

Autor(en): Geng, R., Chen, M., & Tian, J.

Titel: In-Process Usability Problem Classification, Analysis and Improvement

Jahr: 2014

Laufende Nr. im Literaturverzeichnis: 01

Schlagwörter: Usability, Defect classification and analysis, Software process, Human
Computer Interaction, Usability inspection and testing

Kurzbeschreibung: Das Paper stellt eine derzeit existierende Methode aus der Softwareentwicklung zur Fehlerauffindung, die *Orthogonal Defect Classification* (OCD), vor und zeigt, dass diese nicht ausreichend auf Usability Probleme eingeht. Aus diesem Grund hat das Team eine Art Gerüst bzw. Rahmen entwickelt, mit dem Usability Probleme aufgedeckt, klassifiziert und analysiert werden können. Darüber hinaus ist diese Methode nicht ausschließlich auf die Software-Branche beschränkt, sondern umfasst auch jegliche Art von *Human Computer Interface* (HCI). Um zu garantieren, dass Usability Probleme eingeordnet werden können, werden die Problemen in sieben Kategorien klassifiziert: Produktbezogene und aufgaben-bezogene Probleme, Lern-, Durchführungs- und Wahrnehmungsprobleme, sowie Attribute der Auswirkung des Fehlers und dessen Auslöser. Mit Hilfe dieser sieben Attribute kann eine Korrelation zwischen den Fehlern und des Entwicklungsprozesses aufgezeigt werden. Durch die Analyse und Zuordnung der Probleme kann außerdem eine quantitative Aussage über die Häufigkeit des Auftretens der vorliegenden Probleme getroffen werden.

Zur Evaluation und Validierung der entwickelten Methode bzw. des Problemgerüsts wurde eine Case Study, basierend auf 30 Usability Problemen, die mit Hilfe einer Heuristischen Evaluation, und 40 Problemen aus laborbasierten Usability-Tests, durchgeführt. Dazu wurde die Usability hinsichtlich der Vervollständigungsrate der Aufgabe, der benötigten Zeit und der Benutzerzufriedenheit analysiert. 20 Studenten führten zur Erhebung der Daten sechs spezifische Aufgaben in einer vorgegebenen Zeit durch. Nach Abschluss der Durchführung wurden die Teilnehmer zu den vollzogenen Aufgaben befragt. Zur Klassifizierung der Probleme wurden drei Experten mit dem Modell vertraut gemacht. Die Unterschiede der Ergebnisse von Experten und Studenten wurden im Anschluss gegenübergestellt.

Mit Hilfe dieses neu entwickelten Gerüsts werden Usability Probleme während des gesamten Entwicklungsprozesses identifiziert, klassifiziert und analysiert. Durch das Ergänzen des bereits bestehenden ODC, konnte das Anwendungsgebiet erweitert und eine genauere Spezifikation der Probleme erreicht werden. Jedoch sind noch Verbesserungen in Hinsicht auf die Klassifizierung der Probleme mit bestimmten Messmethoden erforderlich, um eine noch effizientere Methode zu schaffen.

Erstellt von: André Brandewiede (332283)

Autor(en): Tarkkanen K., Harkke V., & Reijonen P.

Titel: Are We Testing Utility? Analysis of Usability Problem Types

Jahr: 2015

Laufende Nr. im Literaturverzeichnis: 04

Schlagwörter: Usability problem, Utility problem, Problem classification, Usability testing

Kurzbeschreibung: Im Rahmen dieses Papers wurden Klassifizierungen von Usability

Problemen definiert, mit dessen Hilfe der Designprozess von Produkten durch direkteres Feedback verbessert werden soll. Das Ziel war die Verbesserung der Problemidentifikation, analyse und –berichterstattung. Aus drei empirischen Usability-Tests, welche mit offenen Testaufgaben durchgeführt wurden, wurde eine Kategorisierung von Problemtypen vorgenommen. Die dabei getesteten Systeme waren Werkzeuge aus der Arbeitsumgebung der Benutzer. Bei der Analyse wurde besonderes Augenmerk darauf gelegt den Unterschied zwischen Usability und Utility (Brauchbarkeit) zu verdeutlichen. Durch die gezielte Abtrennung von Usability und Utility Problemen ist es möglich den Produktdesignern in der spezifischen Produktentwicklungsphase zu helfen eine Optimierung durchzuführen. Die Datenerhebung wurde mit Prototypen aus der Gesundheitsbranche und deren späteren Nutzern durchgeführt. Zwei der Prototypen lagen als Papierprototypen vor und der dritte auf einem Tablet. Alle Tests wurden mit der Think-Aloud-Methode durchgeführt, wobei mindestens zwei Supervisor anwesend waren. Die Teilnehmerzahl schwankte zwischen vier und sechs Teilnehmern. Zusätzlich wurden Video- und Audioaufnahmen gemacht. Und alle gefundenen Probleme wurden dokumentiert und an den Designer/Entwickler weitergeleitet. Im Rahmen des Projektes wurden 173 Usability Probleme identifiziert und in elf Kategorien unterschieden. Um dies zu ermöglichen, fertigten Wissenschaftler Codes an, die im Nachhinein diskutiert wurden.

Für die einzelnen Kategorien entwickelte das Team spezifische Erklärung um eine bessere Zuordnung und Analyse zu ermöglichen. Für die einzelnen Fälle wurden die gefundenen Probleme klassifiziert und die Häufigkeiten ermittelt. Außerdem wurde unterschieden wie viele Probleme der Usability und wie viele der Utility zuzuordnen sind.

Diese neue Methode zur Klassifizierung hat den Vorteil, dass es sich nicht nur auf die Usability Probleme, sondern gleichzeitig auch auf die Utility Probleme konzentriert und diese gezielt unterscheidet. Dadurch kann eine Umgestaltung des Systems leichter verwirklicht und zugeordnet werden. Die meisten bisherigen Methoden berücksichtigen die Utility Probleme nicht getrennt voneinander. Durch die offene Testgestaltung werden jedoch mehr Utility Probleme produziert als bei einer vordefinierten Aufgabe.

Diese Klassifizierung ermöglicht außerdem gezielt Probleme im Design und der Handhabung von den Problemen des Arbeitsprozessablaufs zu unterscheiden.

Erstellt von: André Brandewiede (332283)

Autor(en): Vilbergsdóttir, S., Hvannberg, E. & Law, E.

Titel: Classification of Usability Problems (CUP) Scheme: Augmentation and Exploitation

Jahr: 2006

Laufende Nr. im Literaturverzeichnis: 05

Schlagwörter: Usability, defect classification, think-aloud, user test, validity, reliability, acceptance

Kurzbeschreibung: Das Paper baut auf dem von den gleichen Autoren designten Modell zur Klassifizierung von Usability Problemen (*Classification of Usability Problems*, CUP) auf.

Dieses Modell wurde entwickelt, um durch Aufdeckung und Klassifizierung von Usability Problemen die Produktqualität und eine Prozessverbesserung zu realisieren. Dies geschieht durch präzise und konkrete Informationen an die Produktentwickler. In dem Paper wird außerdem eine genaue Definition und Erklärung zu der Klassifizierung und derer Attribute vorgenommen. Um Fehler einer Benutzeroberfläche aufzudecken wurde ein Think-Aloud-User-Test durchgeführt. Die Beispiel-Benutzeroberfläche gehört zu einem Universitäts-Management-Tool, welches von Studenten und Lehrkräften gleichermaßen benutzt wird. Elf Teilnehmer, bestehend aus fünf Lehrkräften und sechs Studenten, führten die Tests durch. Die beiden Benutzergruppen mussten jeweils neun Aufgaben bewältigen, welche sich jedoch unterschieden. Während der Durchführung waren die Teilnehmer angehalten laut zu denken und sie wurden währenddessen aufgenommen. Dabei wurden die Zeit, die Häufigkeit von Frustrimenten und die Anzahl der Hilfestellungen gezählt. Insgesamt wurden 71 Usability Probleme identifiziert. Die Klassifizierung der Probleme wurde so angelegt, dass sie nicht länger als 90 Minuten andauerte. Für Wiederholbarkeits-Tests wurden acht weitere Teilnehmer in verschiedenen Altersgruppen und mit verschiedenen Knowhows ausgewählt. In dem ersten Termin erhielten die Teilnehmer eine Präsentation über das Thema Klassifizierung. und im zweiten mussten die Teilnehmer eine Klassifizierung von 21 Usability Problemen vornehmen. In den drei Tagen zwischen den beiden Terminen wurden die Teilnehmer gebeten, das System bereits kennenzulernen und sich mit der Materie auseinanderzusetzen. Zusätzlich zur Reliabilität wurde auch die Validität in einem zusätzlichem Test untersucht. Diese Tests wurden von zwei Software-Entwicklern durchgeführt. Die Daten wurden dabei in semi-strukturierten Interviews und Treffen in einer Zeitspanne von drei Monaten erhoben. Es wird deutlich, dass die Ergebnisse stark abhängig von den durchführenden Personen und deren Expertise sind. Ein Experte auf dem Gebiet identifiziert andere Usability Probleme als ein Laie und analysiert diese gegebenenfalls anders.

Insgesamt kann gesagt werden, dass die gefundenen Ergebnisse einer solchen CUP definitiv das Feedback für den Entwickler verbessern kann, sodass eine Evaluation und ein verbessertes Design möglich wird. Außerdem gibt die Studie einen guten methodischen Rahmen, um das System hinsichtlich Reliabilität und Validität zu untersuchen. Es wird jedoch auch klar, dass ein Training der Testpersonen notwendig ist, um die Usability Probleme korrekt einordnen zu können. Außerdem sollten die Gruppen von Experten und Novizen relativ gleichgroß sein, um eine bessere Beurteilung der Ergebnisse in Abhängigkeit der Expertise durchführen zu können. Eine Triangulationsstudie ist ein gutes Mittel, um die Feedbackgestaltung und Verbesserungsvorschläge für einen Entwickler zu verdeutlichen.

Erstellt von: André Brandewiede (332283)

Autor(en): Hornbaek, K., & Frokjaer E.

Titel: Comparing Usability Problems and Redesign Proposals as Input to Practical Systems Development

Jahr: 2005

Laufende Nr. im Literaturverzeichnis: 03

Schlagwörter: Usability evaluation, redesign, think aloud, metaphrs of human thinking, empirical study, usability inspection

Kurzbeschreibung: Im Rahmen der durchgeführten Studie wurden verschiedene Ansätze zur Überarbeitung von bestehenden Systemen verglichen und ausgewertet. Dabei wurden Techniken wie die Think-Aloud-Methode und die Heuristische Evaluation zur Problemfindung eingesetzt. Das Ziel der Autoren war es, die Unterschiede zwischen der Beschreibung von Usability Problemen und den Vorschlägen zur Überarbeitung des Systems hervorzuheben. Dazu bewerteten Entwickler die Ergebnisse, denn in den meisten Fällen werden eher quantitative Aussagen über gefunden Probleme geboten, aber weniger anwendbare Verbesserungen für den Entwickler. Die Methode bietet somit einen Ansatz nicht nur eine Evaluierung durchzuführen, sondern diese auch praktikabel anzuwenden. Die Usability Problembereiche wurden von 43 Studenten durchgeführt, die eine Jobportalseite aus

Dänemark untersucht haben. Um die „Methaphors of human thinking“- und die Think-Aloud-Methode vergleichen zu können, wandten zwei Gruppen von jeweils 21/22 Studenten je eine dieser Techniken an. Anschließend sollten die Studenten begründete Designverbesserungsvorschläge anfertigen. Die Bewertungen der Vorschläge wurden von vier Entwicklern vorgenommen und nach ihrer Nützlichkeit bewertet. Zusätzlich wurde betrachtet, wie häufig ein bestimmtes Problem auftrat und es wurde diskutiert inwiefern die Vorschläge eine notwendige Veränderung darstellen oder ob diese gar keinen Nutzen für das Portal haben. Die Entwickler wurden im Nachhinein in Einzelinterviews zu den Auswertungen befragt. Bei der Auswertung wurde festgestellt, dass zwischen den einzelnen Evaluierungstechniken kein bedeutender Unterschied vorliegt. Durch die Interviews wurde erkannt, dass die Entwickler bereits mit den Usability Problemen vertraut waren, aber die meisten von ihnen wurden zuvor aus Entwicklersicht als nicht wichtig erachtet. Entwickler sahen die Erklärungen der Studenten somit eher als Anreiz zur Priorisierung und weniger als neuen Informationsgewinn. Anders als die beschriebenen Usability Probleme konnten die Verbesserungsvorschläge großes Interesse bei den Entwicklern wecken. Sie wurden sogar als deutlich wichtiger erachtet als die eigentlichen Usability Probleme.

Mit ihrer Ausarbeitung hat das Team gezeigt, dass anders als bislang angenommen, nicht das Auffinden der Usability Probleme den Entwicklern hilft, sondern ausgearbeitete Verbesserungsvorschläge nachhaltiger und hilfreicher sind. Eine Kombination aus beiden Informationen ist somit besonders nützlich um Systeme optimale zu gestalten. Dieser Ansatz wird in den meisten derzeit existierenden Methoden nicht beachtet, wodurch das Paper nützliche Ergebnisse generieren konnten. Als Schwäche muss herausgehoben werden, dass die Techniken ausschließlich von Akademikern durchgeführt wurden, die jedoch nicht die einzigen Nutzer einer solchen Plattform sind.

Erstellt von: André Brandewiede (332283)

Autor(en): Ham, D.-H.

Titel: A New Framwork for Characterizing and Categorizing Usability Problems

Jahr: 2008

Laufende Nr. im Literaturverzeichnis: 02

Schlagwörter: Usability problem classification, Design Process, IT-Systems, usability inspection,

Kurzbeschreibung: In dem Paper von Ham wird ein Konzeptentwurf für einen Modell vorgestellt mit dem Usability Probleme systematisch klassifiziert werden können. Zusätzlich soll eine direkte Verbindung von den Usability Problemen zum Designprozess und den zu erfüllenden Aufgaben geschaffen werden. Dabei ist das Anwendungsgebiet auf die ITBranche beschränkt. Dazu wurden die bestehenden Klassifizierungsmethoden kurz vorgestellt und deren Kriterien aufgezeigt. Zusammenfassend wurde gesagt, dass unabhängig der Art der Klassifizierung, drei grobe Einteilungen existieren: Usability Tests, Usability Untersuchungen und Usability Inspektionen. Dabei wurde klargestellt, dass, je nach Vorhaben, eine bestimmte Datenerhebungsmethode ausgewählt werden sollte.

Außerdem wurde ein Zusammenhang zwischen Problemen und den Usability Evaluierungsstudien festgestellt, welcher verdeutlicht wie wichtig die Verbindung der Klassifizierung und dem Designprozess für die Qualität von Systemen ist.

Das vorgestellte Modell besteht aus drei Teilen und umfasst den Benutzungskontext, das Designwissen und die Designaktivitäten.

Nach Ham ist der Schlüssel für eine gute Usability von IT-Systemen das Auffinden von Usability Problemen und das gehaltvolle Feedback mit Verbesserungsvorschlägen für den Designprozess.

Das Vorgestellte Modell gibt zwar einen groben Rahmen vor, wie eine erfolgreiche Umsetzung von Verbesserungen umgesetzt werden kann. Allerdings gibt das Modell kein genaues Vorgehen vor, mit dem Entwickler an den Optimierungsprozess herangehen können. Deshalb ist zu bemängeln, dass ein reines Modell ohne durchgeführte Case Studies oder andere Versuche vorgestellt wird. Hingegen ist der Grundgedanken, dass der Designprozess und die Usability Probleme stark zusammenhängen und diese systematisch zusammen betrachtet werden sollten, richtig.

Erstellt von: André Brandewiede (332283)

Autor(en): Andre, T. S., Hartson, H. R., Belz, S. M. & McCreary, F. A.

Titel: The user action framework: a reliable foundation for usability engineering support tools

Laufende Nr. im Literaturverzeichnis: 06

Schlagwörter: user action framework; Usability-Evaluation; Evaluations-Methoden

Kurzbeschreibung: Im Paper wird das User Action Framework (UAF) vorgestellt, beschrieben und auf seine Reliabilität untersucht. Das UAF wird beschrieben als ein Rahmenkonzept für Usability-Testing und -Forschung, dass sich für verschiedene Tools (z. B. zur Klassifizierung, Evaluation oder Dokumentation) einsetzen lässt und so Ergebnisse der verschiedenen Tools vergleichbar und übersetzbar macht. Das UAF basiert auf der Struktur des „Theory of Action Model“ und soll eine gute Klassifizierung von Usability Problemen ermöglichen, wodurch im weiteren Verlauf ohne Verschwendung von Ressourcen passende Lösungen der Probleme produziert werden können. Es werden verschiedene Herangehensweisen an die Klassifizierung von Usability Problemen beschrieben. Weiterhin wird auf die Entwicklung des UAF eingegangen: Durch Modifizierung und Kombination des „Theory of Action Model“ sowie der „Usability Problem Taxonomy“ und der „Usability Problem Classifier“ entstand der Interaction Cycle, der die kognitiven und physischen Handlungen des Users in der Interaktion mit einer Maschine darstellt. Dabei wird dieser in drei Phasen

eingeteilt: *planning, physical action, assessment*. Die drei Phasen des Interaction Cycle finden sich im UAF wieder und bilden die oberste Ebene der (weitgehend) hierarchischen Struktur der Klassifizierung. Diese wird mit jedem Level spezifischer. Durch diese hierarchische Klassifizierung erfolgt eine systematische Einteilung der Probleme, wodurch sich diese sinnvoll vergleichen lassen. Das UAF stellt sich als reliable Klassifizierungsmethode heraus und erreicht bessere Reliabilität als die Klassifizierung nach Heuristiken und nach den zuvor vorgestellten Klassifizierungsansätzen. Lediglich auf der 6. Ebene der hierarchischen Struktur konnten keine besseren Werte in der Reliabilität im Vergleich zu anderen Verfahren gefunden werden.

Stärken und Schwächen: Als Stärke des UAF lässt sich zum einen die Genauigkeit der Klassifizierung nennen. Durch die hierarchische Struktur wird gewährleistet, dass es eine hohe Übereinstimmung in der Kategorisierung gibt und vor allem, dass andere Personen, wie zum Beispiel Entwickler, die im Verlauf die gefundenen Usability-Probleme beheben müssen, eine klare Vorstellung des Problems bekommen und es wenig Verständnis-Probleme oder Informations-Verluste durch eigene Interpretation etc. gibt. Eventuell als Schwäche lässt sich nennen, dass das Durchführen des UAF zeitaufwändig im Vergleich zu anderen Methoden ist.

Erstellt von: Anna-Magdalena Wisotzki, Matrikel-Nr. 384273

Autor(en): Keenan, S. L., Hatrson, H. R., Kafure, D. G. & Schulman, R. S.

Titel: The Usability Problem Taxonomy: A Framework for Classification and Analysis

Laufende Nr. im Literaturverzeichnis: 08

Schlagwörter: usability problem classification, usability problem analysis, problem prioritization

Kurzbeschreibung: Im Paper wird die Usability Problem Taxonomy (UPT) vorgestellt, mit der Usability-Probleme im GUI klassifiziert werden sollen. Die UTP wurde ausgehend von über 600 Usability-Problemen entwickelt, wobei aus diesen eine hierarchische Klassifikationsstruktur entwickelt wurde. Diese besteht aus den zwei Komponenten „artifact“ und „task“, die hierarchisch in primäre Kategorien und Subkategorien gegliedert sind. Die artifact-Komponente enthält die primären Kategorien visualness, language und manipulation, die task-Komponente task-mapping und task-facilitation. Bei der Klassifizierung eines Problems werden beide Komponenten unabhängig genutzt, d. h. ein Problem wird durch beide Komponenten klassifiziert, wobei die Autoren betonen, dass ein Problem komplett und genau beschrieben werden muss, um ausreichend klassifiziert werden zu können. Es wird ein UTP Web Tool beschrieben, das die Vorgehensweise der UTP, sowie die einzelnen Komponenten und Kategorien erklärt und mit welchem die Klassifizierung durchgeführt werden kann. Eine Reliabilitäts-Analyse ergab eine zufriedenstellende Übereinstimmung bei der Klassifizierung. Die UTP soll in drei Bereichen hilfreich sein: In der Untersuchung von Problem-Datensätzen (z. B. Lösen von mehreren Problemen gleichzeitig innerhalb eines Systems, Identifizierung von Problem-Typen & von globalen Problemen), zur Komplementierung anderer Analyse-Techniken (z. B. Problem-Schweregrad, Programm-Performance, User-Zufriedenheit) und zur Priorisierung der Probleme. Zum Abschluss beschreibt das Paper noch zwei reale Anwendungsbeispiele der UTP

Stärken und Schwächen: Schwäche: Einschränkung auf GUI

Erstellt von: Anna-Magdalena Wisotzki, Matrikel-Nr. 384273

Autor(en): Hvannberg, E. T. & Law, L.-C.

Titel: Classification of Usability Problems (CUP) Scheme

Laufende Nr. im Literaturverzeichnis: 07

Schlagwörter: usability problem classification, heuristic evaluation, user test

Kurzbeschreibung: Im Paper wird das Classification of Usability Problems (CUP) Scheme vorgestellt. Es ist gezielt dafür gedacht, Usability-Probleme aus Usability-Tests zu kategorisieren, um Entwicklern und Designern ein besseres Feedback für die Problembehebung geben zu können. Für die Kategorisierung der Probleme werden Usability-Tests nach mehreren Merkmalen ausgewertet (kurze Beschreibung, Aktivität zur Behebung des Problems, Trigger, Auswirkung, erwartete Phase der Behebung, tatsächliche Phase der Behebung, Qualität des Problems, Ursache, Fehler-Prävention), wobei bei fast allen Merkmalen Skalen oder Antwortmöglichkeiten vorgegeben sind. Die Phasen der User Interface Entwicklung werden dabei gesondert beschrieben, um eine präzise Einordnung zu gewährleisten. Zur Analysierung von Tests-Methoden, die mit dem CUP-Scheme ausgewertet werden können, wird eine Studie beschrieben, bei der heuristische Evaluation mit einem Cognitive Walkthrough verglichen wurde. Die Reliabilität einiger Merkmale war zwar gut, allerdings war sie insgesamt eher durchwachsen.

Stärken und Schwächen: Stärke: die Verständigung zwischen Usability-Testern und Entwicklern wird verbessert, wodurch gezieltere Änderungen des Interface durchgeführt werden können. Schwäche: die Beurteiler müssen gut geschult sein, um das CUP-Scheme gut anwenden zu können. Die Merkmale sind nicht immer intuitiv gut beantwortbar; das CUP-Scheme ist durch seinen Aufbau nur sehr eingeschränkt nutzbar

Erstellt von: Anna-Magdalena Wisotzki, Matrikel-Nr. 384273

Autor(en): Khajouei, R., Peute, L. W. P., Hasman, A & Jaspers, M. W. M.

Titel: Classification and prioritization of usability problems using an augmented classification scheme

Laufende Nr. im Literaturverzeichnis: 09

Schlagwörter: usability problem classification, & prioritization, user-computer interface

Kurzbeschreibung: Das Ziel des Papers ist es, ein Schema zu entwickeln, welches Usability-Probleme eindeutig klassifiziert und sie im gesamten Prozess des (Re)-designs nachvollziehbar macht, sie gemäß ihres Effekts auf die kognitive & physische Anstrengung der User einstuft und den potentiellen Effekt auf das Endergebnis der User-System-Interaktion berücksichtigt. Um diesen drei Ansprüchen gerecht zu werden, wird das Schema aus dem User Action Framework (UAF), einem Schweregrad-Ranking und einer Einschätzung des Einflusses auf das Endergebnis zusammengesetzt. Dabei dient das UAF der Kategorisierung der Probleme und die anderen beiden Maße der Priorisierung der Behebung. Das UAF berücksichtigt die Planung, Übersetzung, Ausführung und Bewertung der Aufgabe durch den User und Beurteilung können durch die quasi-hierarchische Kategorisierung die Art des Problems genau definieren. Bezüglich des Schweregrads des Problems werden die Frequenz, der Einfluss auf den User und die Persistenz berücksichtigt. Hier sollten allerdings vier

unabhängige Beurteiler herangezogen werden, um eine zufriedenstellende Einschätzung zu erreichen. Die Einschätzung des Einflusses auf das Endergebnis der Interaktion ist hier spezifisch auf das untersuchte System (Gesundheitswesen) ausgerichtet und dementsprechend wurden relevante Kategorien gebildet. Zur Einschätzung betrachten die Beurteiler alle Problembeschreibungen mit den zugehörigen System-Status. Für andere Systeme müssen die Kategorien neu definiert werden. Bezüglich der UAF wurden gute Inter-Rater-Reliabilitäten gefunden und durch die beiden zusätzlichen Einschätzungen konnte die Priorisierung der Problemlösung erleichtert werden.

Stärken und Schwächen: Stärke: zusätzliche Erhebung der Konsequenzen des Problems außerhalb der User-System-Interaktion – Dieser Punkt wird in konventionellen Usability-Tests nicht berücksichtigt, kann aber von großer Bedeutung sein, gerade bei System mit hohem Gefahr-Potential. Schwäche: Die Durchführung ist sehr zeitaufwändig

Erstellt von: Anna-Magdalena Wisotzki, Matrikel-Nr. 384273

Autor(en): Tarkkanen, K., Harkke, V. & Reijonen, P.

Titel: Are We Testing Utility? Analysis of Usability Problem Types

Laufende Nr. im Literaturverzeichnis: 04

Schlagwörter: usability problems, utility problems, classification, usability testing

Kurzbeschreibung: In dem Artikel steht die Differenzierung der Utility (= Funktionalität) von der Usability (Nutzbarkeit der Funktionalität) im Fokus, wobei diese stark zusammenhängen. Ohne System-Funktionen kann keine gute Usability erreicht werden und ohne gute Usability kann die potentielle Funktionalität des Systems nicht real erreicht werden. Es werden verschiedene Klassifikations-systems für Usability-Probleme vorgestellt (Orthogonal Defect Classification, User Action Framework, Classification of Usability Problems scheme). Durch einen Cognitive Walkthrough mit offenen Aufgaben an Prototypen-Programmen wurden Probleme aufgedeckt, abstrahiert und in Kategorien zusammengefasst (z. B. missing, misinterpreted, inadequate), unabhängig von bereits existierenden Klassifikations-Methoden. Dabei stellte sich heraus, dass sowohl Usability- als auch Utility-Probleme gefunden werden, was eine wichtige zusätzliche Informationsquelle im Design-Prozess eines Systems darstellen kann.

Stärken und Schwächen: Stärke: Erweiterung der Problem-Findung um Unterscheidung von Utility und Usability ist sehr nützlich, um zu entscheiden wie mit einem gefundenen Problem umgegangen werden kann. Schwäche: Die Vorgehensweise stellt kein neues System zur Untersuchung von Usability- oder Utility-Problemen dar, sondern liefert nur einen Anstoß zur weiteren Forschung. Die gewonnen Kenntnisse könnten in bestehende Systeme eingepflegt werden, um den Umgang mit gefundenen Problemen noch besser zu gestalten.

Erstellt von: Anna-Magdalena Wisotzki, Matrikel-Nr. 384273

Autor(en): Pataki, K. & Thüring, M.

Titel: Gewichtung von Usability-Kriterien

Jahr: 2007

Laufende Nr. im Literaturverzeichnis: 10

Schlagwörter: Usability-Scoring-Prozedur, Usability-Scores, Software-Evaluation, Analytisch hierarchischer Prozess, ISONORM 9241-110

Kurzbeschreibung: Das Paper bezieht sich auf die Frage, ob Nutzer Usability-Kriterien vor dem Hintergrund verschiedener Nutzungskontexte, Aufgaben und Ziele als verschieden relevant wahrnehmen. So wird darauf hingewiesen, dass eine Aggregation von Bewertung und Gewichtung zu einer drastischen Veränderung des Gesamturteils und der damit verbundenen Präferenzordnung führen kann. Als Grundlage für die Gewichtung dient die analytisch hierarchische Prozedur von Saaty. Die Kombination dieser Prozedur mit dem Fragebogen ISONORM 9241/110-Short führt zur im Paper vorgestellten „Usability Scoring Prozedur“. Die Studie wurde anhand von Websites aus den Nutzungskontexten Onlinebanking, -Reisereservierung und-auktion durchgeführt. Mit dem Fragebogen ISONORM 9241-110 wurde eine selbstgewählte Website bewertet, bevor die Gewichtung der Aspekte anhand der Analytisch hierarchischen Prozedur betrachtet wurde. Es stellte sich heraus, dass die verschiedenen Aspekte in Abhängigkeit des Nutzungskontexts als unterschiedlich relevant wahrgenommen wurden. Die unterschiedliche Gewichtung veränderte jedoch nicht das Gesamtranking der vergleichbaren Websites.

Stärken und Schwächen: Eine Schwäche der Studie, die auch in dem Paper selbst angesprochen wird, ist, dass wenig relevante Aspekte möglicherweise ein Gesamturteil gar nicht so sehr beeinflussen, selbst wenn die Gewichtung nicht beachtet wird. Das macht es schwer, eine Differenz zwischen den Ergebnissen einer Bewertung mit oder ohne Gewichtung zu finden. Probanden, die einen Usability-Aspekt generell wenig beachten, neigen möglicherweise dazu, diesen auch nicht mit extremen Skalenwerten zu bewerten, sondern die Bewertung eher an die restlichen Einschätzungen anzupassen. Beispiel: Wenn Aufgabenangemessenheit, Steuerbarkeit und Erwartungskonformität sehr gut sind, wird die Fehlertoleranz möglicherweise auch mit gut bewertet, auch wenn diese sehr schlecht ist. Das kann einerseits aus dem Grund geschehen, dass die Fehlertoleranz nicht stark beansprucht wird, da durch die hohe Erwartungskonformität eine geringe Fehlerquote und so ein unzureichendes Bild der Fehlertoleranz entstehen. Andererseits kann ein User, der einem Produkt bisher ausschließlich gute bis sehr gute Skalenwerte gegeben hat, Hemmungen empfinden, einem anderen Aspekt desselben Produkts plötzlich schlechte Werte zuzuteilen. Auf diese Weise erhält man sowohl mit als auch ohne Gewichtung ähnliche Resultate.

Erstellt von: Lena Anne-Katrin Mrowetz, Matrikel-Nr. 358601

Autor(en): Sachse, K. & Thüning, M.

Titel: ReMUS: Ein Usability-Scoring-Verfahren für rechnerbasierte Sicherheitstechnik

Jahr: 2011

Laufende Nr. im Literaturverzeichnis: 14

Schlagwörter: Rechnerbasierte Sicherheitstechnik, Usability-Scoring, Analytic Hierarchy Process, Softwaresicherheit, ISONORM 9241-110, Usability Evaluation

Kurzbeschreibung: In dem Paper wird die Entwicklung rechnerbasierter Sicherheitstechnik thematisiert. Das entwickelte, zum Zeitpunkt der Entstehung des Papers noch nicht validierte Tool basiert auf der Bewertung von Sicherheit und Gebrauchstauglichkeit. Die Sicherheitstechnik wird dabei zum einen mit Hilfe von entsprechenden Normen evaluiert, wobei die hier vorgegebenen Kriterien vor denen der softwareergonomischen Normen zu priorisieren sind. Daher wird bei der Entwicklung des Tools darauf Wert gelegt, nicht nur eine Messung, sondern auch eine Gewichtung der Kriterien zu ermöglichen. Neben den domänenspezifisch auszuwählenden Normen wird die Sicherheit auch auf Basis von Expertenfragebögen (eigene Entwicklung) bewertet. Die Benutzbarkeit wird aufgrund der ISO 9241-10 und der Ergebnisse von Nutzerfragebögen (z.B. ISONORM 9241/110-S) evaluiert und eine Gewichtung mit den Analytic Hierarchy Process von Saaty durchgeführt. Da Softwaresicherheit im Wesentlichen aus zwei Aspekten besteht („security“ im Sinne von Schutz/Sicherheit nach außen und „safety“ im Sinne von innerer Sicherheit), muss in der Entwicklung des Tools beides beachtet werden. Es wird festgestellt, dass die Beurteilungen durch User nicht ausreichen, da die fraglichen Elemente im normalen Anwendungskontext nicht auftreten. Laut Fazit müssten statt dessen Experten in einem simulierten, die relevanten Aspekte betreffenden Kontext die Sicherheit bewerten. Aus Wahrscheinlichkeit und Ausmaß des Schadens ergibt sich ein Risikowert für die Priorisierung.

Stärken und Schwächen: Durch die Verwendung des ISONORM-9241/110-S wird eine dimensionsspezifische Bewertung und Gewichtung erlaubt, was angesichts der möglicherweise lebensrettenden Eigenschaften der relevanten Software eine hohe Bedeutung hat. Leider gibt es bezüglich der sicherheitskritischen Aspekte derzeit keine solcher vorgegebenen Dimensionen. Es muss also darauf vertraut werden, dass das zu entwickelnde Instrument, welches auf Seite 4 angesprochen wird, mit Software-Sicherheitsexperten entwickelt wird, um eine ebenso praktische Einteilung zu gewährleisten. Erst so kann man gewährleisten, dass die Komplexität einer Abwägung von sicherheitsrelevanten Items/Dimensionen soweit herunter gebrochen wurde, dass eine konsistente und nachvollziehbare Entscheidung des Users möglich ist.

Erstellt von: Lena Anne-Katrin Mrowetz, Matrikel-Nr. 358601

Autor(en): Saaty, T.L.

Titel: How to Make a Decision. The Analytic Hierarchy Process

Jahr: 1990

Laufende Nr. im Literaturverzeichnis: 13

Schlagwörter: Analytic Hierarchy Process, Entscheidungen, Gewichtung

Kurzbeschreibung: Das Paper zum Thema Entscheidungsfindung geht auf die Frage nach einer absoluten Bewertung (im Gegensatz zur vergleichenden Wertung) ein. Zu diesem Zweck wird der „Analytic hierarchy process“ (AHP) vorgestellt, der sowohl die Rating als auch die vergleichende Methode einbezieht. Die auf einer neunstufigen Skala basierende Methode ist begründet durch die Feststellung, dass die menschliche Intuition in Bezug auf die Bewertung und Einstufung komplexerer Optionen kein ausreichendes Instrument ist. So werden die Ansprüche an einen geeigneten Lösungsansatz vorgestellt sowie die Bedingungen, unter denen Entscheidungen getroffen werden können. Der Prozess der Entscheidungsfindung wird in sechs Schritte untergliedert. Die Stärke des menschlichen Urteilsvermögens, übersichtliche

Probleme hinreichend zu bewerten, wird genutzt, um aus der Vielzahl der Unterproblemlösungen die Gesamtlösung abzuleiten. Auf die Bedeutung der Modellbildung des Problems wird hingewiesen. Kosten und Nutzen werden separat behandelt. Im Verlauf des Entscheidungsprozesses wird jeweils darauf hingewiesen, dass die gewählten Kriterien auf Basis ihrer relativen Gewichtung in die Gesamtbewertung einfließen. Ebenso spielt die Konsistenz der Entscheidungen eine wichtige Rolle, da ab einem gewissen Maß an Inkonsistenz der AHP nicht mehr aussagekräftig genug ist. Als Maß dient der CR (consistency ratio). Als Kritik am AHP wird die Umkehrung der Präferenzen und Rankings aufgeführt.

Stärken und Schwächen: Die sehr umfangreiche Methode des AHP hat den wesentlichen Vorteil, dass die natürliche Stärke des Menschen genutzt wird, kleine übersichtliche Probleme zu lösen und Entscheidungen zu fällen. Diese Intuitivität gekoppelt mit dem logischen Aufbau der Methode stellt eine wesentliche Stärke dar. Die Unterteilung in Nutzen und Kosten ist ebenfalls eine Stärke, da die Komplexität des Problems herunter gebrochen wird. Eine Schwierigkeit, die die Methode nicht abgreift, ist die Modellierung des Problems. Es wird vorausgesetzt, dass alle Aspekte des Problems bekannt sind und berücksichtigt werden. Einen Teil dieser Problematik findet sich in dem Unterpunkt, bzw. Anforderung der Konsistenz wieder, die für den Erfolg der Methode vorausgesetzt wird. Die Methode bezieht sich also ausschließlich auf Probleme, dessen Optionen sowie deren Kosten und Nutzenaspekte bereits bekannt sind.

Erstellt von: Lena Anne-Katrin Mrowetz, Matrikel-Nr. 358601

Autor(en): Pataki, K.

Titel: Usability-Scoring auf Basis multiattributer Entscheidungsverfahren, S. 25-38

Jahr: 2009

Laufende Nr. im Literaturverzeichnis: 11

Schlagwörter: Analytic Hierarchy Process, Usability Scoring Procedure, Entscheidungen

Kurzbeschreibung: Das Kapitel 1.3 „Usability Scoring Procedure“ in der Doktorarbeit von Kristin Pataki behandelt die Gewichtung der Aspekte des Befragungsinstruments ISONORM 9241/10. Die Gewichtung wird dabei anhand des Analytic Hierarchy Process (AHP) vorgenommen. Auf Grundlage der sieben Gestaltungsgrundsätze der DIN EN ISO 9241-10 ergeben sich insgesamt 35 itemspezifische Usability-Aspekte. Diese itemspezifischen Aspekte werden hierarchisch gewichtet. Dazu werden sowohl die dimensions- als auch itemspezifischen Aspekte paarweise verglichen. Zu diesem Zweck werden Nutzer befragt. Aufgrund der einfacheren Gestaltung wird eine bottom-up-Gewichtung vorgeschlagen (zunächst Item-, dann Dimensionsbewertung). Aus den jeweiligen Paarvergleichen können Wichtigkeitsurteile erstellt werden, die auf Grundlage der 42 Usability-Aspekte (7 Dimensionen, 35 Items) relative Größen darstellen (Verhältnisse der Aspekte zueinander). Die Abbildung der Wichtigkeitsurteile geschieht in einer Paarvergleichsmatrix mit dem Wert 1 auf der Diagonalen und auf der unteren Hälfte dem reziproken Wert der oberen Urteile. Voraussetzung für den Erfolg dieses Näherungsverfahrens ist, dass der User sich in seinen Urteilen weitestgehend konsistent zeigt. Dies wird anhand eines Konsistenzindex bestimmt. Als Kritik an AHP wird aufgeführt, dass die Saaty-Skala bei sich sehr stark unterscheidenden Relativbedeutungen per se zu inkonsistenten Wichtigkeitsurteilen führt. Da Alternativvorschläge nicht zu einem zufriedenstellendem Ergebnis geführt hatten, wurde für

das Usability Scoring Procedere dennoch die neunstufige Saaty-Skala übernommen. Die Gesamtgewichtung ergibt sich aus der Multiplikation der itemspezifischen mit den dimensionsspezifischen Gewichten. Dadurch werden Items, die zu einer weniger bedeutsamen Dimension gehören, in Relation weniger gewichtet als Items aus bedeutsameren Dimensionen. Zur Überführung der Wichtigkeitsurteile in Usability-Scores (USP-Werte) sind mehrere Aggregationsstufen zu durchlaufen. Nach der Erhebung von Wert- und Wichtigkeitsurteilen folgen die Normierung itemspezifischer Werturteile, die Berechnung der gewichteten itemspezifischen Werturteile und anschließend die Berechnung des Usability-Scores.

Stärken und Schwächen: Das gesamte Verfahren basiert auf der ISONORM 9241/10, was den Vorteil hat, dass ein bereits erprobter und bewährter Fragebogen verwendet wird. So erübrigt sich die Frage nach angemessenen Kriterien. Das Kapitel 1.3 der Doktorarbeit beleuchtet, wie groß die Herausforderung der guten Auswahl einer korrekten Skala ist. An dieser Stelle weist die Studie eine Schwäche auf, da, falls große Unterschiede zwischen den Einschätzungen durch den User liegen, da die Skala dann zu Inkonsistenz führt. Dies ist allerdings auch ein allgemeiner Kritikpunkt am verwendeten AHP. Wie ebenfalls deutlich gemacht wird, ist die Konsistenz der User-Bewertungen von entscheidender Bedeutung. Es muss also dafür gesorgt werden, dass dem User ein umfassendes und klares Bild von den zu untersuchenden Objekten hat sowie eine ausreichende Kenntnis von Usability-Aspekten, um die erfassten Bewertungen und Gewichtungen anwenden zu können. Über die bottom-up-Vorgehensweise werden die sonst unvoreingenommenen User an die Thematik herangeführt und ausreichende Kenntnisse der zur Orientierung dienenden Kenntnisse gewährleistet.

Erstellt von: Lena Anne-Katrin Mrowetz, Matrikel-Nr. 358601

Autor(en): Tergan, S.

Titel: Qualitätsbeurteilung von Bildungssoftware mittels Kriterienkatalogen. Problemaufriss und Perspektiven

Jahr: 2001

Laufende Nr. im Literaturverzeichnis: 15

Schlagwörter: Bildungssoftware, Kriterienkataloge, Anwendungsprobleme

Kurzbeschreibung: In dem Artikel von Sigmar-Olaf Tergan werden verschiedene Anforderungen und Probleme an auf Bildungssoftware bezogene Kriterienkataloge thematisiert. Es wird klargelegt, wofür Kriterienkataloge erstellt werden und welchen Ansprüchen sie gereichen müssen. Sowohl im summativen als auch formativen Kontext können diese Kataloge im Rahmen der Qualitätssicherung eine unterstützende Funktion erfüllen. Das Paper bezieht sich nicht auf allgemeine Probleme wie der unmittelbaren Entwicklung und Auswertung von Fragebögen und Kriterienkatalogen. Es wird vielmehr auf konzeptuelle, inhaltliche, praktische und formale Aspekte eingegangen. Unterschieden werden des Weiteren die Problemtypen Konstruktionsmängel/Konzeptuelle Schwächen, Schwächen bezüglich der Gütekriterien und Anwendungsprobleme. Unter dem Begriff der Konzeptuellen Schwächen werden Probleme wie Probleme der Standardisierung, Übergewicht der technischen Kriterien, die Dominanz eindimensionaler Items und ähnlichem erläutert. Als Schwächen bezüglich der Gütekriterien gelten beispielsweise eine geringe Beurteilerübereinstimmung und unklare Validität und Gewichtung der Kriterien. Als

Anwendungsprobleme werden unter anderem Unhandlichkeit, mangelnde Gerichtetheit der Kriterien für spezifische Verwendungszwecke und fehlende oder strittige Bewertungs- und Gewichtungsverfahren beschrieben und kritisiert. Als Fazit wird eine generelle Eignung von Kriterienkatalogen für eine differenzierte Beschreibung der Merkmale von Bildungssoftware festgestellt. Dies kann jedoch nicht auf die Frage nach der Eignung der Software für gute Lerneffekte ausgeweitet werden, da situative Bedingungen der Software häufig vernachlässigt wird. Neben der Verbesserung und Anpassung der Kriterienkataloge an den jeweiligen situativen und technischen Stand wird eine kombinierte Anwendung von Expertenbeurteilungs- und Evaluationsverfahren vorgeschlagen.

Stärken und Schwächen: Es ist schwer, in einem Paper, das sich auf das Aufzeigen von Problemen fokussiert, Stärken und Schwächen zu finden. Es handelt sich um keine Studie oder Methode, zu deren Umsetzung oder Konzept etwas gesagt werden könnte. Die Theorie, die hinter diesem Paper steckt, wird nicht vorgestellt, sondern untermauert durch die aufgezeigten Probleme. Die vorgestellten Probleme werden selbstverständlich nicht in vollem Maß auf jede einzelne Bildungssoftware zutreffen. Jedoch sind sie auch größtenteils nicht nur auf Bildungssoftware beschränkt, sondern kann auch auf andere Software und dementsprechend entwickelte Kataloge angewandt werden.

Erstellt von: Lena Anne-Katrin Mrowetz, Matrikel-Nr. 358601

Autor(en): Pataki, K.

Titel: Usability-Scoring auf Basis multiattributer Entscheidungsverfahren, S. 39-45

Jahr: 2009

Laufende Nr. im Literaturverzeichnis: 12

Schlagwörter: Analytic Hierarchy Process, Usability Scoring Procedure, Entscheidungen, universelle Gewichtungen

Kurzbeschreibung: Im Kapitel 1.4 wird zunächst die Frage nach dem Einfluss der dimensionsspezifischen Gewichte auf die Identifikation wichtiger itemspezifischer Aspekte gestellt. Es wird Wert darauf gelegt, dass die itemspezifischen Gewichte nur dann hinreichende Parameter darstellen, wenn die dimensionsspezifischen Aspekte ähnliche Gewichtungen aufweisen. Daher wird die Frage diskutiert, ob dimensionsspezifische Gewichte die Identifikation relevanter, itemspezifischer Aspekte steuern können. Es stellt sich heraus, dass dimensionsspezifische Gewichte die Identifikation der itemspezifischen Aspekte beeinflussen. Daher können itemspezifische Gewichte nur unter der Bedingung, dass alle dimensionsspezifischen Aspekte eine gleiche oder ähnliche Gewichtung haben, allein für eine Diagnose zu realisierender Aspekte dienen. Ist dies nicht der Fall, muss ihre Bedeutung auch weiterhin über die Gewichtung der Dimensionen relativiert werden. Nachfolgend wird das Potential der universellen Gewichte thematisiert. Es wird die Möglichkeit aufgezeigt, bei ähnlichen Fällen über Sensitivitätsanalysen einen erneuten Evaluationsprozess zu vermeiden, da die dimensions- und itemspezifischen Gewichte als voreingestellte Parameter behandelt werden. Als problematisch für den Bewertungsansatz wird die Tatsache empfunden, dass das multiattribute Entscheidungsverfahren für individuelle Entscheidungen entwickelt wurde. Diese sind für die Evaluation interaktiver Produkte wenig bedeutsam, da bei derartigen Aufgaben stets eine größere Menge User befragt und deren Werte gemittelt werden. Dies hat auch Auswirkungen auf die Präferenzunabhängigkeit und Redundanzfreiheit, die theoretisch

vorhanden sein sollten. Da beide Anforderungen für den vorliegenden Fall nicht gewährleistet werden können, werden Präferenzunabhängigkeit und Redundanzfreiheit zugunsten von guter Reliabilität und Validität zurückgestellt.

Stärken und Schwächen: Die auch in der Arbeit thematisierte Schwäche der Methode liegt in der Zurückstellung der Präferenzunabhängigkeit und Redundanzfreiheit. Dies geschieht aufgrund der der Validität und Reliabilität und kann daher nur schwer bis gar nicht umgangen werden, falls man nicht den Aufwand der Entwicklung eines neuen Fragebogens mit den Qualitäten des ISONORM 9241-10 auf sich nehmen möchte. Durch die Generalisierung der Gewichtungen ergibt sich die Möglichkeit einer breiteren Anwendung, wobei ein Nachteil der Methode, nämlich der Aufwand für die Probanden 91 Paarvergleiche durchzuführen, verringert wird.

Erstellt von: Lena Anne-Katrin Mrowetz, Matrikel-Nr. 358601

Autor(en): Nielsen, J.

Titel: Severity Ratings for Usability Problems

Jahr: 1995

Laufende Nr. im Literaturverzeichnis: 16

Schlagwörter (3-5): Usability Probleme, Gewichtung, Heuristische Evaluation

Kurzbeschreibung: Die Gewichtung von Usability Problemen ergibt sich nach Nielsen aus der Kombination folgender drei Faktoren: Der Häufigkeit des auftretenden Problems, der Auswirkungen des auftretenden Problems auf den Nutzer, sowie der Fortdauer des Problems. Des Weiteren sind die Auswirkungen auf den Markt zu beurteilen, da manchmal Probleme von geringer Schwere eine große negative Auswirkung auf die Popularität eines Produktes haben können. Bei der Gewichtung empfiehlt Nielsen eine 5-stufige Skala (0=kein Usability Problem, 5= katastrophales Usability Problem), welche alle oben genannten Faktoren umfasst. Nach Nielsen sind die Gewichtung der Probleme von drei Experten ausreichend um reliable Ergebnisse zu erhalten. Nach Nielsen dauert die Gewichtung nach seiner hier beschriebenen Vorgehensweise in heuristischen Evaluationen nur etwa 30 Minuten pro Tester. Sie ist somit eine schnelle und kostengünstige Möglichkeit zur Priorisierung von Usability Problemen. Leider bleibt unklar, unter welchen Gesichtspunkten die oben genannten drei Faktoren zu einer Gesamtskala zusammenzufügen sind. Des Weiteren beruht die Anzahl von 3 Evaluatoren für eine ausreichende Reliabilität der Gewichtungen lediglich auf dem Erfahrungswert von Nielsen. Etliche Studien zeigen jedoch, dass die Korrelation zwischen den Gewichtungen unterschiedlicher Evaluatoren eher gering ausfällt.

Erstellt von: Lion Kochan, 314352

Autor(en): Rubin, J., Chisnell, D.

Titel: Handbook of Usability Testing, Kapitel 11, S. 261 - 263

Jahr: 2008

Laufende Nr. im Literaturverzeichnis: 17

Schlagwörter (3-5): Usability Probleme, Gewichtung, Schwere des Problems, Auftretenswahrscheinlichkeit des Problems

Kurzbeschreibung: Rubin und Chisnell schlagen für die Gewichtung von Usability Problemen eine Kombination aus der Schwere des Problems und dessen Auftretenswahrscheinlichkeit vor. Die Schwere des Problems wird anhand einer 4-stufigen Skala bestimmt (1 = kosmetisches Problem, 5 = Unbenutzbarkeit eines Programmfeatures). Anschließend wird geschätzt, wieviel Prozent der Benutzergruppe von diesem Problem betroffen sein werden und mit welcher Wahrscheinlichkeit dieses Problem in der betroffenen Gruppe auftreten wird. Diese beiden Werte werden nun miteinander multipliziert und deren Produkt ergibt die Auftretenswahrscheinlichkeit, welche in einer 4-stufigen Skala einem Wert zugewiesen wird (1 = Auftretenswahrscheinlichkeit < 10%, 5 = Auftretenswahrscheinlichkeit > 90%). Die aus der Schwere-Skala und der Auftretenswahrscheinlichkeits-Skala ermittelten Werte werden nun miteinander addiert. Das Problem mit der größten Schwere sowie Auftretenswahrscheinlichkeit bekommt somit die höchste Priorisierung bei der Fehlerbehebung. Im Vergleich zu der Skala von Nielsen (siehe Literatur 1) wird in diesem Verfahren konkret erläutert, wie die beiden verwendeten Faktoren (Schwere und Auftretenswahrscheinlichkeit) in die Gewichtung der Usability-Probleme einfließen. Jedoch ist die Bestimmung der Schwere ebenso wenig erläutert wie bei Nielsen, mit dem Unterschied, dass die von Nielsen beschriebenen Faktoren „Auswirkungen des auftretenden Problems auf den Nutzer“ (siehe Literatur 1) sowie „Fortdauer des Problems“ (siehe Literatur 1) nicht einmal erwähnt werden. Darüber hinaus kann davon ausgegangen werden, dass die Auftretenswahrscheinlichkeits-Skala, da sie allein aufgrund einer Schätzung berechnet wird, sehr fehlerbehaftet ist und somit keine hohe Korrelation zwischen mehreren Evaluatoren zu erwarten ist.

Erstellt von: Lion Kochan, 314352

Autor(en): Molich, R., Jeffries, R.

Titel: Comparative expert reviews: CHI 2003 workshop description.

Jahr: 2003

Laufende Nr. im Literaturverzeichnis: 18

Schlagwörter (3-5): Usability Probleme, Gewichtung, Häufigkeit des Problems, Auswirkungen auf den Nutzer, Fortdauer des Problems

Kurzbeschreibung: Molich und Jeffries nennen in ihrem Paper für die Gewichtung von Usability Problemen die drei Faktoren von Nielsen: Häufigkeit des Problems, Auswirkungen auf den Nutzer sowie Fortdauer des Problems. Im Gegensatz zu Nielsen bildet jeder Faktor eine eigene 3-stufige Skala. Dabei konkretisieren die Autoren für jede Skala, was sie unter dem jeweiligen Faktor verstehen. Für die Häufigkeit wollen sie wissen, wie oft das gefundene Problem auftritt. Dabei kann zwischen selten (1), häufig (2) und manchmal (3) gewählt werden. Bei dem Faktor der Auswirkungen auf den Nutzer wollen sie wissen, wie schwer das Problem ist, wenn es auftritt. Wird der Nutzer nur kurz beeinflusst (1), gibt es signifikante Verzögerungen (2) oder kann gar die Aufgabe nicht durchgeführt werden (3). Beim Faktor der

Fortdauer des Problems fragen sie, ob die Nutzer lernen das Problem zu umgehen. Dabei kann gewählt werden zwischen „Ja, sehr schnell“ (1), „nur nach mehreren Versuchen“ (2) oder „Nein“ (3). Zur Erstellung einer Gewichtung aller Usability-Probleme empfehlen die Autoren die Punkte in den drei Skalen für ein Usability Problem zusammen zu addieren und die Probleme anhand ihrer höchsten Summen zu sortieren. Die Skala von Molich und Jeffries kann folglich als Weiterentwicklung der Skala von Nielsen (Literatur 1) gesehen werden, da sie im Vergleich zu Nielsen zuerst jeden Faktor auf einer Skala bewerten und die drei Skalen im Anschluss zu einem Wert zusammenfassen.

Im Vergleich zu Rubin und Chisnell decken sie mit den drei Faktoren ein größeres Spektrum ab (Rubin und Chisnell -> Schweregrad und Auftretenswahrscheinlichkeit). Die 3-stufigen Skalen von Molich und Jeffries bieten gegenüber 4- oder 5-stufigen Skalen weniger Unterscheidungsmöglichkeiten und teilweise sind die Antwortoptionen verwirrend bzw. unpräzise (Faktor Häufigkeit -> selten, häufig, manchmal). Auch die Beschreibungen der Faktoren wirken zu kurz geraten. Vergleicht man die Konzeption des Faktors „Häufigkeit des Problems“ von Molich und Jeffries mit dem Faktor „Auftretenswahrscheinlichkeit“ von Rubin und Chisnell, so wird deutlich, dass Molich und Jeffries diesen Faktor in meinen Augen zu oberflächlich betrachten.

Erstellt von: Lion Kochan, 314352

Autor(en): Sim, G., & Read, J. C.

Titel: The damage index: an aggregation tool for usability problem prioritisation.

Jahr: 2010

Laufende Nr. im Literaturverzeichnis: 19

Schlagwörter (3-5): Usability Probleme, Gewichtung, Heuristische Evaluation, Damage Index

Kurzbeschreibung: Sim und Read präsentieren in ihrer Studie den Damage Index als Möglichkeit zur Priorisierung von Usability Problemen. Der Damage Index eines Usability Problems wird berechnet mit Hilfe des Mittelwertes der Schwere über alle Evaluatoren (als Beispiel wird die Skala von Nielsen genannt, wobei die Berechnung der Schwere nicht an eine bestimmte Skala gebunden ist) multipliziert mit der Anzahl der Evaluatoren (oder Gruppen von Evaluatoren), die den Fehler während der heuristischen Evaluation gefunden haben. Das entstandene Produkt wird nun geteilt durch das Produkt aus dem höchsten Wert der Schwere-Skala (im Falle der Nielsen-Skala wäre dies eine 4) und der Gesamtanzahl an Evaluatoren (bzw. der Gruppen). In der Skala von Rubin und Chisnell wurde die Priorisierung der Usability Probleme aus der berechneten Schwere und einer theoretisch berechneten Auftretenswahrscheinlichkeit bestimmt. Der Damage Index bezieht ebenfalls die Schwere mit ein, verwendet als zweiten Faktor jedoch nicht die theoretische, sondern die empirische Auftretenshäufigkeit des Problems. Im Vergleich zu Rubin und Chisnell muss hier also kein Wert für die Auftretenshäufigkeit geschätzt werden, sondern dieser lässt sich direkt aus den Ergebnissen der Evaluatoren bestimmen. Durch die verwendete Formel lassen sich zudem subjektive Ausreißer korrigieren. Hat beispielsweise eine Minderheit von Evaluatoren ein Problem bedeutend anders bewertet als die restliche Gruppe, so wird dies über die Damage Index Formel entsprechend korrigiert. Zudem bringt der Damage Index Vorteile, wenn es eine große Anzahl an Problemen gibt, die in der Schwere-Skala den gleichen Wert bekommen haben. Durch das Hinzuziehen der Häufigkeit des Problems lässt sich eine weitere

Gewichtung innerhalb eines Wertebereiches einer Schwere-Skala vornehmen. Dies zeigen Sim und Read in einer durchgeführten „Single heuristic Evaluation“ und einer „Multiple Heuristic Evaluation“. Kritisch anzumerken ist, dass bei der Bestimmung der Schwere den Evaluatoren keine Skala vorgegeben wird, sondern diese frei wählbar ist. Was jedoch dazu führen kann, dass bei Verwendung unterschiedlicher Skalen große Unterschiede bei der Zuordnung der Probleme zu einer entsprechenden Stufe auftreten. Die hierbei entstandenen Verzerrungen sind nun Berechnungsgrundlage für den Damage Index. Unklar bleibt auch, wieso nur die Auftretenshäufigkeit als extra Faktor definiert wird und weitere Faktoren wie die Fortdauer des Problems oder die Auswirkungen auf den Benutzer keine Beachtung finden. Des Weiteren muss im Anschluss der durchgeführten heuristischen Evaluation auf subjektiver Basis entschieden werden, welche der genannten Probleme auf ein und dasselbe Problem verweisen. Dieser Wert ist entscheidend bei der Berechnung des Damage Index und unterliegt aufgrund seiner subjektiven Bestimmung einer großen Unsicherheit. Der Damage Index bietet in meinen Augen trotz der Kritikpunkte eine sinnvolle Alternative zum Modell von Rubin und Chisnell, da es durch die Verwendung der empirischen Auftretenshäufigkeit eine bessere Anwendung für die Praxis garantiert, als die Verwendung der theoretischen Auftretenswahrscheinlichkeiten, wie bei der Skala von Rubin und Chisnell. Im Vergleich zu der Skala von Molich und Jeffries bezieht sie jedoch zu wenig Faktoren mit ein und bietet durch die freie Wahl der Schwere-Skala viel Raum für Verzerrungen bei der Bestimmung des Damage Index.

Erstellt von: Lion Kochan, 314352

Autor(en): Wilson, C.

Titel: Defining, categorizing, and prioritizing usability problems: UPA 2003 Idea Market.

Jahr: 2003

Laufende Nr. im Literaturverzeichnis: 20

Schlagwörter (3-5): Usability Probleme, Gewichtung, Priorisierung

Kurzbeschreibung: Die in diesem Paper von Wilson vorgestellten Aspekte zur Gewichtung von Usability Problemen sind in einer sogenannten „Idea Market Session“ der UPA (Usability Professionals' Association) im Jahre 2003 von einer Gruppe von Usability Experten gemeinsam diskutiert und die entsprechenden Ergebnisse von Wilson in diesem Paper zusammengetragen worden. Zu allererst werden hier verschiedene Faktoren besprochen, welche die Gewichtung von Usability Problemen beeinflussen. Diese sind wesentlich umfangreicher als die bisher genannten Faktoren. Neben den bisher genannten (Häufigkeit, Auswirkungen, Fortdauer etc.) Problemen werden vor allem unternehmensrelevante Faktoren genannt, wie zum Beispiel ob das gefundene Problem ein zentrales Business-Ziel beeinflusst, ob das Problem dem Ansehen der Firma schaden könnte, ob das Problem zu einem Datenverlust führt, ob Software-Entwickler bereit sind, das Problem zu lösen oder ob das Problem rechtliche Konsequenzen für die Firma haben könnte. Des Weiteren gibt das Paper Anregungen, wie eine höhere Inter-Rater Reliabilität erreicht werden kann. Neben beispielsweise einem Training für alle Evaluatoren, wie genau auftretende Probleme zu gewichten sind, schlägt Wilson selbst vor, ein formales Dokument anzufertigen, in welchem die Stufen der Schwere-Skala definiert und mit Beispielen praxisnah erläutert werden. Die von

Wilson beschriebene 4-stufige Skala von leicht (4) bis schwer (1) unterscheidet sich zwar nicht sehr stark von Skalen wie der von Nielsen, dafür wird jede Stufe sehr genau definiert und mit etlichen praxisrelevanten Beispielen beschrieben. In meinen Augen ist die von Wilson beschriebene Skala in Verbindung mit den genannten Faktoren die praxisnaheste Lösung von allen bisher vorgestellten Verfahren zur Gewichtung von Usability Problemen. Sie ist nicht nur die faktenreichste Ausarbeitung, sondern bietet auch bei der Beurteilung der Schwere des Problems die detailreichste Beschreibung. Dennoch muss ich kritisch anmerken, dass keinerlei Richtlinien genannt werden, wie die unterschiedlichen Faktoren zu kombinieren oder zu gewichten sind. Eine Formel wie beim Damage Index vermisste ich hier gänzlich.

Erstellt von: Lion Kochan, 314352

Autor(en): Dumas, J. & Redish, G.

Titel: A practical guide to usability testing

Jahr: 1999

Laufende Nr. im Literaturverzeichnis: 21

Schlagwörter (3-5): Usability Probleme, Gewichtung, Priorisierung, Reichweite, Schwere

Kurzbeschreibung: Dumas und Redish gewichten Usability Probleme nach zwei Dimensionen. Die erste Dimension bezeichnen sie als Reichweite des Problems, die entweder lokal oder global vorliegen kann. Dabei wird dem Beheben von globalen Problemen eine größere Wichtigkeit zugeschrieben als den lokalen Problemen. Als ein lokales Problem bezeichnen sie beispielsweise einen undeutlich beschriebenen Menü-Eintrag. Bei einem globalen Problem wäre die gesamte Menüführung verwirrend und Unterpunkte generell schwer aufzufinden. Das Erkennen globaler Probleme erfordert laut Dumas und Redish einen gewissen Grad an Erfahrung, da diese sonst nur als Ansammlung lokaler Probleme wahrgenommen werden. Die zweite genannte Dimension der Autoren ist die der Schwere des Problems. Als Vorschlag liefern sie eine 4-Stufige Skala (1 = Probleme verhindern die Durchführung einer Aufgabe, 4 = Probleme von geringer Schwere, die meist auf eine Erweiterung des Produktes hindeuten). Im Anschluss werden alle nach den 2 Dimensionen bewerteten Problemen zuerst nach ihrer Schwere sortiert und anschließend in jeder Schwere-Stufe nach globalen und dann nach lokalen Problemen sortiert. So ergibt sich für Probleme der Schwerstufe 1 und einer globalen Reichweite die höchste Priorität. Die zwei Dimensionen zur Gewichtung von Usability Problemen wurden gut beschrieben und die genaue Gewichtung klar geschildert. Unklar bleibt jedoch, aus welchen Faktoren sich die Schwere des Problems zusammensetzt, also nach welchen Kriterien die Probleme in die genannten 4 Stufen eingeteilt werden. Die Autoren geben zwar zu jeder Stufe eine kurze Beschreibung sowie ein Beispiel, aus den obigen Literaturzusammenfassungen wird jedoch deutlich, wie vielschichtig die Dimension der Schwere in der Praxis ist. Allein in dem Paper von Wilson (Literatur 4) werden über 10 Faktoren genannt, nach denen die Schwere eines Usability Problems beschrieben werden kann. Dumas und Redish erwähnen jedoch auch, dass statt der von ihnen genannten Skala auch eine firmeninterne Schwere-Skala verwendet werden kann und der zweite Faktor (Reichweite) somit die bereits vorhandene Gewichtungs-Skala sinnvoll ergänzen kann. In Unternehmen, in denen bereits eine Skala zur Bestimmung der Schwere besteht, kann der

Ansatz von Dumas und Redish durchaus verwendet werden, um das bestehende System um die Dimension Reichweite zu erweitern bzw. zu verbessern.

Erstellt von: Lion Kochan, 314352

Autor(en): Foraker Labs

Titel: Defining the Usability Problem

Jahr: 2015

Laufende Nr. im Literaturverzeichnis: 22

Schlagwörter (3-5): Usability Probleme, Problemraum, Anwendungsfälle, Geschäftsziele

Kurzbeschreibung: Der Artikel besagt, dass man das Usability Problem zunächst erkennen und klar definieren muss, bevor man eine Webseite oder Applikation verbessern kann. Dazu muss verstanden werden, wie und wofür das Produkt genutzt werden soll, was durch Gespräche mit Domäne Experten und Investoren bewerkstelligt werden kann. Auch heuristische Evaluationen können Usability Probleme aufdecken. Indem man erkennt, was der Benutzer mit dem Produkt erreichen will, kann der Problemraum genauer definiert werden. Anhand der wichtigen und häufig auftretenden Aufgaben, werden Anwendungsfälle festgelegt, mithilfe derer dann das Produkt optimiert werden kann. Die Anwendungsfälle können sich aus Beobachtung und Befragung der Nutzer ergeben.

Stärken und Schwächen: Laut Artikel sollen die Funktionen und Nutzbarkeit des Produkts an den Geschäftszielen ausgerichtet sein. Da die Geschäftsziele von Entscheidungsträgern wie Managern, Marketingspezialisten und Investoren festgelegt werden und nicht von Usability-Experten, könnte die Gebrauchstauglichkeit darunter leiden. Allerdings sollte es im Interesse der Entscheidungsträger liegen, wenn das Produkt optimal von der antizipierten Zielgruppe angenommen wird.

Erstellt von: Martin Krabbe 384180

Autor(en): Lavery, D., Cockton, G., Atkinson, M.

Titel: Comparison of evaluation methods using structured usability problem reports

Jahr: 1997

Laufende Nr. im Literaturverzeichnis: 23

Schlagwörter (3-5): Usability Problem, Definition, Problem Report

Kurzbeschreibung: Das Paper stellt eine neue Report-Struktur für Usability Probleme vor, die bessere Übereinstimmung zwischen analytisch erwarteten Problemen und den Problemen, die tatsächlich in empirischen Nutzungstest auftreten, herstellen soll, da bisherige Ansätze, die untersucht wurden, die keine adäquaten Übereinstimmungen erzielen konnten. Die neuentwickelte Definition für Usability Problem, auf der die Report-Struktur aufbaut, lautet sinngemäß übersetzt: „*Ein Usability Problem ist ein Aspekt des Systems und/oder eine Anforderung an den Nutzer, die es unangenehm, ineffizient, beschwerlich oder unmöglich für den Nutzer macht, seine Ziele in typischen Nutzungssituation zu erreichen.*“ Das Problem wird

dazu in vier Komponenten aufgeschlüsselt: (1) Die *Ursache* (z.B. ein Design-Fehler), (2) ein möglicher *Misserfolg/Missverständnis* (Breakdown) in der Interaktion, (3) die daraus *resultierende Folge* und (4) der *Kontext*, in dem es passiert. Durch eine Designänderung kann die Ursache behoben, der Breakdown vermieden und die Folge verändert werden. Basierend auf der Definition wurde ein strukturiertes Report Format erstellt, welches sich bei Tests als Verbesserung gegenüber den zuvor untersuchten Report Formaten erwies.

Stärken und Schwächen: Die Definition liefert eine gute Grundlage, um Usability Problemen von Problemen anderer Art abzugrenzen. Die für den Report vorgenommene Unterscheidung zwischen Breakdown in der Interaktion und resultierenden Folgen, findet sich jedoch nicht klar in der Definition wieder.

Erstellt von: Martin Krabbe 384180

Autor(en): Manakhov, P., Ivanov, V.

Titel: Defining Usability Problems

Jahr: 2016

Laufende Nr. im Literaturverzeichnis: 24

Schlagwörter (3-5): Usability Probleme, Definition, Evaluation, Anforderungen

Kurzbeschreibung: Die Autoren des Papers stellen fest, dass viele der bisher genutzten Definitionen von Usability Problemen zu unspezifisch sind und nicht alle relevanten Aspekte aus Human-Computer-Interaction-Sicht abdecken. Eine gute Definition und somit Klarheit über das Problem sei jedoch essentiell für die Evaluation. Die Definition soll von nicht-HCI-basierten Problemen abgrenzen, des Weiteren für Vergleichbarkeit und Einheitlichkeit sorgen und gut an Lernende vermittelbar sein. Die im Paper festgelegten fünf Anforderungen an die Definition sind im Hinblick auf bisherige Defizite: (1.) Alle HCI Phänomene einbeziehen, (2.) Usability Probleme von Problemen anderer Art abgrenzen, (3.) das Problem und seine Ursache separat benennen, (4.) klarstellen, dass eine Kombination von Faktoren das Problem verursachen kann und (5.) die Definition von Lösungsvorschlägen trennen. Die aus den Anforderungen abgeleitete Definition lautet:

«A usability problem is a set of negative phenomena, such as user's inability to reach his/her goal, inefficient interaction and/or user's dissatisfaction, caused by a combination of user interface design factors and factors of usage context.»

Die Ursache (Kombination der Faktoren) und die Wirkung (negative Phänomene in Effektivität, Effizienz und Zufriedenheit) sind somit differenziert dargestellt und treten in jedem Usability Problem auf.

Stärken und Schwächen: Die von den Autoren entwickelte Definition deckt die selbst gestellten Anforderungen umfassender ab, als Definitionen anderer Autoren, die zum Vergleich herangezogen wurden. Ob damit alle denkbaren Anforderungen zum Thema abgedeckt sind bleibt fraglich, ebenso ob die Ursache in jedem Fall sowohl in mindestens einem Interface Design Faktor als auch einem Faktor des Nutzungskontexts liegen muss.

Erstellt von: Martin Krabbe 384180

Autor(en): Nielsen, J.

Titel: Usability 101: Introduction to usability

Jahr: 2012

Laufende Nr. im Literaturverzeichnis: 25

Schlagwörter (3-5): Usability, Definition, Verbesserung

Kurzbeschreibung: Der Artikel definiert Usability anhand von fünf Qualitätskomponenten. Diese sind Erlernbarkeit, Effizienz, Merkbarkeit, Fehlerverhalten und Zufriedenheit. Ein weiteres wichtiges Qualitätsmerkmal sei Utility, welches beschreibt, ob die vorhandenen Features das bieten, was der Nutzer benötigt. Usability und Utility geben an, wie brauchbar (useful) das Produkt ist. Usability ist insbesondere im Webauftritt wichtig, da der Nutzer leicht auf Konkurrenzangebote ausweichen kann. Zur Verbesserung der Usability werden User-Tests empfohlen, bei dem repräsentative Nutzer unter Beobachtung repräsentative Aufgaben durchführen. Durch frühzeitige und per iterativem Design mehrfach durchgeführte Tests können Usability Probleme erkannt und verbessert werden. Dabei können zu Beginn schlicht Papier-Prototypen eingesetzt werden.

Stärken und Schwächen: Bei relativ kostengünstigem Aufwand (ca. 10% des Budgets), lassen sich durch Usability-Maßnahmen deutliche Verbesserungen der Qualitätsmetriken erzielen. Der Aufwand hält sich in Grenzen, da in der Regel fünf Testpersonen pro Zwischenversion ausreichen und die Tests direkt im Unternehmen durchgeführt werden können.

Erstellt von: Martin Krabbe 384180

Autor(en): Sarodnick, F., Brau, H.

Titel: Usability-Probleme

Jahr: 2016

Laufende Nr. im Literaturverzeichnis: 26

Schlagwörter (3-5): Usability Problem, Definition, Domäne-Berücksichtigung, Offensichtlichkeit

Kurzbeschreibung: Das Kapitel stellt fest, dass eine Definition von Usability Problemen diese klar von anderen Problemarten abgrenzen muss. Dies betrifft etwa Probleme bei Personen, die nicht zur Zielgruppe des Systems gehören, wenn dieses speziell für Domäneexperten konzipiert ist. Des Weiteren werden Probleme ausgegrenzt, die bei unsachgemäßer Benutzung auftreten. Als Definition ergibt sich daraus, angelehnt an der von Lavery et al. (1997) : „*Ein Usability-Problem liegt vor, wenn Aspekte eines Systems es Nutzern mit hinreichender Domänenerfahrung unangenehm, ineffizient, beschwerlich oder unmöglich machen, in einem typischen Anwendungskontext die Ziele zu erreichen, für deren Erreichung das System erstellt wurde.*“ Darüber hinaus wird anhand von zwei Beispielen noch aufgezeigt, dass schwerwiegende Usability-Probleme nicht unbedingt leichter zu entdecken sind als Leichte. Die Offensichtlichkeit ist somit nicht direkt von der Schwere des Problems abhängig.

Stärken und Schwächen: Die vorgebrachte Definition schließt bestimmte nicht-relevante Problemarten aus und erweitert damit die Definition von Lavery et al. (1997). Jedoch schließt die Definition keine Anforderungen an den Nutzer als Problemursachen mit ein.

Erstellt von: Martin Krabbe 384180

Literaturverzeichnis:

- (01) Geng, R., Chen, M., & Tian, J. (2014). In-Process Usability Classification, Analysis and Improvement. In *14th International Conference on Quality Software* (pp. 240 - 245).
- (02) Ham, D.-H. (2008). A New Framework for Characterizing and Categorizing Usability Problems. In *EKC2008 Proceedings of the EU-Korea Conference on Science and Technology* (pp. 345 – 352).
- (03) Hornbaek, K., & Frokjaer E. (2005). Comparing Usability Problems and Redesign Proposals as Input to Practical Systems Development. *Methods & Usability*, 391 - 400
- (04) Tarkkanen K., Harkke V., & Reijonen P. (2015) Are We Testing Utility? Analysis of Usability Problem Types. In *Design, User Experience, and Usability* (pp. 269 - 280) Los Angeles: DUXU.
- (05) Vilbergsdóttir, S., Hvannberg, E. & Law, E. (2006). Classification of Usability Problems (CUP) Scheme: Augmentation and Exploitation. *NordiCHI*, 281 - 290.
- (06) Andre, T. S., Hartson, H. R., Belz, S. M., & McCreary, F. A. (2001). The user action framework: a reliable foundation for usability engineering support tools. *International Journal of Human-Computer Studies*, 54(1), 107-136.
- (07) Hvannberg, E. T., & Law, L.-C. (2003). Classification of Usability Problems (CUP) Scheme. *Interact*, 655—662.
- (08) Keenan, S. L., Hatrson, H. R., Kafure, D. G., & Schulman, R. S. (1999). The Usability Problem Taxonomy: A Framework for Classification and Analysis. *Empirical Software Engineering*, 4, 71-104.
- (09) Khajouei, R., Peute, L. W. P., Hasman, A., & Jaspers, M. W. M. (2001). Classification and prioritization of usability problems using an augmented classification scheme. *Journal of Biomedical Informatics*, 44, 948-957.
- (10) Pataki, K. & Thüring (2007). Gewichtung von Usability-Kriterien. In *Prospektive Gestaltung von Mensch-Technik-Interaktion* (pp. 417-422). Berliner Werkstatt Mensch-Maschine-Systeme. Berlin.
- (11) Pataki, K. (2009). Usability-Scoring auf Basis multiattributer Entscheidungsverfahren. Die Gewichtung von Aspekten der software-ergonomischen Qualität (pp. 25-38). Dissertation. Technische Universität Berlin. Berlin.
- (12) Pataki, K. (2009). Usability-Scoring auf Basis multiattributer Entscheidungsverfahren. Die Gewichtung von Aspekten der software-ergonomischen Qualität (pp. 39-45). Dissertation. Technische Universität Berlin. Berlin.

- (13) Saaty, T. L. (1990). How to Make a Decision. The Analytic Hierarchy Process. *European Journal of Operational Research*, 48, 9-26.
- (14) Sachse, K. & Thüring, M. (2011). ReMUS: Ein Usability-Scoring-Verfahren für rechnerbasierte Sicherheitstechnik. In *Berliner Werkstatt Mensch-Maschine-Systeme*, Vol. 7. Berlin.
- (15) Tergan, S. (2001). Qualitätsbeurteilung von Bildungssoftware mittels Kriterienkatalogen. Problemaufriss und Perspektiven. *Unterrichtswissenschaft* 29 (4), 319-341
- (16) Nielsen, J. (1995). Severity ratings for usability problems. Internetressource: <https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/> [18.12.16].
- (17) Rubin, J., & Chisnell, D. (2008). Analyze Data and Observations. In *Handbook of usability testing : How to plan, design, and conduct effective tests* (2.nd ed.) (pp. 245–268). Indianapolis, IN: Wiley.
- (18) Molich, R. & Jeffries, R. (2003). Comparative expert reviews: CHI 2003 workshop description. Internetressource: http://www.dialogdesign.dk/tekster/cue4/cue4_proposal.pdf [18.12.16].
- (19) Sim, G., & Read, J. C. (2010). The damage index: an aggregation tool for usability problem prioritisation. In *Proceedings of the 24th BCS Interaction Specialist Group Conference* (pp. 54-61). British Computer Society.
- (20) Wilson, C. (2003). Defining, categorizing, and prioritizing usability problems: UPA 2003 Idea Market. July 11, 2003. Internetressource: https://web.archive.org/web/20030808033626/http://www.upassoc.org/conferences_and_events/upa_conference/2003/usability%20problems_wilson.doc [18.12.16].
- (21) Dumas, J. & Redish, G. (1999). Tabulating And Analyzing Data. In *A practical guide to usability testing* (pp. 309-330). London: Intellect Books, Revised Edition.
- (22) Foraker Labs (2015). Defining the Usability Problem. Internetressource: <http://www.usabilityfirst.com/about-usability/defining-the-usability-problem/> [10.12.16].
- (23) Lavery, D., Cockton, G., Atkinson, M. (1997). Comparison of evaluation methods using structured usability problem reports. *Behaviour & Information Technology*, 16 (4/5), 246-266.
- (24) Manakhov, P., Ivanov, V.D. (2016). Defining Usability Problems. In *2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 3144-3151). New York: ACM.
- (25) Nielsen, J. (2012). Usability 101: Introduction to usability. Internetressource: <https://www.nngroup.com/articles/usability-101-introduction-to-usability/> [09.12.16].
- (26) Sarodnick, F., Brau, H. (2016). Usability-Probleme. In *Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung* (pp. 25-26). Bern: Hogrefe Verlag.