

Quantitative Methoden zur Messung von User Experience

Martin Schrepp
User Experience
SAP SE
Walldorf, Germany
martin.schrepp@sap.com

Bernard Rummel
User Experience
SAP SE
Walldorf, Germany
bernard.rummel@sap.com

ABSTRACT

Ziel des Workshops ist es, Personen zusammenzubringen, die an der Entwicklung und Anwendung von Verfahren zur quantitativen Messung von User Experience arbeiten. Es sollen Erfahrungen ausgetauscht werden und offene Fragestellungen und Probleme angesprochen werden. Mögliche Themen betreffen methodische Fragen bei der Entwicklung und Validierung solcher Verfahren und Problemstellungen in Bezug auf ihre praktische Anwendung in konkreten Projekten.

CCS CONCEPTS

Human-centered computing → HCI design and evaluation methods

KEYWORDS

User Experience, Quantitative Methoden, UX messen, Fragebögen

1 Fragestellung

Die Messung der Benutzererfahrung (User Experience, kurz UX) ist kein Selbstzweck. Es gibt eine Reihe von unterschiedlichen und sehr wichtigen Fragestellungen, die eine quantitative Messung von UX erfordern (siehe z.B. [1]):

- *Kontinuierliches UX Monitoring*: Hat eine Neugestaltung des Produkts die UX verbessert? Ist die neue Version in einem objektiven Sinne besser als die Vorversion? Eine reliable und valide Methode zur Messung von UX kann diese Fragen durch einen statistischen Vergleich zweier UX Messungen leicht beantworten.
- *Vergleich mit Wettbewerbern auf dem Markt*: Wie gut ist die UX des eigenen Produkts verglichen mit der UX der Wettbewerber? Der Unterschied zur ersten Frage besteht darin, dass das Produkt hier mit einer Reihe von anderen Produkten verglichen wird.
- *Prüfen, ob ein Produkt eine ausreichende UX aufweist*: Erfüllt das Produkt die allgemeinen Erwartungen der Nutzer bzgl.

UX? Solche allgemeinen Erwartungen entstehen durch die Verwendung von häufig verwendeten Produkten. Falls die verwendete Messmethode einen Benchmark (d.h. einen Datensatz mit Messergebnissen vieler Produkte) bereitstellt, kann diese Frage einfach durch einen Vergleich des eigenen Ergebnisses mit den Daten im Benchmark beantwortet werden.

- *Klären, in welchen Bereichen man investiert*: Oft hat man eine einfache Verbindung zwischen bestimmten Design-Elementen und UX Aspekten, zum Beispiel zwischen der Navigationsstruktur eines Produkts und der Effizienz der Bedienung oder zwischen der verwendeten Terminologie und der einfachen Verständlichkeit. Eine Messung der UX Aspekte kann damit zeigen, welche Veränderungen das größte Verbesserungspotential haben.
- *Return on Investment*: Hat sich eine Investition in die Verbesserung der UX gelohnt? Investitionen zur UX Verbesserung lassen sich meist leicht durch eine Kennzahl messen, z.B. Entwicklungsaufwand in Tagen oder ein investierter Geldbetrag. Eine sinnvolle Diskussion der Frage, ob sich eine solche Investition gelohnt hat, erfordert es, die Verbesserung der UX ebenfalls numerisch auszudrücken.

Es gibt eine Reihe von Methoden, UX oder einzelne Aspekte von UX zu messen. Eine sehr weit verbreitete Methode sind UX Fragebögen. Auch auf der Mensch und Computer wurden in den letzten Jahren mehrere neue UX Fragebögen vorgestellt, z.B. AttrakDiff2 [2], UEQ [3], UEQ+ [4], meCUE [5], VISAWI [6], oder Web-CLIC [7], die auch häufig in Praxisprojekten eingesetzt werden.

Neben Fragebögen sind aber noch eine Reihe anderer Verfahren vorhanden, die numerische UX Kennwerte liefern, z.B. das Messen von Bearbeitungszeiten für Tasks [8] oder deren Schätzung über Verfahren der kognitiven Modellierung, z.B. GOMS [9]. Auch Methoden, die aus der automatischen Erfassung von Click-Pfaden oder Mausbewegungen Rückschlüsse auf UX Qualitäten ziehen, sind hier zu nennen.

Der Workshop ist als Fortsetzung der beiden sehr gut besuchten Workshops zum Thema UX Fragebögen auf der Mensch und Computer 2018 bzw. 2019 (siehe [10] bzw. [11]) geplant. Allerdings wird das Themenspektrum jetzt explizit auf alle Methoden zur Messung von UX ausgeweitet.

Wie auch in den beiden Vorgänger-Workshops sollen die Vorträge offene Probleme bei der Konstruktion, Validierung und

Anwendung von quantitativen Methoden zur Messung von UX thematisieren. Natürlich sind insbesondere auch Vorträge willkommen, die an die im Workshop in Dresden 2018 bzw. Hamburg 2019 herausgearbeiteten Problemstellungen anknüpfen und ggfs. Lösungsmöglichkeiten aufzeigen.

Die eingereichten Papers können in Deutsch und Englisch abgefasst werden und werden bei Annahme im Workshop-Band veröffentlicht.

Beispiele für mögliche Beiträge und Diskussionspunkte sind:

- Methodische Probleme in Bezug auf quantitative Verfahren, speziell in Bezug auf Objektivität, Reliabilität und Validität der Messergebnisse.
- Innovative Methoden zur Messung und neue Ansätze zur Darstellung und Präsentation der Ergebnisse quantitativer Verfahren.
- Herausforderungen und Probleme bei der Anwendung solcher quantitativer Methoden zur Messung von UX in der Praxis.
- Sicherstellen der Datenqualität, speziell bei Online-Befragungen.
- Etablieren eines auf quantitativen UX Daten basierenden Qualitätsprozesses im Unternehmen.

Natürlich sind auch andere Fragestellungen möglich, solange sie sich im oben beschriebenen Inhaltsbereich bewegen.

2 Geplanter Ablauf

- Einleitung und kurze Beschreibung des Ablaufs durch die Organisatoren.
- 7 kompakte Vorträge (maximal 15 Minuten), die jeweils ein konkretes Problem oder Problemfeld bei Entwicklung und beim Einsatz von quantitativen Methoden zur Messung von UX herausarbeiten.
- Diskussion offener methodischer Fragestellungen durch die Teilnehmer.

3 Beiträge

Folgende Beiträge werden im Workshop vorgestellt.

Anna-Lena Meiners, Andreas Hinderks & Jörg Thomaschewski: Korrelationen zwischen UX Fragebögen

Verschiedene UX Fragebögen liefern eine Kennzahl, die die Gesamtwahrnehmung der User Experience beschreibt (UX KPI). Es wird in einer Studie zu YouTube untersucht, ob die KPIs von User Experience Questionnaire+ (UEQ+), System Usability Scale (SUS) und Net Promotor Score (NPS) vergleichbare Ergebnisse liefern. Im Ergebnis zeigte sich, dass die UEQ+ Skala „Intuitive Bedienung“ hoch mit der SUS KPI korreliert. Alle anderen Korrelationen weisen moderate bis geringe Werte auf. Aus den Ergebnissen kann geschlossen werden, dass die KPIs der einzelnen Fragebögen nicht allgemein vergleichbar sind.

Heike Sandkühler, Martin Schrepp & Jörg Thomaschewski: UX Messung mithilfe des UEQ+ Frameworks

Der Beitrag beschreibt einen Ansatz zur kontinuierlichen Messung der User Experience eines digitalen Zeitungsangebots mit dem UEQ+. Der UEQ+ ist eine Sammlung von UX-Skalen, die beliebig kombiniert werden können. Damit kann man einen exakt auf die Erfordernisse des zu evaluierenden Produkts zugeschnittenen Fragebogen erzeugen. Es wird gezeigt, wie die relevanten UEQ+ Skalen im Projekt ausgewählt wurden, wie der entsprechende Online-Fragebogen ins Produkt eingebettet wurde und wie die Ergebnisse zur Verbesserung des Angebots eingesetzt werden.

Bernard Rummel: Kostbare Zeit - Analyse und Modellierung von Bearbeitungszeiten

Bearbeitungszeit ist die bei weitem beliebteste Metrik zur Erfassung von Effizienz als Komponente der Gebrauchstauglichkeit. Statistische Besonderheiten von Zeitverteilungen sind jedoch unter UX-Praktikern noch wenig bekannt. In diesem Beitrag wird eine Anleitung gegeben, wie man durch geeignete Datenvisualisierung zufällige von nicht-zufälligen Einflüssen auf Bearbeitungszeiten unterscheiden und geeignete Modelle für quantitative Aussagen auswählen kann.

Meinald Thielsch & Gerrit Hirschfeld: In der Kürze liegt die Würze - Der Web-CLIC Short zur Bewertung digitaler Inhalte

In digitalen Medien ist der Inhalt entscheidend. Für die Messung der subjektiven Wahrnehmung von Inhalten steht mit dem Web-CLIC auch ein sorgfältig konstruierter Fragebogen zur Verfügung. Der Beitrag beschreibt die Validierung einer Kurzversion des Web-CLIC Fragebogens mit nur 4 Items. Die Unidimensionalität wurde mit einer konfirmatorischen Faktoranalyse getestet. Hohe Korrelationen zwischen der verkürzten und der Vollversion deuten darauf hin, dass der Web-CLIC-S sinnvoll als Vorhersage für den Gesamtwert der Vollversion genutzt werden kann. Weiterhin werden Ergebnisse zur internen Konsistenz, Retest-Reliabilität und Konstruktvalidität berichtet.

David Obremski, Carolin Wienrich & Astrid Carolus: What the user's voice tells us about UX - Analyzing parameters of the voice as indicators of the User Experience of the usage of voice assistants

This paper investigates voice parameters as indicators of the user experience during their interaction with voice-based technology. Following the requirements of a non-intrusive measurement and an automatic analysis, a tool is developed, which measures and analyses speech recordings of users during interactions with smart speakers. The tool distinguishes between user- and voice assistant speech and analyses mean amplitude and pitch of user's speech as parameters reflecting the user experience of users. A user study presents first steps of a validation of the first version and reveals limitations and requirements for further developments.

Martin Schrepp: A Comparison of UX Questionnaires - What is their underlying concept of user experience?

The paper analyzes a sample of 40 established UX questionnaires concerning their differences and similarities. This analysis shows the heterogeneity of the concept of user experience in the field

and the current problems to determine the exact semantic meaning behind UX scales. It is discussed how the current practice to develop new UX questionnaires adds to this inhomogeneity and lack of common understanding on what we measure when we measure UX.

Marc Busch, Marco della Schiava, Christin Kipke & Benedict Salzbrunn: Adaptive Item-Vorgabe als Innovation in der modernen UX Messung

Der Beitrag beschreibt die Konstruktion von UX-Fragebögen nach den Prinzipien der probabilistischen Testtheorie. Dieser Ansatz ermöglicht die Items adaptiv vorzugeben und damit die Anzahl zu beantwortender Items pro Teilnehmer zu reduzieren. Die adaptive Vorgabe der Items wird ermöglicht durch den Einsatz des dichotom-logistischen Rasch-Modells, bei dem kein deterministischer Zusammenhang zwischen den Bewertungen der Items und deren Personenparameter (z.B. wahrgenommene Nützlichkeit eines digitalen Services) angenommen wird, sondern ein zugrunde liegendes Merkmal, welches die Wahrscheinlichkeit der Bewertungen der Items durch die Versuchspersonen - neben anderen Einflüssen - beeinflusst. Dies ermöglicht Vergleiche der einzelnen Personenparameter, wobei nicht alle Personen dieselben Items bewertet haben müssen.

REFERENCES

- [1] Schrepp, M.; Hinderks, A. & Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios. In: Marcus, A. (Ed.): Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Lecture Notes in Computer Science, Volume 8517, S. 383-392, Springer International Publishing.
- [2] Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Ziegler, J. & Szwillus, G. (Hrsg.), Mensch & Computer 2003. Interaktion in Bewegung, S. 187-196, Stuttgart, Leipzig: B.G. Teubner.
- [3] Laugwitz, B.; Schrepp, M. & Held, T. (2006). Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. A.M. Heinecke & H. Paul (Eds.): Mensch & Computer 2006 - Mensch und Computer im Strukturwandel. Oldenbourg Verlag, S. 125 - 134
- [4] Schrepp, M., & Thomaschewski, J. (2019). Eine modulare Erweiterung des User Experience Questionnaire. Mensch und Computer 2019-Usability Professionals.
- [5] Minge, M. & Riedel, L. (2013). meCUE - Ein modularer Fragebogen zur Erfassung des Nutzungserlebens. In: S. Boll, S. Maaß & R. Malaka (Hrsg.): Mensch und Computer 2013: Interaktive Vielfalt, S. 89-98. München, Oldenbourg Verlag.
- [6] Moshagen, M. & Thielsch, M. T. (2010). Facets of visual aesthetics. International Journal of Human-Computer Studies, 68, S. 689-709.
- [7] Thielsch, M. T. & Hirschfeld, G. (2018). Evaluation von Online-Inhalten mit dem Web-CLIC Fragebogen. In: Dachselt, R. & Weber, G. (Hrsg.), Mensch und Computer 2018 - Tagungsband. Bonn: Gesellschaft für Informatik e.V.
- [8] Rummel, B. (2017). Beyond average: Weibull analysis of task completion times. Journal of Usability Studies, 12(2), 56-72.
- [9] John, B. E., & Kieras, D. E. (1996). The GOMS family of user interface analysis techniques: Comparison and contrast. ACM Transactions on Computer-Human Interaction (TOCHI), 3(4), 320-351.
- [10] Schrepp, M. & Rummel, B. (2018). UX Fragebögen - Konstruktion und praktischer Einsatz. In: Dachselt, R. & Weber, G. (Hrsg.), Mensch und Computer 2018 - Workshopband. Bonn: Gesellschaft für Informatik e.V. DOI: 10.18420/muc2018-ws16-0103.
- [11] Schrepp, M. & Rummel, B., (2019). Konstruktion und praktischer Einsatz von User Experience Fragebögen. Mensch und Computer 2019 - Workshopband. Bonn: Gesellschaft für Informatik e.V. DOI: 10.18420/muc2019-ws-109

Organisatoren

Martin Schrepp

Dr. Martin Schrepp studierte Mathematik und Psychologie an der Universität Heidelberg. 1990 Abschluss als Diplom-Mathematiker. 1990 – 1993 Promotion in Psychologie. Seit 1994 bei der SAP AG tätig. Er ist einer der Autoren des User Experience Questionnaire (UEQ) und hat Erfahrungen in der praktischen Anwendung zahlreicher anderer UX Fragebögen. Er ist auch Autor einer Reihe von Beiträgen zu methodischen Fragen im UX Bereich.

Bernard Rummel

Bernard Rummel studierte Psychologie an der Universität Kiel (Diplom 1990). Nach neun Jahren am Schifffahrtsmedizinischen Institut der Marine kam er 2000 zu SAP. Er arbeitet weiterhin im DIN-Normenausschuss Ergonomie – Benutzungsschnittstellen an Normenreihen wie der ISO 9241, sowie als Fachprüfer des UXQB für Usability Testing und Evaluation (CPUX-UT). Seit 2011 beschäftigt er sich bei SAP mit der Quantifizierung von Gebrauchstauglichkeit und Usability Benchmarking.