

Tren Yolcu Bilet Sisteminde Bir Kullanılabilirlik Testi Uygulaması

Kadir Herkilođlu¹, Esen Özbayramođlu¹, Pınar Alptekin Binici¹,
Zeynep Gönülan¹, Yıldız Ebru Öktensoy¹, Volkan Arısoy¹, Burak Yolaçan¹

¹ HAVELSAN A.Ş. Mustafa Kemal Mah. 2120 Cad. No:39 Çankaya, Ankara, Türkiye
{kherkiloglu, eozbayramoglu, pabinici,
zgonulalan, eoktensoy, vkapkac, byolacan}@havelsan.com.tr

Özet. Kullanılabilirlik, ilgili sistemin, Uluslararası Standartlar Organizasyonu'nda (ISO-9241-11) verilen tanımı kapsamında son kullanıcı ile olan arayüzü üzerinden etkili, verimli ve tatmin edici bir biçimde kullanılabilir olma yeteneđi olarak tanımlanmaktadır. Bu çalışmada, HAVELSAN A.Ş. tarafından geliştirilen bir Tren Yolcu Bilet Sistemi yazılımı, söz konusu yeteneđin ne ölçüde açığa çıktığını belirlemek amacıyla, kullanılabilirlik testine tabi tutulmuştur. Bu bağlamda, test öncesi uygun profilde ve sayıda denekler seçilip, bu deneklere çeşitli görevler verilerek yazılımı kullanması ve sesli düşünme yönetimi ile geri dönüş yapılması sağlanmıştır. Test gözlemcilerinin elde ettiđi veriler kullanılarak ISO 9241-11 kapsamında ürün etkinliđi ile ürün verimliliđi ölçülmüş ve Sistem Kullanılabilirlik Ölçeđi (SKÖ) skor çalışması yapılarak kullanıcı memnuniyet seviyesi belirlenmiştir. Ayrıca, söz konusu yazılım ürününün iyileştirme ihtiyacı olabilecek potansiyel alanları analiz edilerek önerilerde bulunulmuştur.

Anahtar Kelimeler: Kullanılabilirlik, Kullanılabilirlik Testi, Tren Yolcu Bilet Sistemi

Usability Testing Application on a Train Passenger Ticket System

Kadir Herkilođlu¹, Esen Özbayramođlu¹, Pınar Alptekin Binici¹,
Zeynep Gönülalan¹, Yıldız Ebru Öktensoy¹, Volkan Arısoy¹, Burak Yolaçan¹

¹ HAVELSAN A.Ş. Mustafa Kemal Mah. 2120 Cad. No:39 Çankaya, Ankara, Turkey
{kherkiloglu, eozbayramoglu, pabinici,
zgonulalan, eoktensoy, vkapkac, byolacan}@havelsan.com.tr

Abstract. According to International Standards Organization (ISO-9241-11), usability is defined as the ability of the related system to be used via the interface between the system and the end user effectively, efficiently, and satisfactorily. In this work, usability testing is applied to a Train Passenger Ticket System developed by HAVELSAN A.Ş. to reveal the level of such ability. In this context, appropriate number of test subjects from different profiles is preliminary determined. They are asked to use the software by performing certain given tasks and to give feedback by using think-aloud method while doing the given tasks. Product effectiveness and efficiency according the ISO 9241-11 standard are measured by using the data acquired by the test observers and user satisfaction level is determined by doing a System Usability Scale (SUS) score study. Moreover, suggestions for the design team of the software are produced by analyzing the potential areas of the software product that need improvement.

Keywords: Usability, Usability Testing, Train Passenger Ticket System

1 Giriş

Kullanılabilirliğin en yaygın kullanılan operasyonel tanımı Uluslararası Standartlar Organizasyonu'nda (ISO) verilmektedir. ISO-9241-11 (1994)'e [1] göre kullanılabilirlik; "Bir ürünün, belirli bir kullanım bağlamında, belirli kullanıcılar tarafından, belirli amaçları gerçekleştirmek üzere, etkili (effective), verimli (efficient) ve tatmin edici (satisfactory) bir biçimde kullanılabilmesi"dir. Etkinlik; kullanıcıların amaçlarını ve görevlerini doğru ve tam olarak tamamlama düzeyleri, Verimlilik; amaçlara ve görevlere ulaşırken harcanan kaynakları (zaman/çaba), Memnuniyet ise kullanıcıların sistem kullanımı ile ilgili pozitif tutumları ve rahatlıkları ile ölçülmektedir.

Bu çalışmada, HAVELSAN A.Ş. tarafından geliştirilen Tren Yolcu Bilet Sistem yazılımı, kullanılabilirlik testine tabi tutulmuştur. Bu bağlamda, ilgili sistem kapsamındaki kullanılabilirlik testi, Etkinlik, Verimlilik ve Memnuniyet boyutlarında ölçümler alınmasında belirleyici olmuştur. Tipik olarak bir grup potansiyel kullanıcı adayı tespit edildikten sonra HAVELSAN A.Ş. kullanılabilirlik test mühendisleri tarafından ilgili test gerçekleştirilmiştir. Bu kapsamda kullanıcılardan bir dizi olağan olarak yapılagelen görevleri icra etmeleri istenmiştir. İlgili oturumların testlerinin kayıt altına alınmasından sonra sistem yazılımının geliştirme ve iyileştirme ihtiyacı olabilecek potansiyel alanları analiz edilmiştir.

Bu bildirin takip eden ikinci bölümünde uygulanan test metodolojisine değinilmektedir. Üçüncü bölüm, elde edilen test sonuçları ve bu sonuçların analizi üzerinedir. Son olarak, dördüncü bölüm bu deneyim sonrası ilgili sistem kapsamında ulaşılan sonuçları özetlemektedir.

2 Test Yöntemi

Test hazırlık aşamasında, kullanılabilirlik çalışması yapılacak yazılım arayüzü belirlendikten sonra, test sırasında kullanılacak araçlar, anketler, gizlilik sözleşmesini içeren denek rıza formu, hedef kullanıcı kapsamındaki denek grubu ve bu deneklerin yapacakları görevler (Görev 1: Bilet rezervasyonu, Görev 2: Rezerve edilmiş bileti satın alma, Görev 3: Rezerve yapmadan bilet satın alma, Görev 4: Bilet açığa alma) belirlenmiştir. Bu amaçla, 32 aday arasından trenle seyahat deneyimleri, eğitim düzeyi, cinsiyet, yaş vb. kriterler kapsamında hazırlanan Denek Aday Profil Anketi (Tablo 1 ve Tablo 2) kullanılarak hedef kullanıcı profiline puantaj olarak en çok yakınsayan 14 test deneği belirlenmiştir. Çalışma kapsamındaki ilgili nicel veriler ISO'nun tanımında yer alan etkinlik (görev tamamlama başarısı), verimlilik (görev tamamlama süresi bazlı kaynak kullanımı) ve memnuniyet kriterleri baz alınarak değerlendirilmiştir. Nitel veriler ise Sesli Düşünme Protokolü [2] ve gözlem aracılığıyla toplanarak işlenmiştir.

Test sırasında denek kullanıcıya sırayla görevler verilerek her görev için görevi başarıp başaramadığı, yaptığı hata sayısı ve süre verileri tutulmuştur. Test süresince deneğin sesli düşünmesi sağlanarak arayüzle ilgili nerelerde sorun yaşadığı ve arayüz hakkında ne düşündüğü öğrenilmiştir. Bu esnada deneğin geri bildirimleri test göz-

lemcileri tarafından gözlemci kayıt formu aracılığıyla kaydedilmiştir. Her görev sonrası derecelendirme anketi (ASQ-After Scenario Questionnaire) ve test bitiminde kullanıcının arayüz hakkındaki memnuniyetini öğrenmek için Sistem Kullanılabilirlik Ölçeği (SKÖ) / System Usability Scale (SUS) bazlı memnuniyet anketi uygulanmıştır. Test sonrasında, bütün veriler toplanarak analiz edilip kullanılabilirlik sorunları ile değişiklik/iyileştirme önerilerini de içerek şekilde rapor hazırlanmıştır.

Tablo 1. Denek Aday Profil Belirleme Anketi- I

Denek No	Cinsiyetiniz?	Eğitim Düzeyiniz?	Ne sıklıkla şehirlerarası seyahat edersiniz?	Trenle şehirlerarası ne sıklıkla seyahat edersiniz?	Seyahat sınıfı tercihiniz nedir?	Tren seyahatleriniz genellikle kaç saat sürüyor?
1	1	1	3	3	2	2
2	2	3	1	1	2	2
3	2	3	3	1	2	2
4	2	3	3	3	2	2
5	2	5	2	3	1	2
6	1	3	1	4	2	-
7	1	4	2	1	2	2
8	2	4	2	1	2	1
9	1	4	2	1	2	2
10	1	4	1	1	1	2
11	1	4	1	1	2	2
12	2	4	1	1	1	2
13	2	4	2	1	2	2
14	1	3	1	4	-	-

Tablo 2. Denek Aday Profil Belirleme Anketi- II

Denek No	Trenle ulaşım amacınız nedir?	Kişisel bilgisayarınızı internet erişimi için günde kaç saat kullanıyorsunuz?	İnternet erişimi olan kişisel bilgisayarınızı en çok hangi amaçla kullanıyorsunuz?	Kişisel bilgisayar işletim sistemlerinden hangilerini kullanma deneyiminiz bulunmaktadır?	Çalışmakta olduğunuz sektör hangisidir?	Yaş?
1	2	2	2,5,8,9	3	1	1
2	1,2	2	1,5,6,7,8,9	1,2,3,4,5	1	3
3	1	1	1,2,4,5,9	1,2,3,5	1	2
4	1	2	1,2,5,8	3,5	1	3
5	1	2	5,7,8,9	1,2,4,5	1	4
6	3	1	2,9	1,2,3,4,5	1	1
7	2	3	1,3,5,6,8,9	1,2	1	1
8	2	1	2,6,7,8,9	1,2,3,4,5	1	1
9	2	2	1,2,3,6,8,9	1,2,3,5	1	1
10	1,2	1	6,7,8,9	1,2,5	1	3
11	2	2	1,2,3,6,7,8,9	4	7	1
12	1	1	2,5,9	1,3,5	1	2
13	1	2	3,5	1,2,3,4,5	1	2
14	-	2	3,6,7,8,9	1,2,3	1	2

3 Test Sonuçları

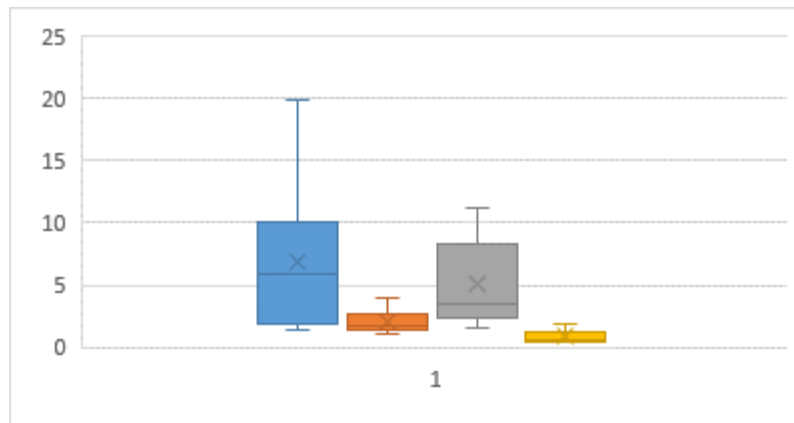
Tüm denekler (D-1 ~ D-14) “Görev-3” görevi dışındaki görevleri (Görev-1, Görev-2, Görev-4) %100 başarıyla tamamlamışlardır. Deneklerin yaklaşık %93’i Görev-3’ü başarıyla tamamlayabilmiştir. Görev-1’in deneklerin istedikleri tarih ve istedikleri yere bilet rezervasyonu yapmalarını gerektiren en uzun görev olduğu (ortalama 06:55) izlenmiştir. Bununla birlikte bu görevin tamamlanmasının yaklaşık 20 dakikadan (19:50) 1,5 dakikaya yakın (01:20) değişen sürelerde yapılabildiği görülmüştür. Görev 4’ün en kısa görev olduğu (ortalama 00:51) değerlendirilmiştir. Deneklerin görevleri tamamlamakta harcadıkları zaman dağılımı Şekil 1 ve Şekil 2’de verilmektedir.

ISO 9241’e göre etkinlik ölçeği, kullanıcıların uygulamayı kullanarak yapmaları beklenen görevleri ne düzeyde başarabildiklerini ifade etmektedir. Verimlilik ise belirlenen görevi yapmak için kullanıcı tarafından kullanılan kaynaklar baz alınarak ölçülmektedir. Bu kapsamda test edilen ürünün verimliliği, etkin deneklerin (görevi başarıyla tamamlayan) harcadığı zamanın tüm deneklerin (etkin olan ve etkin olmayan) harcadığı zamana oranı olarak tanımlanmaktadır. Bu bağlamda çalışmada kullanılan metriklerin formülasyon içerikleri sırasıyla Tablo 3 ve Tablo 4’de verilmektedir [3].

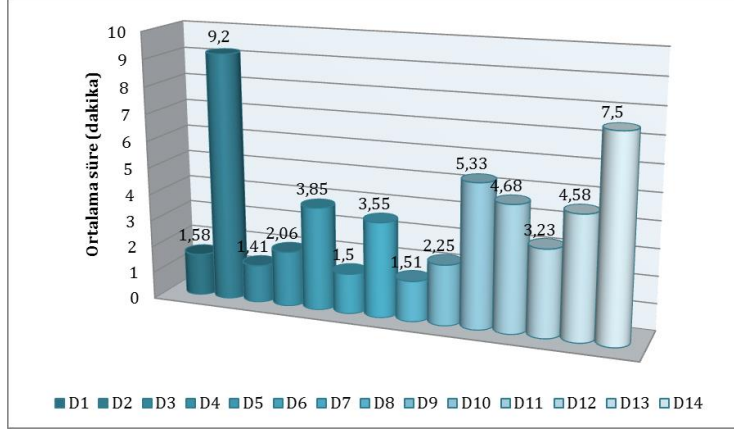
Test gözlemcilerinin elde ettiği veriler kullanılarak ISO 9241-11 kapsamında ürün etkinliği %98.21 ve ürün verimliliği %94.97 olarak ölçülmüştür. Ürün etkinliğinin harf cinsinden sınıflandırması için [6]’da verilen derecelendirme skalası kullanılabilir.

Deneklerin tüm görevleri tamamlamasından sonra uygulanan Sistem Kullanılabilirlik Ölçeği (SKÖ/SUS) [4] tabanlı memnuniyet anketinde çıkan sayısal sonuçlara uygulanan hesaplamalar üzerinden memnuniyet oranı bulunmuştur. İlgili anket 10 sorudan oluşmaktadır [Bkz. Tablo 5] ve 1-5 aralığında puanlama ile yapılmaktadır. Memnuniyet anketindeki sorulardan tek numaralı sorular olumlu, çift numaralı sorular olumsuz yapıda kurguludur.

Deneklerin ankete verdiği cevaplardan yola çıkarak hesaplama (Bkz. Tablo 6) [5] yapıldığında memnuniyet değerinin 73,21 olduğu hesaplanmıştır. Elde edilen değer [72.6 – 74.0] aralığında yer alırken uygulamanın kullanıcı nezdinde “B-“ sınıfı memnuniyet sağladığı değerlendirilmektedir (Bkz. Tablo 7) [6].



Şekil 1: Görev Süreleri Dağılımı



Şekil 2. Denek Bazlı Ortalama Görev Süreleri

Tablo 3. Etkinlik Metriği

N	Görev Sayısı	
R	Denek Sayısı	
n_{ij}	i numaralı deneğin j numaralı görevi başarıyla tamamlanma durumu	0: başarısız
		1: başarılı
Ürün Etkinliği (Product Effectiveness)	$\bar{E} = \frac{\sum_{j=1}^R \left(\sum_{i=1}^N (n_{ij}) \right)}{RN} * 100$	

Tablo 4. Verimlilik Metriği

N	Görev Sayısı	
R	Denek Sayısı	
n _{ij}	<i>i</i> numaralı deneğin <i>j</i> numaralı görevi başarıyla tamamlanma durumu	0: başarısız
		1: başarılı
t _{ij}	<i>i</i> numaralı deneğin <i>j</i> numaralı görevi tamamlanmak için harcadığı saniye cinsinden süre	
Zaman Bazlı Verimlilik (Time Based Efficiency)	$\bar{p} = \frac{\sum_{j=1}^R \left(\sum_{i=1}^N (n_{ij} t_{ij}) \right)}{\sum_{j=1}^R \left(\sum_{i=1}^N (t_{ij}) \right)} * 100$	

Tablo 5. SKÖ(SUS) Anket Soruları

Soru No	Soru
1	Bu ürünü sıklıkla kullanacağımı düşünüyorum.
2	Ürünü gereksiz bir şekilde karmaşık buldum.
3	Ürünün kolay kullanıldığını düşündüm.
4	Bu ürünü kullanabilmek için teknik bir kişinin desteğine ihtiyacım olabileceğini düşünüyorum.
5	Üründeki çeşitli fonksiyonları iyi entegre olmuş biçimde buldum.
6	Üründe çok fazla tutarsızlık olduğunu düşündüm.
7	Birçok insanın bu ürünü hızlı bir şekilde kullanabileceğini düşünüyorum.
8	Ürünün kullanımını çok hantal buldum.
9	Ürünü kullanırken kendimden emindim.

Tablo 6. Memnuniyet Metriği

R	Denek Sayısı
Q ⁺	Anketteki pozitif soru sayısı
Q ⁻	Anketteki negatif soru sayısı
p _{ij} ⁺	<i>j</i> deneği tarafından <i>i</i> numaralı (pozitif) soruya verilen cevabın ağırlığı [0..4]
p _{ij} ⁻	<i>j</i> deneği tarafından <i>i</i> numaralı (negatif) soruya verilen cevabın ağırlığı [0..4]
Ortalama Memnuniyet (Average Satisfaction)	$\bar{S} = \frac{\sum_{j=1}^R \left(\sum_{i=1}^{Q^+} \left(\frac{p_{ij}^+}{4} \right) \right) + \sum_{j=1}^R \left(\sum_{i=1}^{Q^-} \left(\frac{p_{ij}^-}{4} \right) \right)}{(Q^+ + Q^-) * R} * 100$

Tablo 7. SKÖ(SUS) Puanı Memnuniyet Sınıflandırması

SKÖ(SUS) Skor Değeri	Memnuniyet Derecesi
84,1 -100	A+
80.8 – 84.0	A
78.9 – 80.7	A-
77.2 - 78.8	B+
74.1 -77.1	B
72.6 – 74.0	B-
71.1 – 72.5	C+
65.0 – 71.0	C
62.7 - 64.9	C-
51.7 – 62.6	D
0.0 – 51.6	F

4 Sonuç

Bu çalışmada, HAVELSAN A.Ş. tarafından Tren Yolcu Bilet Sistemi yazılımına kullanılabilirlik testi uygulanmıştır. Çalışma için önce uygun profilde denekler seçilip, bu deneklere çeşitli görevler verilerek uygulamayı kullanması ve geri dönüş yapması sağlanmıştır. İlgili uluslararası standart (ISO-9241-11) kapsamında teste tabi yazılımın etkinlik ve verimlilik metrikleri ile ölçeklendirilmesi sağlanmış ve memnuniyet derecesi tespit edilmiştir. Test sırasında deneklerin yaşadıkları problemlere ve kullanıcı arayüzünde belirsizlik olduğunu düşündükleri noktalara dikkat çeken ve genel kullanım kolaylığını geliştirebilecek, üç görev kapsamında yüksek önem dereceli 16 adet değişiklik önerisi, dört görev kapsamında orta önem dereceli 24 adet değişiklik önerisi ve bir görev kapsamında düşük önem dereceli 5 adet iyileştirme önerisinde bulunularak yazılımının olgunlaştırılması yönünde katkı sağlanmıştır.

Kaynakça

1. ISO, SFSEN. "9241-11. 1998." Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)–Part II Guidance on Usability.
2. Boren, Ted, and Judith Ramey. "Thinking aloud: Reconciling theory and practice." IEEE transactions on professional communication 43.3 (2000): 261-278.
3. Alsabhan, Waleed. Designing a human-centred, mobile interface to support real-time flood forecasting and warning system. Diss. Brunel University London, 2017: 241-244.
4. System Usability Scale (SUS), 2018, <http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>.
5. Brooke, John. "SUS-A quick and dirty usability scale." Usability evaluation in industry 189.194 (1996): 4-7.
6. Sauro, Jeff, and James R. Lewis. Quantifying the user experience: Practical statistics for user research. Morgan Kaufmann, 2016: 42, 204.