

Hans-Christoph Hobohm (Hrsg.)

Informationswissenschaft zwischen virtueller Infrastruktur und materiellen Lebenswelten

**Information Science between
Virtual Infrastructure and Material Lifeworlds**

Unter Mitarbeit von Judith Pfeffing

Proceedings des 13. Internationalen Symposiums
für Informationswissenschaft (ISI 2013)

Potsdam, 19. bis 22. März 2013

vwh

Verlag Werner Hülsbusch
Fachverlag für Medientechnik und -wirtschaft

Eignen sich Tree-Tests und Firstclick-Tests für die nutzerzentrierte Evaluierung der Informationsarchitektur?

Katja Eck², Ben Heuwing¹, Christa Womser-Hacker¹

¹ Universität Hildesheim

Institut für Informationswissenschaft und Sprachtechnologie

Marienburger Platz 22, D-31141 Hildesheim

{heuwing|womser}@uni-hildesheim.de

² Wolters Kluwer Deutschland GmbH

Freisinger Straße 3, D-85716 Unterschleißheim/München

KEck@wolterskluwer.de

Zusammenfassung

Neue Methoden für die Überprüfung der Informationsarchitektur (Navigationsdesign und Informationsstruktur) versprechen, schnell und einfach Ergebnisse zu einem frühen Zeitpunkt im Entwicklungsprozess zu liefern. Ein Vergleich mit einem formativen Nutzertest bestätigt dies weitgehend und zeigt Besonderheiten und Grenzen der Methoden auf.

Abstract

New approaches to evaluating information architecture (navigation design and information structure) promise to deliver results in a swift and convenient way in the early stages of the development process. This analysis confirms these claims and demonstrates the limits and special characteristics of these tools in comparison to formative user testing.

In: H.-C. Hobohm (Hrsg.). Informationswissenschaft zwischen virtueller Infrastruktur und materiellen Lebenswelten. Tagungsband des 13. Internationalen Symposiums für Informationswissenschaft (ISI 2013), Potsdam, 19.–22. März 2013. Glückstadt: Verlag Werner Hülsbusch, 286–297.

1 Einleitung

Bei der Entwicklung von Informationssystemen sind die Methoden und Prozesse der nutzerzentrierten Gestaltung eine wichtige Voraussetzung für gebrauchstaugliche Systeme (ISO 2009). Die Informationsarchitektur umfasst die Systematisierung der Inhalte und der Funktionen (Informationsstruktur) sowie ihre Abbildung in der Benutzeroberfläche (Navigationsdesign). Diese werden oft schon früh in der Entwicklung gestaltet und sollten unter Einbeziehung von Nutzern evaluiert werden.

In einem iterativen Entwicklungsprozess müssen Evaluierungen besonders einfach und kostengünstig durchgeführt werden, um mehrere Entwicklungs-Iterationen und den Vergleich von verschiedenen Gestaltungsvarianten zu ermöglichen. Neue Methoden wie Firstclick-Tests und Tree-Tests versprechen eine einfache Überprüfung der Informationsstruktur und des Navigationsdesigns. Die Anwendung dieser Methoden wird durch die Verfügbarkeit von Online-Tools zusätzlich vereinfacht. Aber liefern diese Verfahren auch reproduzierbare Ergebnisse? Welche Arten von Fehlern können erkannt werden und wie groß ist dabei die Abdeckung? In dieser Studie werden diese Fragen anhand eines Vergleichs mit den Ergebnissen eines formativen Nutzertests im Labor überprüft und Hinweise für den Einsatz der Werkzeuge in der Praxis gegeben.

2 Methoden für die benutzerzentrierte Evaluierung von Informationsarchitekturen

Benutzerorientierte Evaluierungsmethoden stellen das tatsächliche Nutzungsverhalten und die subjektive Einstellung repräsentativer Benutzer in den Mittelpunkt der Überprüfung. Zu allgemeinen Methoden, mit denen auch Aspekte der Informationsarchitektur erfasst werden können, gehören aufgabenbasierte Nutzertests und der Einsatz von Fragebogenmethoden. Fragebogenverfahren konzentrieren sich auf die subjektive Einstellung der Nutzer. Nutzertests versuchen zusätzlich Beobachtungsdaten zu erheben und sind vergleichsweise aufwendig in der Durchführung, ermöglichen aber hilfreiche Einblicke in die Sichtweise der Nutzer (Schweibenz 2011: 14). Bei Remote-

Nutzertests werden Teilnehmer online entweder synchron von einem Testleiter begleitet oder bekommen Aufgaben asynchron in textueller Form.

Eine weitere etablierte Methode ist das Card Sorting (Schüßler/Mandl/Womser-Hacker 2009; Toub 2000). Die Teilnehmer gruppieren dabei Inhalte nach eigenen Aspekten und benennen die dadurch gebildeten Kategorien (offenes Card-Sorting). Um eine existierende Systematik aus Gruppen und Bezeichnungen zu überprüfen, kann ein Card-Sorting auch mit vorgegebenen Kategorien durchgeführt werden (geschlossenes Card-Sorting).

In den letzten Jahren wurden weitere innovative Methoden und dafür geeignete Werkzeuge vorgeschlagen, mit denen speziell die Informationsarchitektur evaluiert und verbessert werden kann. Diese kombinieren Aspekte des Card-Sorting mit Nutzertests: Nutzer bekommen einfache Aufgaben gestellt und interagieren mit reduzierten Repräsentationen der zu untersuchenden Systeme um diese zu lösen. Die Aufgabenorientierung kann dabei sehr hilfreich sein, gleichzeitig ist das Nutzerfeedback jedoch reduziert und ähnelt in dieser Beziehung asynchronen Remote-Tests.

2.1 Firstclick-Test

Das Navigationsdesign lässt sich mit dem Firstclick-Testverfahren vorab analysieren.

In Firstclick-Tests werden Nutzer aufgefordert, den Bereich eines angezeigten Screenshots durch Anklicken zu markieren, hinter welchem sie die Lösung einer bestimmten Fragestellung vermuten. Durch diesen Test kann nachvollzogen werden, ob Nutzer wichtige Seitenelemente wahrnehmen und verstehen. Der Test kann in verschiedenen Entwicklungsphasen genutzt werden, vom einfachen Wireframe bis hin zum fertiggestellten System.

In der Untersuchung wurde das Online-Tool *Chalkmark* verwendet (s. Abb. 1). Vergleichbare Anwendungen sind *usabilla* oder *The Clicktest*.¹ Die Verwendung der Tools ist jeweils ähnlich. Zunächst muss der Test mit Aufgabentexten und entsprechenden Bilddateien konfiguriert werden. Durch das Versenden eines Links an die Testteilnehmer ist die Vorarbeit abgeschlossen. Die Ergebnisse sind jederzeit einsehbar und werden visuell in Form von Heatmaps dargestellt, welche die Häufigkeit von Klicks in bestimmten Bereichen anzeigen.

¹ www.optimalworkshop.com/chalkmark.htm; usabilla.com; theclicktest.com
<19.02.2013>

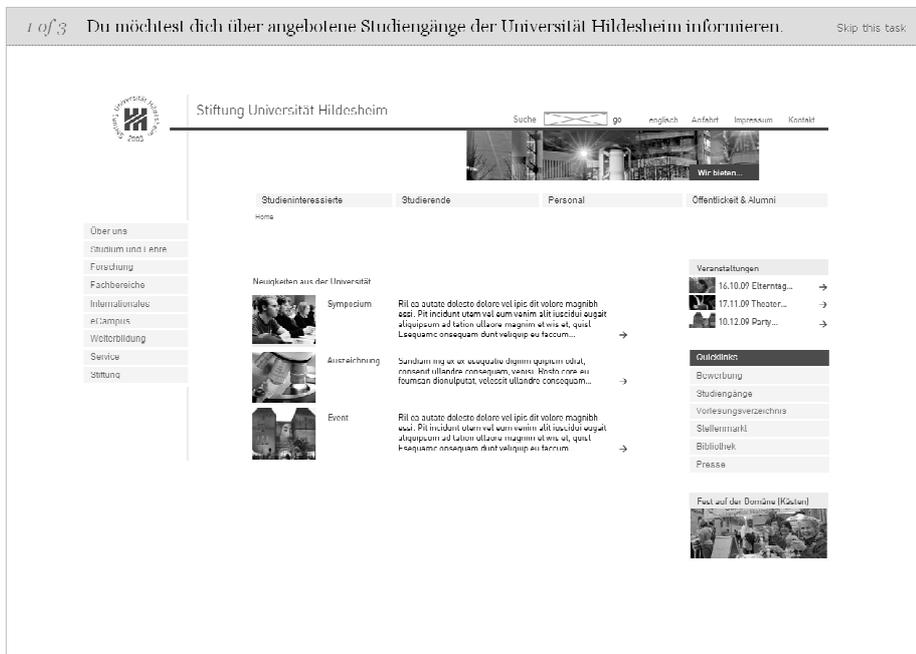


Abb. 1 Aufgabe mit Screenshot in Chalkmark

Grundannahme des Verfahrens ist, dass bei der Durchführung einer Aufgabe ein zur Lösung führender erster Mausklick eine hohe Vorhersagekraft für den späteren Erfolg hat. Dies konnte für Nutzertests bestätigt werden (Bailey et al. 2009: 5).² In einer Fallstudie (Peters/Schumacher 2010) wird ein vergleichbares Werkzeug vorgestellt, mit dessen Hilfe zwei Online-Tools unter Verwendung derselben vier Aufgaben und unter zusätzlicher Verwendung begleitender Fragen miteinander verglichen werden.

2.2 Tree-Tests

Eine häufig Tree-Testing genannte Methode hat sich aus dem geschlossenen Card-Sorting entwickelt (Spencer 2003) und wurde später als Online-Tool umgesetzt (O'Brien 2009).

Bei einem Tree-Test müssen Nutzer Inhalte innerhalb einer vordefinierten, vereinfacht dargestellten hierarchischen Struktur suchen. Es werden dabei nur die Unterebenen der jeweils ausgewählten Kategorie angezeigt. Der

² Anmerkung: Einer der Autoren bietet kommerziell ein entsprechendes Werkzeug an.

Test dient dazu, den strukturellen Aufbau eines Informationssystems zu evaluieren. Er eignet sich daher im Unterschied zum Card Sorting vor allem für die späteren Stadien einer Entwicklung, in denen schon ein Entwurf für eine Informationsstruktur vorhanden ist. Der Nutzer bekommt eine Aufgabe angezeigt und kann dann im Strukturbaum navigieren, um sich für das Konzept zu entscheiden, in dem er die Antwort auf eine Fragestellung erwartet. Es werden dabei keine Inhalte angezeigt und auch keine Hinweise darauf gegeben, ob die Auswahl korrekt war.

Im Rahmen dieser Studie ist *Treejack* zur Anwendung gekommen (s. Abb. 2). Alternativ eignet sich beispielsweise *C-Inspector*³. Für die Durchführung werden zunächst ein Begrüßungstext, eine Baumstruktur und Aufgabenstellungen festgelegt. Daraufhin kann auch hier ein Link zur Testumgebung an verschiedene Testpersonen verschickt werden. Die Auswertung fällt aufgrund der komplexeren Testumgebung umfangreicher aus als bei den Firstclick-Tests. Die Ergebnisse (Taskerfolg, Bearbeitungszeit) können visuell in Diagrammen dargestellt oder in tabellarischer Form eingesehen werden. Ersichtlich wird dabei, welche Menüoptionen zum Einstieg in die Suche genutzt wurden und welche Navigationspfade Testnutzer wählten. Zusammengefasst wird außerdem, welche Seiten als Ziel markiert und ob sie direkt erreicht wurden.

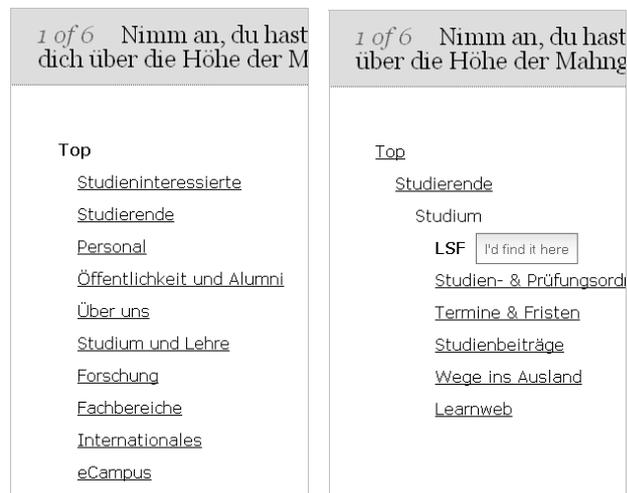


Abb. 2 Tree-Test mit Treejack aus Teilnehmersicht: links die erste Auswahlebene, rechts die geöffnete Hierarchiestruktur mit Auswahllink

3 www.optimalworkshop.com/treejack.htm, www.c-inspector.com <19.02.2013>

Wegen der einfachen Durchführung ist es möglich, mehrere Varianten einer Informationsstruktur zu vergleichen. Meier/Wolff (2012) untersuchen dieses Verfahren mit einem Fokus auf die Rekrutierung von Teilnehmern durch Crowdsourcing und stellen dabei zu Labortests vergleichbare Ergebnisse in Bezug auf den Aufgabenerfolg fest.

Die dargestellten Methoden und dabei verwendeten Tools wurden bisher nur selten beschrieben und wissenschaftlich untersucht. Daher sollen hier die vorgestellten Methoden Firstclick-Tests und Tree-Tests verglichen und die Ergebnisse anhand von Nutzertests validiert und qualitativ untersucht werden. Als Ergebnis können Empfehlungen zum Einsatz und zur Auswahl der Methoden gegeben werden.

3 Vorgehen

Ziel der Untersuchung war eine Einstufung der Validität der vorgestellten Testverfahren und die Einschätzung ihrer Anwendbarkeit in der Praxis. Als Vergleichsmaßstab wurden dafür die Ergebnisse eines Labortests zu denen der Online-Tests in Beziehung gesetzt. Untersucht wurde ein Entwurf für eine neue Version der Website der Universität Hildesheim. Auf der Basis einer ausführlichen Zielgruppen- und Logfile-basierten Taskanalyse wurden dafür zunächst realistische Testszenarien entwickelt.

Designentwürfe des Neuentwurfs der Website bildeten die Arbeitsgrundlage für den Firstclick-Test (3 Aufgaben). Für den Tree-Test (6 Aufgaben) kam eine umfangreiche Struktur mit 174 Konzepten auf fünf Hierarchieebenen zum Einsatz. Die Teilnehmer bearbeiteten zunächst den kürzeren Firstclick-Test mit einem Klick pro Aufgabe, bevor sie zum umfangreicheren Tree-Test weitergeleitet wurden. Dabei lösten sie Tasks wie „Du möchtest dich im Learnweb für einen Kurs anmelden“ im Firstclick-Test oder „Du möchtest dich über die Öffnungszeiten des Immatrikulationsamtes informieren, um deinen Studentenausweis abholen zu können“ im Tree-Test.

Die Tests wurden über einen Zeitraum von 27 Tagen von insgesamt 41 Studierenden und 14 Studieninteressierten abgeschlossen. Gewertet wurden die Ergebnisse aller abgeschlossenen Tasks (61 Studierende und 18 Studieninteressierte hatten die Tests ursprünglich begonnen). Die Ergebnisse der Studien wurden einzeln in Hinblick auf mögliche Probleme bei der Nutzung analysiert.

Um die Validität der Ergebnisse zu überprüfen, wurde parallel ein Labor-test durchgeführt. Klassische Nutzertests können aufgrund der kontrollierten Testsituation als Vergleichsmaßstab für Usability-Methoden gelten (z. B. John/Marks 1997: 194). Hier wurden sowohl die Testszenarien des First-click-Tests als auch die des Tree-Tests mit 23 Teilnehmern durchgeführt (15 Studierende, 8 Studieninteressierte). Testobjekt war ein für diesen Zweck entwickelter interaktiver Prototyp der neuen Version der Website. Dieser stellte mehr Interaktionsmöglichkeiten zur Verfügung als die Testobjekte in den Online-Experimenten. So waren hier beispielsweise dynamische Drop-Down-Menüs implementiert. Die Tests wurden in Räumen der Universität Hildesheim, teilweise auch bei Teilnehmern vor Ort durchgeführt. Die Ergebnisse wurden in Hinblick auf die Übereinstimmung der gefundenen Probleme verglichen.

4 Vergleich der Ergebnisse der Online-Tools

Obwohl beide Online-Tools eine ähnliche Vorarbeit und einen ähnlichen Ablauf der Tests aufweisen, unterscheiden sie sich doch deutlich in Bezug auf die Art der Ergebnisse. Bei der Analyse wurde deutlich, dass sie sich für unterschiedliche Zielsetzungen eignen und sich dabei gut ergänzen.

4.1 Ergebnisse der Firstclick-Experimente

Für die Analyse der Firstclick-Tests steht, neben den Bearbeitungszeiten der Aufgaben, primär eine Heatmap zur Verfügung, welche die Mausclicks der Teilnehmer auf dem untersuchten Screenshot darstellt. Daraus lässt sich vor allem auf Probleme schließen, welche die Gestaltung einzelner Seiten und vor allem der Navigationsbereiche betreffen.

So konnte überprüft werden, ob bestimmte Seitenelemente wie etwa Menüoptionen wahrgenommen wurden. Dabei war die Wahrnehmung bestimmter Navigationselemente in den Online Tests sogar besser als in den Labor-tests (z. B. der Quicklinks in der rechten Seitenspalte). Dies kann an einer entspannteren Umgebung für die Nutzer gelegen haben, was Ergebnisse aus anderen Vergleichsstudien vermuten lassen (vgl. Tullis et al. 2002: 8). Weiterhin ließ sich auf Probleme mit der Terminologie bestimmter Menüpunkte

schließen, wenn andere, designspezifische Gründe als unwahrscheinlich ausgeschlossen werden konnten.

Ein offensichtlicher Nachteil der Methode war, dass durch die Screenshot-Vorlage dynamische Elemente in den Experimenten nicht zur Verfügung standen und daher nicht in die Analyse einbezogen werden konnten. Die Vorteile lagen in der schnellen Bearbeitungszeit und der Erfassung des spontanen, ersten Eindrucks der Teilnehmer.

4.2 Ergebnisse der Tree-Test-Experimente

Der Tree-Test bot bei der Analyse der Testergebnisse zahlreiche Möglichkeiten, um vor allem strukturelle Nutzungsprobleme aufzudecken. Erste Ergebnistendenzen ließen sich bereits aus Visualisierungen des Taskerfolges und der Bearbeitungszeit erkennen. Für eine tiefergehende Analyse eigneten sich vor allem die von den Nutzern gewählten Bearbeitungspfade.

Daraus ließ sich vor allem nachverfolgen, ob bestimmte Inhalte von den Nutzern an anderer Stelle innerhalb der Seitenstruktur eingeordnet wurden. Weiterhin wurde durch die große Streuung der Ergebnisse für einige Aufgaben deutlich, dass sich bestimmte Inhalte nicht einer einzelnen Kategorie zuordnen ließen und möglicherweise einen eigenen Menüpunkt verdienen. Außerdem ließen sich auch bei dieser Methode Verständnisprobleme in Bezug auf die verwendeten Bezeichnungen erkennen.

Durch die vereinfachte Darstellung eines Strukturbaums zum Lösen verschiedener Aufgaben waren Probleme des Navigationsdesigns im Vergleich zum Firstclick-Test nicht identifizierbar. Auch Verweise innerhalb der Website konnten nicht abgebildet und analysiert werden. Der Vorteil der Methode lag vor allem in der strukturellen Überprüfung und umfangreichen Analysemöglichkeiten.

5 Validierung durch Labortest

Im Anschluss an die Auswertung der Ergebnisse der Online-Experimente wurden diese mit den Ergebnissen aus dem Labortest des interaktiven Prototyps verglichen.

5.1 Taskerfolg

Bei den Firstclick-Tests zeigten sich für die ersten beiden Aufgaben ähnliche Ergebnisse (s. Abb. 3, Task 1–3). Bei der Bearbeitung der dritten Aufgabe waren jedoch die Teilnehmer des Labortests häufiger erfolgreich. Grund hierfür könnte ein Lerneffekt aus den vorherigen Aufgaben sein, da die Darstellung dynamischer Menüfunktionen vermutlich die Orientierung auf der Website erleichterte.

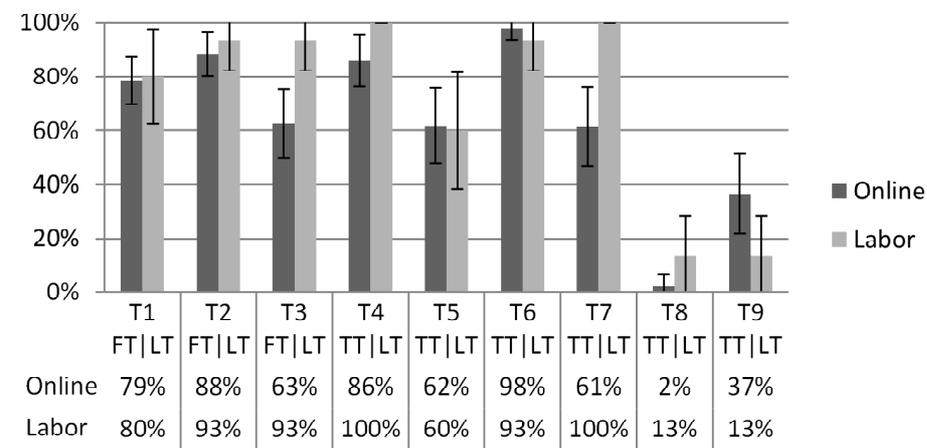


Abb. 3 Aufgabenerfolg Studierende bei Online-Tests ($n = 61-41$) für Firstclick-Test (FT: Task 1–3) und Tree-Test (TT: Task 4–9) im Vergleich zum Labortest ($n = 15$) mit Konfidenzintervall für $\alpha = 0,1$

Bei den Tree-Tests (Abb. 3, Task 4–9) waren die Ergebnisse insgesamt ebenfalls vergleichbar zum Labortest, nur in Task 7 ergaben sich größere Abweichungen, ebenfalls aufgrund von Lerneffekten im Labor. Bei den Tree-Tests zeigte sich in einigen Tasks die Tendenz, dass die Teilnehmer im Vergleich zum Labortest häufiger erfolgreich abschlossen, etwa wenn sich Schwierigkeiten erst beim Auffinden der Information auf der Zielseite ergaben (Task 9).

Die Ergebnisse für die Dauer der Durchführung stellten sich dagegen als wenig vergleichbar heraus, da die Online-Experimente teilweise durch die Teilnehmer unterbrochen und später wieder aufgenommen wurden.

5.2 Vergleich der qualitativen Analyseergebnisse

Das wichtigste Kriterium für die Anwendung der Online-Experimente in der Praxis ist, dass nicht fälschlicherweise Nutzungsprobleme ermittelt werden, die wahrscheinlich bei der Nutzung des tatsächlichen Interfaces keine Rolle spielen (false positives). Insgesamt wurden neun der zehn durch die Tree-Tests ermittelten Nutzungsprobleme auch in den Labortests bestätigt. Dazu gehörten vor allem Fälle, in denen Inhalte an anderen Stellen in der hierarchischen Struktur vermutet wurden. Bei den Firstclick-Tests wurden insgesamt vier Probleme ermittelt, von denen drei bestätigt wurden. Ein Problem, welches im Labor nicht bestätigt wurde, war der häufige Versuch, Test im Online-Test oder Studentenjobs im Zielgruppenmenü für Studierende zu finden (drei von 18 Studieninteressierten), was der Struktur der getesteten Seite nicht entsprach.

Ein weiteres wichtiges Kriterium ist die Abdeckung der ermittelten Probleme bezogen auf die Ergebnisse des Nutzertests im Labor. Viele der Nutzungsprobleme, die im Labortest am interaktiven Prototypen deutlich wurden, konnten in beiden Online-Experimenten nicht gefunden werden (17 von 27). Ein Grund dafür war, dass die Repräsentationen der Website, die in den Online-Experimenten getestet wurden, einen geringeren Umfang hatten. Weitere Probleme waren aufgrund der Einschränkungen der Online-Experimente nicht erfassbar. Allerdings wurden die schwerwiegendsten Probleme der Labortests auch in den Online-Experimenten aufgedeckt.

5.3 Aufwand für die Online-Experimente

Für eine Einschätzung der Online-Experimente ist neben Qualität und Abdeckung der Ergebnisse auch der einzusetzende Aufwand zu berücksichtigen. Der materielle Aufwand liegt deutlich unter dem eines Labortests. Trotz der Gebühren für die Nutzung kommerzieller Tools werden Kosten bei der Rekrutierung der Teilnehmer und für ein geeignetes Usability-Labor gespart. Der zeitliche Aufwand für die Vorbereitung der Experimente ist vergleichbar mit der Vorarbeit für Labortests, während die Durchführung und Auswertung insgesamt weniger Zeit beansprucht.

6 Diskussion und Ausblick

Ein großer Teil der im Labortest ermittelten Probleme wurde in den Online-Experimenten nicht erfasst. Ein Teil der Probleme hätte eventuell mit umfassenderen Online-Experimenten mit mehr Aufgaben erfasst werden können, was bei den Online-Tools durchaus realistisch ist (vgl. Bailey et al. 2009: 6). Einige Arten von Nutzungsproblemen liegen jedoch außerhalb des Erfassungsbereichs beider Tools. Die Tools sind somit kein Ersatz für reguläre formative Nutzertests mit Prototypen. Sie haben vor allem die Vorteile einer höheren Flexibilität und einer schnelleren Rückmeldung durch die Nutzer.

Die Ergebnisse, die mit den Online-Tools erhoben wurden, scheinen jedoch ausreichend belastbar für den Praxiseinsatz zu sein: Nur wenige der in den beiden Online-Experimenten gefundenen Probleme wurden im Labortest nicht bestätigt und es wurden vor allem die schwerwiegenden Probleme gefunden.

Der Nutzen der Verfahren kann daher vor allem in ihrer frühen und wiederholten Anwendung in einem iterativen Entwicklungsprozess gesehen werden, wenn die bestehenden Entwürfe noch einfach geändert werden können. Zusätzlich gehen wir aufgrund der Hilfestellungen durch die Tools und die größere Objektivität der Online-Experimente davon aus, dass für die Durchführung weniger Erfahrung und Methodenwissen notwendig ist, als bei klassischen Nutzertests.

Zu untersuchen bleibt, ob die Verfahren auch auf andere Informationssysteme und umfangreichere Informationsstrukturen angewendet werden können. Weiterhin kann in Fallstudien der iterative Einsatz der Methoden und ihr Einfluss auf den weiteren Entwicklungsprozess (Downstream Usability) untersucht werden.

Literaturverzeichnis

- Bailey, R. W.; Wolfson, C. A.; Nall, J.; Koyani, S. (2009). Performance-Based Usability Testing: Metrics That Have the Greatest Impact for Improving a System's Usability. In: Proceedings of the 1st International Conference on Human Centered Design. San Diego, CA: Springer, 3–12.

- ISO (2009). Human-centred design for interactive systems. Ergonomics of human system interaction Part 210 (ISO 9241-210). ISO – International Organization for Standardization.
- John, B. E.; Marks, S. J. (1997). Tracking the effectiveness of usability evaluation methods. In: Behaviour & Information Technology 16 (4), 188–202.
- Meier, F.; Wolff, C. (2012). IA von Websites: asynchrone Remote-Tests und Laborstudien im Vergleich. In: Reiterer, H.; Deussen, O. (Hrsg.). Mensch & Computer 2012: interaktiv informiert – allgegenwärtig und allumfassend!? München: Oldenbourg.
- OBrien, D. (2009). Tree Testing – A quick way to evaluate your IA. <http://www.bboxesandarrows.com/view/tree-testing> <5.11.2012>.
- Peters, K.; Schumacher, R. (2010). Access-Task Surveys: A Low-Cost Method for Evaluating The Efficacy of Software User Interfaces. In: Beyond the Usability Lab. San Francisco: Morgan Kaufmann, 241–250
- Schweibenz, W. (2011). Grundlagen des Usability-Engineerings – Aspekte der Evaluation von Benutzerfreundlichkeit von Bibliothekswebsites. In: Bekavac, B.; Schneider, R.; Schweibenz, W. (Hrsg.). Benutzerorientierte Bibliotheken im Web: Usability-Methoden, Umsetzung und Trends. Berlin: de Gruyter, 9–30.
- Spencer, D. (2003). Card-Based Classification Evaluation. http://www.bboxesandarrows.com/view/card_based_classification_evaluation <5.11.2012>.
- Schüßler, C.; Mandl, T.; Womser-Hacker, C. (2009). Die Effektivität virtueller Tools für die Usability Methode Card Sorting. In: Wandke, H.; Kain, S. (Hrsg.). Mensch & Computer 2009: Grenzenlos frei!? München: Oldenbourg, 273–282.
- Toub, S. (2000). Evaluating Information Architecture – A Practical Guide to assessing Website Organization. http://argus-acia.com/white_papers/evaluating_ia.pdf <14.01.2013>.
- Tullis, T.; Fleischman, S.; McNulty, M.; Cianchette, C.; Bergel, M. (2002). An empirical comparison of lab and remote usability testing of web sites. In: Usability Professionals Association Conference 2002, Florida.