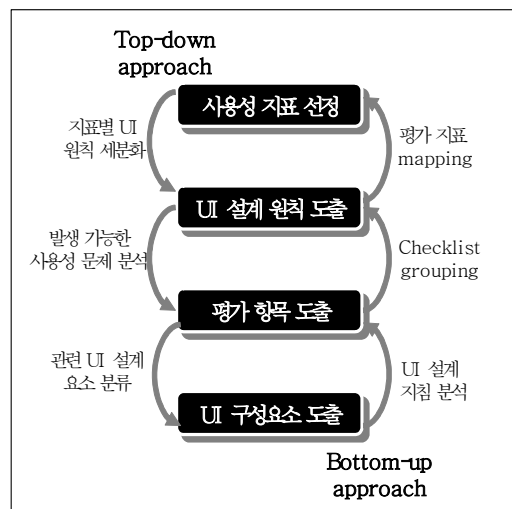


그림 1. 평가 체계 구성 절차  
대응시키거나 새로운 UI 설계 원칙들을 도출, 대응시켰다.

### 3. 사용성 지수 평가 체계

#### 3.1 평가 체계의 구성

구성된 UI 지수 평가 체계의 구조는 그림 2 와 같다. 최상위에는 여러 사용성 측면 별로 사용성 수준을 지수적으로 파악하기 위한 8 개의 사용성 지표들이 존재하며, 각 지표에는 사용성 문제의 발생 원인이나 평가의 근거가 되는 일반적인 UI 설계 원칙들이 할당되어 있으며(총 18 개), 각 원칙들에 대하여 UI 전문가가 실제로 O/X 여부를 판정해야 하는 평가 항목들이 존재한다(총 70 개). 또한 각 평가 항목은 관련된 UI 구성요소들이 대응되어 있다. UI 구성요소들은 서비스 특성에 따라 평가 항목을 설정할 수 있는 기준이 된다.

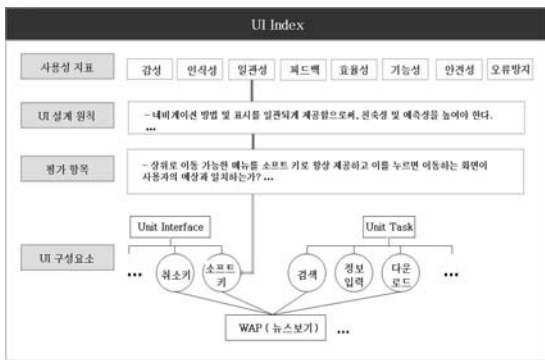


그림 2. 제안된 사용성 지수 평가 체계의 구성

사용성 수준을 보다 정밀하게 파악하기 위해서는 사용자가 체감하는 사용성 관련 측면들을 나누어 살펴볼 필요가 있다. ISO 9241-11 에서는 사용성과 관련한 체감 요소를 효과성(effectiveness), 효율성(eficiency), 주관적 만족도(satisfaction)라는 3 가지 요소로 정의하였다[7]. Nielsen 은 학습성(learnability), 효율성 (efficiency), 기억성(memorability), 에러(error), 만족도(satisfaction) 과 같은 5 가지의 요소로 정의하였다[1]. 그 외에도 많은 연구를 통해 다양한 사용성 지표들이 정의되어 왔다[8-11].

본 연구에서는 기존의 사용성 지표들의 개념적 관련성을 바탕으로 요인 분석을 수행하여 무선인터넷 환경에서의 중요한 사용성 측면들이 상호배타적이면서도 빠짐없이 포함될 수 있도록 8 개의 사용성 지표들을 표 1 과 같이 선정하였다. 특히 무선인터넷 서비스는 다른 web 서비스나 단말 기능과는 달리 과금이란 요소가 존재하며, 사용자에게 신뢰성 있는 과금 정보를 제공하는 것이 중요한 UI 요소 중 하나라고 할 수 있다. 본 연구에서는 이러한 측면을 안전성이란 지표로 구성하였다.

표 1. 선정된 사용성 지표

사용성 지표	의미
감성 (Satisfaction)	사용자가 시스템을 이용하는 과정에서 신뢰감을 갖고, 만족감을 느끼도록 함.
인식성 (Recognition)	시스템에 표현되는 각종 정보의 의미가 쉽게 이해됨
일관성 (Consistency)	작업 절차나 컨트롤의 조작방식 및 정보의 표현 형태가 일반적인 방법과 일치함
피드백 (Feedback)	사용자의 행위나 시스템의 작업 결과에 대한 정보가 명확히 제시됨
효율성 (Efficiency)	작업 내용에 비해 작업 절차가 짧고 간단함
기능성 (Functionality)	상황에 따라 필요한 기능이 제공되어야 하며, 조작이 용이함
안전성 (Safety)	사용자를 부적합한 요구부과나 사생활 침해 등의 위험으로부터 보호함
에러방지 (Error)	실수로 인해 발생될 심각한 결과를 방지할 수 있으며, 문제 발생시 이전 단계로 되돌릴 수 있음

각 지표 별로 할당된 UI 설계 원칙은 해당 사용성 지표를 향상시키기 위해서 지켜져야 할 UI 가이드라인으로 볼 수 있다. 가이드라인들은 무선 인터넷 서비스에서 중요한 요소들을 중심으로 구성되었다. 예를 들어 무선 인터넷 서비스에서는 작은 디스플레이 환경으로 인해 화면간 이동이 빈번하게 일어나게 되며, 조작 방식과 이동 경로가 일관되게 제시될 필요가 있다. 이러한 관점에서 일관성 지표 아래에 “네비게이션 방법 및 표시를 일관되게 제공함으로써 친숙성과 예측성을 높여야 한다” 라는 UI 설계 원칙을 배정하였다.

각 설계 원칙에는 구체화된 평가 항목들이 할당되어 있으며, 평가자는 평가항목들을 기준으로 평가를 진행하게 된다. 평가 항목을 통해 평가가 이루어짐에도 불구하고 설계 원칙을 도입한 이유는 설계 원칙들이 사용성 지표를 설명하는 역할을 수행할 뿐만 아니라, 장기적으로 환경이 변화하여 평가 항목이 수정, 추가, 제거 되더라도 설계 원칙 관점에서 일관된 관리를 함으로써 평가 항목들의 효율적인 할당을 하기 위함이다.

각 평가 항목들에는 UI 구성요소들이 1:1 로 정의되어 있다. UI 구성요소는 단위 Interface 개체(예: 취소키, 소프트키 등)와 관련한 ‘Unit Interface’ 와 절차나 기능 구조(예: 검색, 정보 입력, 다운로드 등)와 관련한 ‘Unit task’ 로 구분되어 있다. 이러한 UI 구성요소를 평가 항목과 대응시킨 이유는 본 평가 체계에서의 평가 항목들이 가능한 O/X 로 판정 내릴 수 있도록, 구체적인 내용으로 기술되어 있어 서비스 특성에 따라 평가 항목들을 선별해서 적용 해야 할 필요가 있기 때문이다. 즉, 모든 서비스에 전체 평가 항목을 적용하는 것이 아니라, 서비스가 가진 기능 구조에 따라서 평가 되어야 할 평가 항목이 달라진다. 이러한 구조는 평가자가 전체 평가 항목들의 내용을 읽고 난 후 평가 항목을 선정하는 것보다 서비스의 기능이나 관련된 UI 구성 요소만 파악하면 평가 항목의 선정이 가능하므로 전체 평가 체계상에서

많은 수의 구체적인 평가 항목들을 효과적으로 관리/적용할 수 있는 장점이 있다.

### 3.2 평가 절차 및 S/W

평가 절차는 다음과 같다.

- 가. 평가 대상 서비스를 분석하여 unit interface 와 unit task 목록으로부터 서비스에 적용된 인터페이스 개체와 기능 요소를 선별함으로써 평가 항목들을 선정한다.
- 나. 해당 서비스를 이용하면서 선정된 평가 항목들에 대하여 만족/ 불만족 여부를 체크한다.
- 다. 불만족된 평가 항목에 대해서는 문제점이 발생한 위치와 원인, 개선안들을 기술한다.
- 라. 선별된 모든 평가 항목들에 대하여 평가가 완료되면 선택된 전체 평가 항목 대비 만족된 평가항목의 비율로서 해당 서비스의 지표 별 점수 및 총점을 계산한다.
- 마. 기입된 불만족 사항을 정리하여, 사용성이 취약한 지표와 구성요소의 취약 이유를 설명하고, 이에 대한 개선 방안을 제시한다.

제안된 평가 체계는 효율적인 평가자 지원을 위해 S/W 로 구현 되었다. S/W 를 통한 평가 시에는 좌측에 선택된 평가 항목들의 목록으로부터 평가 항목간 이동을 하며, 각 평가 항목에 대해서 사용성 지표와 UI 원칙, 그리고 평가 항목과 사례 이미지들을 참고하여 해당 서비스가 평가 항목을 만족했는지, 불만족했는지를 체크하게 된다(그림 3) 평가가 완료되면 그림 4 와 같이 각 지표별 평가 점수가 계산된 결과를 살펴볼 수 있으며, 지표를 클릭하면 평가자가 기입한 문제점들이 나타나도록 구성되었다.



그림 3 구현된 S/W 의 평가 항목 화면

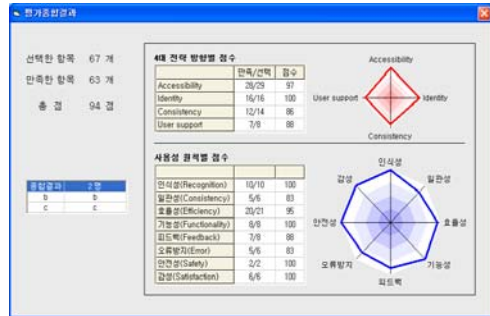


그림 4 구현된 S/W 의 평가 결과 화면

### 4. 제안된 평가 체계의 검증

제안된 평가 체계의 효과성과 정확성을 검증하기 위하여 동일한 서비스에 대하여 사용자 테스트를 통한 실제 사용성 문제점들을 제안된 평가 체계를 통해 파악된 문제점들과 비교하여 검증을 수행하였다. 평가 대상 서비스는 MMS 서비스로써 멀티미디어 메시지를 발송하는 태스크를 대상으로 수행되었다. 사용자 테스트에 참가한 피실험자는 10 대 에서 20 대 사이의 중간 사용자들 5 명 이었다. 각 사용자들이 주어진 태스크를 수행하는 동안 에러를 범하거나, 불편함을 호소하는 경우 실제 사용성 문제점으로 판정하였다. UI 경력 5 년 이상의 평가자 2 인은 제안된 평가 체계를 통해 사용성 문제점들을 규명하였다. 결과를 바탕으로 제안된 평가 체계가 실제 사용성 문제들을 얼마나 빠짐없이 파악해 내는가(coverage)와 파악한 문제점들이 얼마나 실제적인 문제점인가(precision)를 검증하였으며, 두 척도의 정의는 다음과 같다.

- Coverage = 제안된 평가 체계로 찾아낸 실제 문제 수 / 사용자 테스트를 통해 발견한 실제 문제 수
- Precision = 제안된 평가 체계로 찾아낸 실제 문제 수 / 제안된 평가 체계로 찾아낸 전체 문제 수

검증 결과 실제 사용성 문제점 14 개 중 기술적인 내용 및 콘텐츠 부족의 문제를 제외한 총 12 개의 문제점에 대하여 제안된 평가 체계가 문제점을 발견하여 85.71%(12/14)의 coverage 를 보여주었으며, 또한 제안된 평가 체계로 발견한 총 20 개의 문제 중 12 개가 실제 사용자 테스트에서 발생한 문제점과 대응되어 60%(12/20)의 precision 을 보여 주었다.

### 5. 결론

본 연구에서는 사용성의 수준을 정량적으로 산정하기 위한 사용성 지수 평가 체계를 제안하고, 이를 S/W를 구성 하고 그 효과를 검증하였다. 제안된 평가 체계는 실제 무선 인터넷 서비스들의 사용성 평가에 적용되어 지속적으로 UI 품질 수준을 모니터링 하고 개선 요건을 도출하는 도구로

써 활용되었다. 현재 평가 체계는 평가자 별 결과의 변동성을 최소화 하기 위해 O/X로만 평가 항목을 판정케 하고 평가 항목별 가중치는 고려하지 않은 상태에서 지수를 산출하였다. 그러나 실제 고객이 체감하는 사용성 수준은 단순히 문제점의 개수가 아니라 문제가 가진 심각도에 따라 영향을 받을 수 있다. 그러므로 향후에는 각 평가 항목별 또는 문제점에 따라 사용자가 체감하는 어려움의 수준을 반영할 수 있는 평가 체계로 확장이 이루어질 필요가 있다.

## 참고문헌

- [1] Nielsen, J., Usability Engineering, Academic Press, Boston, 1993.
- [2] Kirakowski, J. and M. Corbett, "SUMI: The software usability measurement inventory" British Journal of Educational Technology, Vol. 24, No. 3, pp. 210~212, 1993.
- [3] Chin, J.P., Diehl, V.A., Norman, K.L., "Development of an instrument measuring user satisfaction of the human computer interface", Proc. of SIGCHI '88, New York, 1988.
- [4] Card, S.K., Moran, T.P., Newell, A., The psychology of human computer interaction, Lawrence Erlbaum Associates, Hillsdale, 1983.
- [5] Nielsen, J. and Mack RL, Usability inspection methods, Wiley, New York, 1994.
- [6] Preece, J., Rogers, Y., Sharp, H., Interaction design Beyond human computer interaction, Wiley, New York, 2002.
- [7] ISO/IEC 9241-11 Ergonomic requirements for office work with visual display terminals (VDT)s Part 11 Guidance on usability. ISO/IEC 9241-11 :1998.
- [8] Nielsen, J., Usability Inspection Methods, John Wiley and Sons, New York, 1994, pp. 25~62.
- [9] Dix, A., Finlay, J., Abowd, G., and Beale, R., Human-Computer Interaction, Prentice Hall, Upper Saddle River, 1998.
- [10] Constantine, L.L., "Collaborative Usability Inspections for Software", Proc. of the Software Development, 1994.
- [11] Lauessen, S., and Younessi, H., "Six Styles for Usability Requirements", Proc. Of the 4<sup>th</sup> Int. Workshop on Requirements Engineering: Foundations of Software Quality, pp.1~12, 1998.