

Stephanie Schütze, Roland Streule, Damian Läge

Warum klassische Evaluation oftmals nicht ausreicht – eine Studie zur Ermittlung der Bedeutsamkeit Mentaler Modelle als Evaluationsmethode

Zusammenfassung

Hohe Benutzertauglichkeit und Akzeptanz eines Webseitenservices sind nur dann gewährleistet, wenn diese auf die funktionalen Bedürfnisse, aber auch auf die strukturellen Vorstellungen ihrer Benutzer zugeschnitten sind. Im Kontext einer Webseite zur „Suche nach E-Learning-Produkten im Internet“ wurde untersucht, inwieweit die Kenntnis der Mentalen Modelle potenzieller Nutzer den Softwareentwicklungsprozess positiv beeinflussen kann. Dabei erlaubte die Erhebung Mentaler Modelle mit Hilfe der Struktur-Lege-Technik (SLT) als Evaluationsmethode einen Vergleich mit der tatsächlich entwickelten Webseite. Die Studie zeigte, dass trotz zuvor bereits durchgeführter, klassischer Evaluationen (Usability Tests, Anwendung von Heuristiken und Cognitive Walkthroughs) 40 Funktionsbereiche genannt wurden, die auf der tatsächlichen Webseite nicht vorgesehen oder umgesetzt sind. Daraus folgt, dass die Erhebung von Mentalen Modellen ebenfalls bereits vor dem Software-Entwicklungsprozess durchgeführt werden soll.

1 Mentale Modelle im Kontext der Usability

Evaluationsschritte im Softwareentwicklungsprozess folgen heutzutage häufig dem „Usability Engineering Lifecycle“ (Nielsen, 1993). Dessen erste Phase ist die Analysephase. Darunter versteht Nielsen eine genaue Zieldefinition, Zielgruppenanalyse, Funktionale Analyse und Aufgabenanalyse, welche mit Methoden wie Interviews, Fragebögen, Verhaltensbeobachtungen, Workshops und Brain Storming erhoben werden können. In der zweiten Phase, der Konzeptphase, wird iterativ das Grob- und Feinkonzept entworfen und als Ergebnis ein Prototyp erstellt. Spätere Phasen fokussieren auf Entwicklung des Systems, Einführung und Durchführung klassischer Nutzertests und Reviews sowie laufender Optimierung und Relaunch-Vorbereitung.

Nielsen betont in späteren Arbeiten wiederholt die Bedeutung für einen frühen Nutzereinbezug, welcher in der Anforderungsanalyse über Grundkonzepte in Skizzen auf Papier (z.B. Paper Prototyping als schnellste und billigste Technik) vorgelegt werden soll (Nielsen, 1994, 2003, 2010) und dessen Feedback und weitere Anforderungen, Bedürfnisse und Wünsche der Nutzer

(i.S. einer formativen, benutzerorientierten Evaluation) in der Konzeptionsphase (Problemanalysephase) einbezogen werden sollen (vgl. auch Herczeg, 2005). Ein frühes Testen führe auch zu einer hohen Wahrscheinlichkeit, Fehler früh zu erkennen, was die Kosten für spätere Fehlerbeseitigungen enorm senke: „It’s a rough estimate, but I would say that the benefits from early usability data are at least ten times greater than those from late usability data.“ (Nielsen, 2003).

Ein jüngerer Ansatz (Young, 2008) stellt die nutzerbezogene Anforderungsanalyse strikt vor die Produktentwicklung. Über Beobachtungen wird die potentielle Nutzergruppe und deren Bedürfnisse definiert. Antworten werden oft mithilfe von Personas formuliert. Die Persona stellt ein Nutzerprofil für eine Zielgruppe dar, mit konkret ausgeprägten Eigenschaften und einem konkreten Nutzungsverhalten. Dabei beschränkt sich diese Analyse auf hypothetische Annahmen und verzichtet auf empirische Evidenzen. Als Ergänzung und Unterstützung zu Personas schlägt Young vor, das „Mentale Modell“ der Nutzer zu erforschen. Dabei sollten das Verhalten, die Vorstellungen und Reaktionen der Nutzer im Mittelpunkt stehen: „Mental models along with web analytics and use cases influence your interaction design concepts“ (Young, 2008, S. 31). Diese Methode könne in allen Entwicklungsständen von Nutzen sein, insbesondere vor der Aufstellung von Produkt- und Interaktionskonzepten (ebd., S. 29ff.). Allerdings ist sie sehr zeitaufwändig und fast als umständlich zu bezeichnen.

Der Begriff des Mentalen Modells wird im Kontext von Software Ergonomie und Usability verstanden als „... the user’s internal representations of their interaction with the system“ (Staggers & Norcio, 1993). Bezogen auf ein User Interface bedeutet dies, dass das Mentale Modell eines Benutzers identisch mit dem sogenannten Conceptual Model des Designers sein sollte (Norman, 1983): „A conceptual model is invented to provide an appropriate representation of the target system, appropriate in the sense of being accurate, consistent, and complete.“ (Norman, 1983, S. 7).

Wenn der Entwickler oder Designer weiss, wie das Benutzermodell aussieht, welche Vorstellungen der Nutzer von der Lösung einer Aufgabe hat und welche Ziele er verfolgt, kann eine benutzerfreundliche, effiziente und problemlos bedienbare Software entwickelt werden.

Die Partizipation durch den Nutzer während der Softwareentwicklung spielt dabei eine grosse Rolle (Partizipative Systementwicklung, vgl. Rauterberg, 1991). Ein frühzeitiger Einbezug ist, wie vorher herausgearbeitet, ein wichtiger Faktor, um kostspielige Fehlentwicklungen zu vermeiden. Frühzeitig bedeutet hier, dass bereits vor der Entwicklung des Grobkonzepts die Vorstellungen und Ziele des Nutzers vorliegen sollten. Diese Vorstellungen sollen dann in das Grob- und später in das Feinkonzept einfließen. Erst anschließend wird mit der Programmierung begonnen.

Interviews, Befragungen und Brainstorming zeigen, vor allem wenn standardisiert und vorstrukturiert konzipiert, jedoch immer nur Ausschnitte des kognitiven Gesamtgefüges eines Nutzers auf oder schränken den Nutzer durch das Vorgeben bestimmter Kriterien ein. Auch über offenes Brainstorming lässt sich nicht gewährleisten, dass eine abschließende Sammlung der Vorstellungen und Ziele der Nutzer stattfindet (folgt dieses doch einem assoziativen Ansatz, welcher, einmal eingeschlagen, nur einen Ast des gesamten kognitiven Gefüges abzubilden vermag). Die Erhebung „Mentale Modelle“, wie Young (2008) vorschlägt, kann diese methodische Lücke schließen. Offen bleibt die Wahl der Methode. Auf der Basis existierender Theorien über Mentale Modelle werden wir deswegen einen Weg mittels Struktur-Lege-Technik vorschlagen und am konkreten Beispiel beschreiben, um herauszufinden, ob der Einsatz einen Mehrwert für das Produkt bringen kann.

2 Entwicklung und Evaluation im Projekt edulap

Edulap (Educational Landscapes Psychology) ist eine webbasierte Such- und Orientierungsplattform für E-Learning-Ressourcen, welche im Rahmen des Innovations- und Kooperationsprojekts der Schweizerischen Universitätskonferenz (SUK) am Institut für Psychologie Zürich entwickelt wird (vgl. <http://www.edulap.ch>). Es werden dort alle elektronischen Ausbildungsangebote in der Schweiz (beginnend im Fach Psychologie) in einem Überblickssystem integriert (Streule & Läge, 2008).

Nutzer können auf dieser Plattform mittels Stichworten oder über Filter nach Ressourcen suchen (z.B. nach Institution, Zielgruppe, Sprache, Autor, Fachgebiet). Auch möglich sind semantische Ähnlichkeitsvergleiche eines Eingabetextes mit den im System erfassten Ressourcen. Die Suchresultate werden in leicht interpretierbaren „Landkarten“ aufgrund ihrer inhaltlichen Ähnlichkeit abgebildet. Je näher zwei Suchresultate (Punkte in der Ergebniskarte) beieinander liegen, desto grösser ist die inhaltliche Ähnlichkeit der E-Learning-Angebote (Abb. 1).

Das Orientierungskartensystem schließt damit eine Lücke, da Dozierende und Forscher eine Plattform nutzen können, welche ihnen die Suche nach potenziell nützlichen E-Learning Ressourcen über die Institutionsgrenzen erlaubt. Ganz im Sinne des aktuellen Trends, nämlich der Öffnung und dem Einsatz von Neuen Medien in Lehre und Forschung, erhöht sich durch solch ein System die Visibilität von digitalen Lernressourcen (oder führt diese sogar erstmals ein).

Solch ein Orientierungskartensystem muss, um funktional, effizient und zielführend zu sein (also nach ISO 9241 eine hohe Usability aufzuweisen), unterschiedlichsten, a priori nicht abschließend zu definierenden Anforderungen seitens des

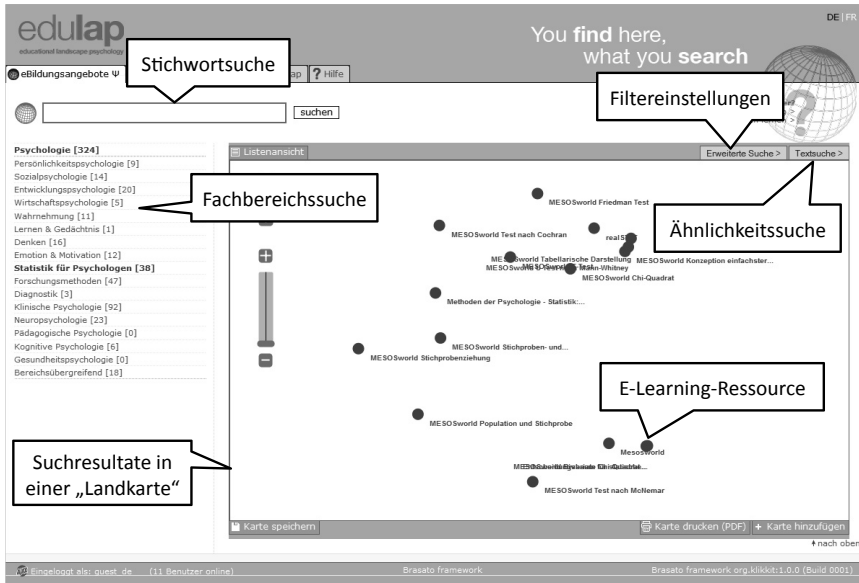


Abb. 1: Screenshot des Orientierungskartensystems edulap

Endnutzers gerecht werden. Die Software des Orientierungskartensystems wurde deswegen anhand des Software Life Cycle Processes i.S. eines klassischen Phasenmodells entwickelt (Schiedermeier, 2007). Die Entwicklung erfolgte inkrementell in mehreren Zyklen. Dabei wurde die Entwicklung in mehrere gleichartige Schritte zerlegt. Die Gesamtfunktionalität des Systems wuchs mit jeder Iteration.

Um nicht nur den praktischen Erfordernissen der Software gerecht zu werden, sondern das Produkt auch einer wissenschaftlichen Überprüfung zu unterziehen, wurden i.S. des Usability Engineerings die Ziele der Usability klar definiert (Anwendergruppen definieren, Zweck des Produktes, Messgrößen der Usability, Messinstrumente/-größen und kritische Werte festlegen, Usability Probleme qualitativ erheben). Dabei stand der Nutzereinbezug im Vordergrund. Auf Basis von existierenden Usability-Methoden wurden Konzept und Realisierung des Orientierungskartensystems immer wieder überprüft. Aus diesen Überlegungen ergab sich ein Evaluationskonzept, welches den Nutzer mit seinen Anforderungen, Erfahrungen, Bedürfnissen und Wünschen in formativen und summativen Evaluationen mit einbezog und in dem gängige Evaluationsmethoden, darunter Cognitive Walkthrough (vgl. Sarodnick & Brau, 2006), zum Einsatz kamen.

Entwicklung und Evaluationen von edulap sind demnach ausgerichtet auf den aktuellen Wissensstand des User Centered Design bzw. des Usability Engineerings und richten sich nach ISO Norm und Evaluationsstandards. Erwartet wurde, dass die im Evaluationskonzept ausgewählten Methoden und die Vorgehensmethoden der Softwareentwicklung ausreichen, um den Nutzerbedürfnissen und -ansprüchen gerecht zu werden. Die Mentalen Modelle der Nutzer sollten sich „indirekt“ im Feedback aus den Usability Tests zeigen und so eine Software schaffen, die einfach zu erlernen, effizient zu benutzen, einfach zu erinnern, angenehm zu bedienen ist und eine geringe Fehlerrate aufweist (Nielsen, 1993) und somit im Einklang mit den Vorstellungen, Wissen und Gefühlen der Benutzer ist. Ob diese Annahme gerechtfertigt und ob die Erhebung Mentaler Modelle vor dem Softwareentwicklungsbeginn einen Mehrwert für die Qualität und Funktionalität des Systems erbracht hätten, wurde im Folgenden in einer empirischen Studie überprüft.

3 Studie zur Ermittlung der Bedeutsamkeit Mentaler Modelle im Softwareentwicklungsprozess

3.1 Untersuchungsansatz und Methoden

In der wissenschaftlichen Forschung findet sich die Heidelberger Struktur-lege-Technik als Erhebungsmethode für Mentale Modelle (Scheele & Groeben, 1988). Ursprünglich für die Pädagogische Psychologie entwickelt, lassen sich Struktur-lege-Techniken (SLT) auch auf softwareergonomischem Gebiet anwenden. Jedoch ist dies bisher nur als Card Sorting (Methode zur Ermittlung von optimalen Navigationsstrukturen) belegt. Somit ist die Erprobung einer vollständigen SLT auf softwareergonomischem Gebiet neu und kann gleichzeitig als Überprüfung dienen, ob sie sich im Bereich Usability bewährt.

Die SLT basiert auf einem Interview, einer Konzeptualisierung des Interviews durch Begriffs- und Relationskarten, einer mit diesen Kärtchen durchgeführten Legung der Struktur (Mentales Modell) und einer abschliessenden Verifizierung des Modells.

3.2 Stichprobe und Durchführung

12 PsychologiestudentInnen der Universität Zürich nahmen an der Untersuchung teil. Das Durchschnittsalter betrug $M = 27.4$ Jahre (Min = 20 J., Max = 45 J.). Acht Personen studierten im 1. und 3. Semester, die weiteren im 5., 7. und 10. Semester.

Die Erhebung der Mentalen Modelle von Nutzern wurde nach Fertigstellung eines ersten, funktionsfähigen Prototyps durchgeführt (vgl. Abb. 1). Das Interviewthema war angelehnt an die benutzte Software von edulap „Suche nach E-Learning-Produkten im Internet“. Um die Konstruktionen der Nutzer zur Internetsuche aufzuspüren, wurde als Befragungsform ein halbstandardisiertes Interview eingesetzt, welches der Versuchsperson einen möglichst großen Spielraum zur Einbringung ihrer Perspektiven ermöglichte. Grundlage bildete ein thematisch strukturierter Interviewleitfaden, der dem Interviewverlauf sowie anzusprechender Themen einen groben Orientierungsrahmen gab.

Interessierende und gemäß der SLT möglichst offen gehaltenen Themengebiete waren Suchaktivitäten und Suchinteressen, Merkmale einer Suche, Suchbedingungen und Suchmethoden, Suchintention und Suchziele, Design und Workflow der Suche, Probleme bei der Suche/Faktoren die die Suche negativ beeinflussen, Suchergebnisse/Ergebnisdarstellung und Folgen und Konsequenzen aus einer Suche (insgesamt 42 Fragen).

Basierend auf dem Interview bereitete der Versuchsleiter das SLT-Strukturbild dreistufig vor: 1. Einordnung der Aussagen in die Themengebiete, 2. Extraktion der Kernaussagen auf Konzeptkarten zur Ermittlung der Hauptkonzepte und 3. Verbindung mit Relationskarten zu Modellstrukturen. In einem Folgetermin erfolgte zusammen mit der Versuchsperson die Validierung des aufgestellten Strukturmodells.

3.3 Resultate

Im Folgenden sind die inhaltlichen Auswertungen bezüglich der Inhalte der Strukturmodelle dargestellt. Diese zeigen, ob die Erhebung der Mentalen Modelle andere Inhalte, Funktionen oder Bedürfnisse beinhalten als im System edulap angedacht oder umgesetzt. Dafür wurden die Begriffe in den Strukturmodellen ausgewertet und ein Kategoriensystem erstellt. Im Mittel nannte jede Versuchsperson 7 Themen, die nicht in edulap enthalten sind (Minimum 3, Maximum 12 Themen). Dabei wurden teilweise gleiche Themen von mehreren Versuchspersonen genannt. Diese Themen lassen sich den Kategorien Suche, Filter, weiteren Funktionen (z.B. Web 2.0), Informationen zur Webseite, Hilfe, Ergebnisdarstellung und Informationen in den Ergebnissen zuordnen.

Abbildung 2 fasst die Themen mit den meisten Nennungen zusammen. Im Vergleich zum vorliegenden Orientierungssystem sind insgesamt 40 genannte Bereiche nicht umgesetzt. Um die Abweichungen der Mentalen Modelle von edulap sichtbar zu machen, wurden diskrepante Themen/Begriffe zusammengefasst (die in edulap, aber nicht im Mentalen Modell oder im Mentalen Modell,

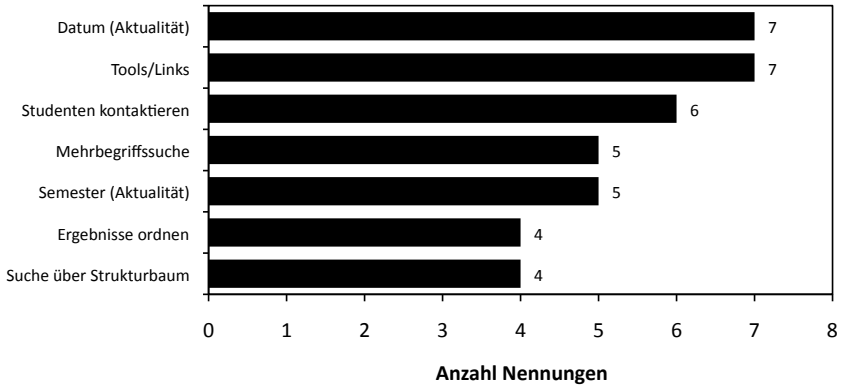


Abb. 2: Meistgenannte Themen, die nicht in edulap angedacht oder umgesetzt sind

aber nicht in edulap sind), Basis für die Ermittlung war ein Strukturmodell des bestehenden edulap-Systems, welches anhand des Kategoriensystems aufgestellt und um Themen, die nicht in Funktionalitäten und Handhabung relevant sind, bereinigt wurde (Abb. 3).

Abbildung 4 zeigt beispielhaft die von einer Versuchsperson gelegte Struktur. (Aus der systematischen Analyse dieser Strukturen greifen wir exemplarisch ein Resultat heraus. Ein Bericht über die Auswertung aller Modelle wäre sicher interessant, sprengt aber den für diesen Beitrag verfügbaren Platz.) Zunächst fällt bei dieser Person die im Vergleich zu den anderen Probanden differenzierte Vorstellung des Funktionsumfangs eines Suchsystems auf. Das SLT-Modell dieser Probandin widerspiegelt prototypisch die Resultate aus Abbildung 2. Die Person wünscht sich Informationen und Funktionen zu Tools/Links (linker oberer Quadrant in Abb. 4) oder Angaben zum Alter des Dokuments/Aktualität (rechter unterer Quadrant) genauso wie Funktionen in Richtung von Web 2.0, wie bspw. der Kontakt zu und der Austausch mit Mitstudierenden (linker oberer Quadrant).

4 Diskussion

Wie die Auswertung mit ihren grossen Unterschieden von Mentalen Modellen und tatsächlich umgesetztem System zeigt, lässt das Evaluationsobjekt edulap Themen, Informationen und Funktionen vermissen, die in den Mentalen Modellen der Versuchsperson enthalten sind. Mit Berücksichtigung des Mentalen Modells sähe demnach die edulap-Webseite trotz vorheriger Durchführung umfassender, klassischer Usability-Tests und Evaluationen in Teilen anders

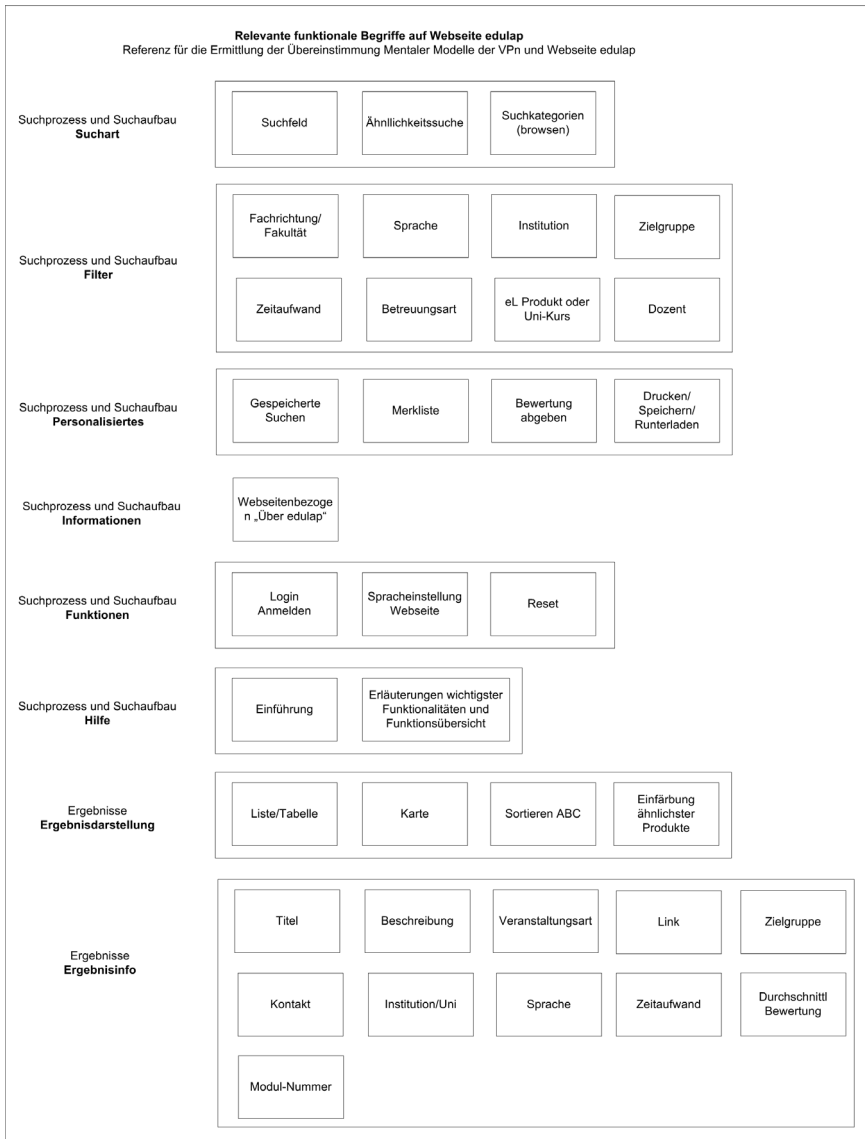


Abb. 3: Strukturmodell von edulap auf der Basis des Prototypen in Abb. 1

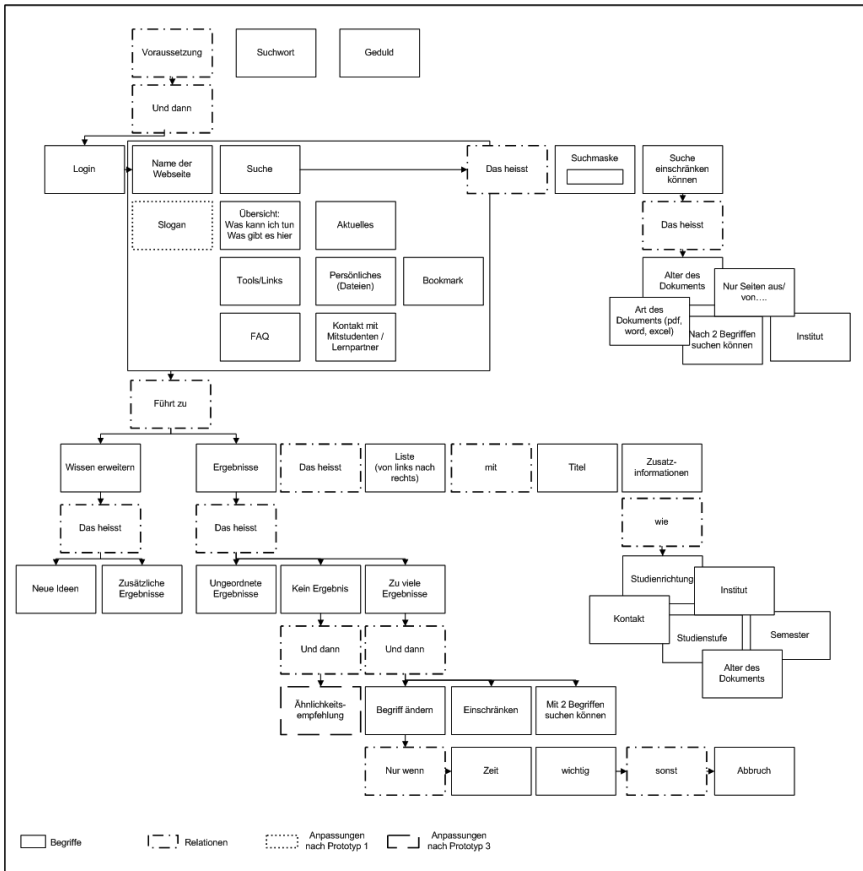


Abb. 4: Strukturlegemodell (Mentales Modell) einer Versuchsperson. Im SLT-Modell speziell markiert sind auch die Begriffe (Funktionen), welche bereits im Entwicklungsprozess und den verschiedenen Prototypenstadien umgesetzt worden sind.

aus: Nur das bereits Vorhandene wurde eben zuvor bewertet. Die Nutzer in den Usability Tests mit offenen Fragen dazu anzuhalten, zusätzliche Vorstellungen, Wünsche und Bedürfnisse zu äussern, wurde wohl faktisch durch die externe Aufmerksamkeitsfokussierung blockiert. Dies legt die Empfehlung nahe, dass die Partizipation des Nutzers 1.) schon vor Entwicklungsbeginn und 2.) über die Erhebung der Mentalen Modelle zum Nutzungskontext der Webseite als Zusatz zu herkömmlichen Verfahren erfolgen sollte. Dies schafft einen deutlichen Informationszugewinn zu Nutzerbedürfnissen und Vorstellungen für das Design- und Entwicklungsteam. Es wird vermieden, dass Designer oder Entwickler etwas „Falsches“ antizipieren, die Entwicklung in suboptimale

Bahnen lenken und dadurch zu erhöhten Kosten für zu spät entdeckte Bedürfnisse oder Usability-Mängel führen. Deshalb sollte die Erhebung Mentaler Modelle als Evaluationsmethode in der Analysephase des Usability Engineering aufgenommen werden, als Ergänzung von Use Cases dienen und in die Anforderungsanalyse und Aufgabenanalyse und das Interaktionsdesign einfließen. So lassen sich spätere Korrekturen oder der spätere Einbau von zusätzlichen Funktionen, die nicht bedacht wurden, umgehen. Deutlich wird anhand dieser Analyse darüber hinaus ein weiterer Vorteil der SLT im Vergleich zu klassischen Anforderungsanalysen: Explizit gemacht werden durch dieses Vorgehen nicht nur die Funktionen an sich, sondern durch die relationale Verknüpfung der Begriffe auch der Workflow innerhalb eines Systems. Die Berücksichtigung dieser Prozesse in der Entwicklung erhöht zusätzlich die Zufriedenheit und die Effizienz in der Handhabung einer Software.

Literatur

- ISO 9241-11 (1998). *Ergonomic requirements for office work with visual display terminals – Part 11: Guidance on usability*. Verfügbar unter: <http://www.it.uu.se/edu/course/homepage/acsd/vt10/ISO9241part11.pdf> [23.02.2011].
- Herczeg, M. (2005). *Software-Ergonomie: Grundlagen der Mensch-Computer-Kommunikation* (Vol. 2). München: Oldenbourg.
- Nielsen, J. (1993). *Usability Engineering*. Boston, Mass.: Acad. Press.
- Nielsen, J. (1994). *Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier*. Verfügbar unter http://www.useit.com/papers/guerrilla_hci.html [11.03.2011]
- Nielsen, J. (2003). *Paper Prototyping: Getting User Data Before You Code*. Verfügbar unter: Jakob Nielsen's Alertbox, <http://www.useit.com/alertbox/20030414.html> [14.04.2003].
- Nielsen, J. (2010). *Mental Models*. Verfügbar unter: Jakob Nielsen's Alertbox, <http://www.useit.com/alertbox/mental-models.html> [18.10.2010].
- Norman, D. A. (1983). Some Observations on Mental Models. In A. L. S. D. Gentner (Ed.), *Mental Models* (pp. 99-127). Hillsdale, NJ.
- Rauterberg, M. (1991). Partizipative Konzepte, Methoden und Techniken zur Optimierung der Softwareentwicklung. *Softwaretechnik-Trends*, 11(3), 104-126.
- Sarodnick, F. und Brau, H. (2006). *Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung*. Bern: Huber.
- Scheele, B & Groeben, N. (1988). *Dialog-Konsens-Methoden zur Rekonstruktion Subjektiver Theorien: Die Heidelberger Struktur-Lege-Technik (SLT), konsensuale Ziel-Mittel-Argumentation und kommunikative Flussdiagramm-Beschreibung von Handlungen*. Tübingen: Francke.
- Schiedermeier, R. (2007). *Vorlesung Programmieren I (5.5 Phasenmodell der Softwareentwicklung)*. Verfügbar unter: <http://www.mi.uni-koeln.de/c/mirror/f7alpha1.informatik.fhmuennenchen.de/~schieder/programmieren-1-ws96-97/phasenmodell.html> [17.03.2009].

- Staggers, N. N., A.F. (1993). Mental Models: concepts for human computer research. *International Journal of Man-machine Studies*, 38, 587-605.
- Streule, R. & Läge, D. (2008). Educational Landscapes: Mapping der elektronischen Ausbildungsangebote eines Faches mit Kognitiven Karten. In S. Zauchner, P. Baumgartner, E. Blaschitz & A. Weissenböck (Hrsg.), *Offener Bildungsraum Hochschule – Freiheiten und Notwendigkeiten* (S. 50-57). Münster: Waxmann.
- Young, I. (2008). *Mental Models – Aligning Design Strategy with Human Behavior*. New York: Rosenfeld Media.