

Ist Usability messbar? Wie kann man mit einfachen Mitteln die Usability einer komplexen Anwendung ermitteln.

Holger Kälble
FIDUCIA IT AG
Fiduciastr. 20
76227 Karlsruhe
holger.kaelble@fiducia.de
www.fiducia.de

Abstract

Mit der Version 3.0 des *agree*[®] BAP hat die FIDUCIA IT AG das Navigations- und Handhabungskonzept erheblich verbessert. Um die komplexe Anwendung mit über 500 Funktionseinstiegen und mehr als 3.000 Masken auch inhaltlich weiter zu optimieren, wurde eine Methode benötigt mit der Schwachstellen in der Usability schnell und einfach ermittelt und dargestellt werden können.

Um dieses Ziel zu erreichen entwickelte die FIDUCIA IT AG ein Prüfverfahren, den „UI-TÜV“ - ohne Berücksichtigung des fachlichen Kontextes.

Dieser Beitrag erläutert die Vorgehensweise und klärt die Frage, wie weit dies überhaupt möglich ist.

Keywords

Usability, Prüf-Verfahren, fachlicher Kontext

1.0 Einleitung

Aufgrund der Umstellung des Bedienkonzeptes auf die Version BAP 3.0 und der begleitenden Untersuchung mittels AttrakDiff™ wussten wir, dass die Verbesserung der Navigation nicht ausreicht. Im nächsten Schritt galt es die fachlichen Inhalte der Masken zu untersuchen und zu optimieren. Bei über 500 Funktionseinstiegen und insgesamt mehr als 3.000 Masken konnte dies nicht mit herkömmlichen Methoden erreicht werden. Die jeweilige Berücksichtigung des fachlichen Kontextes ist zwar Grundvoraussetzung für die Bewertung der Usability - allerdings war dies bei der hohen Komplexität der integrierten Anwendung nicht in einem akzeptablen Zeitrahmen möglich.

Hier stellte sich die Frage: „Ist Usability überhaupt messbar?“ und welche Kriterien müssen für eine standardisierte Methode herangezogen werden?

2.0 Vorgehensweise

2.1 Bestehende Prüfverfahren

„Leitfadenorientierte Prüfverfahren sind für Experten konzipiert, die sich dem System nicht hinsichtlich ihrer Eignung zur Aufgabenerfüllung nähern, sondern aus dem Blickwinkel der Software-Ergonomie. Ihr Vorteil liegt in der vergleichsweise schnellen Durchführbarkeit und der Tatsache, dass für ihre Anwendung kaum methodische Kenntnisse notwendig sind.“ (Hüttner, Wankde & Rätz, 1995)

Allerdings sind die Fragen von bekannten Verfahren wie z.Bsp. EVADIS II, teilweise offen gehalten, so dass die Antworten interpretiert und bewertet werden müssen.

Für uns ist es besonders wichtig, dass die Antworten einfach und ohne großen Aufwand auswertbar sind. Deshalb war es notwendig, die Fragen und auch die Antwortmöglichkeiten zu standardisieren.

2.2 Fragestellung

Zu Beginn stand die große Frage „Was fragen wir?“. Aber wie kann man Usability hinterfragen, wenn man von der Aufgabenstellung des Anwenders nichts weiß?

2.2.1 Rückschlüsse aus Usabilitytests

Zunächst einmal haben wir Usabilitytests herangezogen, die wir für die Anwendung *agree*[®] BAP bereits durchgeführt haben. Dabei ging es darum, Masken die von Anwender bei den Tests beanstandet wurden zu analysieren, welche Gemeinsamkeiten diese haben, um daraus den Umkehrschluss abzuleiten.

Beispiel: Viele der beanstandeten Masken aus den Usabilitytest zeichneten sich dadurch aus, dass sie sehr umfangreich waren und viele Karteikartenreiter hatten. So legten wir den Schluss nahe, dass Seiten mit vielen Karteikartenreitern potentiell eine schlechte Usability haben.

2.2.2 Ableiten von Gesetzmäßigkeiten

Ein zweiter Ansatz war das Ableiten von bekannten Gesetzmäßigkeiten,

Beispiel: Hick's law is a human-computer interaction model that describes the time it takes for a user to make a decision as a function of the possible choices he or she has. (http://en.wikipedia.org/wiki/hick_law, 12.06.2007)

Wir fragen hier nach, ob Masken übermäßig viele Datengruppen besitzen.

Weitere Gestaltgesetze, wie Gesetz der Nähe oder der Ähnlichkeit sind ebenfalls in das Prüfverfahren eingeflossen.

2.2.3 Einhaltung von Standards

Der dritte Ansatz verfolgte die Einhaltung unseres internen Fiducia-Styleguides und von Empfehlungen der DIN ISO 92 41 - 110.

2.3 Bewertung

Insgesamt ergab sich daraus ein Prüfkatalog mit 36 standardisierten Fragen. Dabei legten wir besonderen Wert darauf, dass die Kriterien allgemeingültig sind, damit die unterschiedlichen Masken miteinander verglichen werden können. Bei den Fragen wird das Nichtbeachten der Regeln und Vorgaben mit Fehlerpunkten gewertet, da der Schwerpunkt auf der Ermittlung von Schwachstellen liegt. Bonuspunkte für besonders gute Umsetzungen gibt es deshalb nicht.

Die Fragen haben wir in drei Prüfbereiche unterteilt: Interaktion, Darstellung und Standards.

2.3.1 Interaktion

Hier geht es vor allem um die Darstellung von Dialogen in der Interaktion, um Handlingsprobleme und korrekte Unterstützung bei Fehleingaben.

2.3.2 Darstellung

Hier werden z.B. Fragen nach Anzahl von Karteireitern, Detailmasken, und Spaltenanzahl bei Tabellen gestellt.

2.3.3 Standards

Wie schon unter 2.2.3 beschrieben, geht's es hier hauptsächlich um die Einhaltung von allgemeingültigen Standards und des anwendungsspezifischen Styleguides.

2.4 Gewichtung

Die drei Prüfbereiche werden gleichwertig gewichtet, weil es uns wichtig war, dass das Ergebnis transparent und nachvollziehbar ermittelt werden kann. Die Punkte werden mit einem Faktor für den Schwierigkeitsgrad multipliziert, der sich aus der Komplexität des Funktionseinstiegs ergibt. Weniger umfangreiche Masken werden dabei strenger bewertet, weil es hier weniger Möglichkeiten gibt, Fehler zu machen.

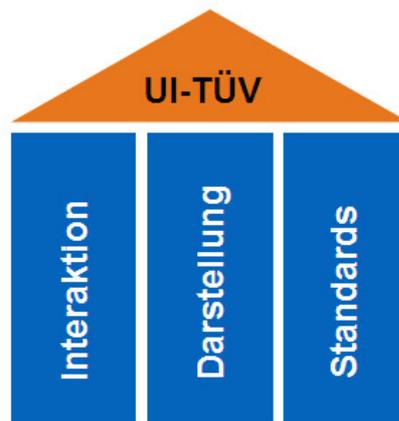


Abb. 1: Aufteilung der Prüfbereiche

2.5 Darstellung des Ergebnisses

Das Ergebnis wird in einfacher Form mit einer Ampel dargestellt, für jeden Prüfbereich eine. Aus den dadurch zehn möglichen Kombinationen

leitet sich dann die Gesamtnote ab. Diese reicht von „nicht bedienbar“ (dreimal rot) bis zu „hervorragend“ (dreimal grün).



Abb. 2: Darstellung der Bewertung des Funktionseinstiegs „Personendaten bearbeiten“

2.6 Einsatz

Mit dem Prüfverfahren „UI-TÜV“ messen wir in der Fiducia IT AG 2007 erstmals, ob die Entwicklungsabteilungen Ihre Ziele zur Verbesserung von ausgewählten Masken erreicht haben. Damit ist es uns gelungen, dieses Prüfverfahren innerhalb des Unternehmens zu etablieren.

3.0 Fazit

Der UI-TÜV ersetzt keinen Usabilitytest – er bietet lediglich eine Möglichkeit, eine oder mehrere Masken schnell, einfach und nachvollziehbar zu prüfen.

Bei der Prüfung konnte es teilweise passieren, dass Masken die positiv bewertet wurden, trotzdem erheblich fachliche Mängel enthalten haben. Diese können natürlich aufgrund des nicht berücksichtigten Kontextes nicht in das Ergebnis einfließen.

Wir haben bei den Prüfkriterien zwar sehr großen Wert auf Allgemeingültigkeit und Vergleichbarkeit gelegt, allerdings sind die Fragen zum größten Teil sehr konkret auf die Anwendung agree©-BAP der FIDUCIA IT AG ausgerichtet. Eine direkte Übertragbarkeit auf andere Anwendungen ist somit nicht möglich, al-

lerdings kann die Herangehensweise, Bewertung und Gewichtung allgemein verwendet werden.

Im Endeffekt konnten wir mit dem UI-TÜV Usability für die Fiducia IT AG messbar machen, weil es uns dadurch gelungen ist, ohne fachliche Diskussionen schnell und über viele Projekte hinweg Schwachstellen aufzuzeigen.

Literatur

Hüttner, Wanke, Rätz (1995): Benutzerfreundliche Software, Paschke-Verlag, Berlin

Sarodnick, Brau (2006): Methoden der Usability Evaluation, Verlag Hans Huber, Bern

http://en.wikipedia.org/wiki/hick_law,
12.06.2007

