

Evaluation bibliothekarischer Online-Angebote

Usability und Usefulness – zwei Seiten einer Medaille

Thomas Weinhold, Jasmin Hügi, René Schneider, Bernard Bekavac

„A product by itself has no value; it has value only insofar as it is used. Use implies users.“

Dumas/Redish (1999:4)

Abstract: Aus Anwendersicht sind für die Qualität bibliothekarischer Online-Angebote zwei Aspekte entscheidend. Einerseits müssen für sie relevante Inhalte zur Verfügung stehen. Andererseits sollte es möglichst einfach und intuitiv möglich sein, diese Inhalte zu finden und auf sie zuzugreifen. Die Themen Usefulness und Usability sind daher für Bibliotheken von grosser Bedeutung. Der Beitrag beschreibt wie diese beiden Aspekte evaluiert werden können und wie Bibliotheken anhand von BibEval und PECI selbstständig ihre Angebote überprüfen können.

Abstract: From the users' point of view two aspects are essential for the quality of library online services. On the one hand, relevant content must be available. On the other hand, it should be simple and intuitive to find and to access this content. Therefore, the topics usability and usefulness are of great importance for libraries. This article describes how these two aspects can be evaluated and how libraries can use BibEval and PECI for assessments of their own services.

1. Einleitung

Online-basierte Angebote von Bibliotheken (Online Library Services) gewinnen immer stärker an Relevanz. Einen wesentlichen Erfolgsfaktor stellt dabei die Usability, beziehungsweise die Anwenderfreundlichkeit oder auch Gebrauchstauglichkeit dieser Angebote dar. Nach Nielsen sind für Benutzungsschnittstellen in diesem Zusammenhang neben der Erfüllung der Nutzerbedürfnisse insbesondere die Erlernbarkeit, die Effizienz, der Wiedererkennungswert sowie die Fehlerrate von besonderer Bedeutung (vgl. Nielsen 1993:25ff). Diese Eigenschaften finden sich grösstenteils auch in der ISO-Norm 9241-11 wieder, wonach Usability definiert wird als „das Ausmass, in dem ein Produkt von einem bestimmten Benutzer verwendet werden kann, um bestimmte Ziele in ei-

nem bestimmten Kontext effektiv, effizient und zufriedenstellend zu erreichen“ (ISO 9241-11 1998). Nielsen (2012) definiert Usability als ein Qualitätsmerkmal, das ausdrückt, wie einfach ein Produkt zu bedienen ist, beziehungsweise wie gut Anwender/-innen mit der Benutzerschnittstelle eines Systems zurechtkommen.

In Bezug auf die von Anwender/-innen wahrgenommene Qualität eines interaktiven Produkts wird zunehmend auch von der sogenannten User Experience gesprochen. Während sich der Begriff der Usability ausschliesslich auf die eigentliche Nutzungssituation bezieht, erweitert die User Experience dieses Konzept um die Antizipation, also die angenommene oder vorgestellte Nutzung sowie die Verarbeitung der Nutzungssituation (vgl. Geis 2010). Seit 2010 ist User Expe-

rience als genormter Begriff in der ISO-Norm 9241-210 enthalten. Definiert wird der Begriff dort als die „Wahrnehmungen und Reaktionen einer Person, die aus der tatsächlichen und/oder der erwarteten Benutzung eines Produkts, eines Systems oder einer Dienstleistung resultieren“ (ISO 9241-210 2010).

2. Ausgangslage und Problemstellung

Die Usability hat sich in den letzten Jahren zu einem der zentralen Qualitätsmerkmale jeglicher interaktiver Produkte, also auch von Software und Websites beziehungsweise Webanwendungen entwickelt. Das Thema ist daher auch für Bibliotheken von zentraler Bedeutung. Denn in jedem Fall ist es so, dass eine schlechte Gebrauchstauglichkeit negative Auswirkungen auf die Nutzerzahlen eines Angebots hat.

Hinsichtlich der von Anwender/-innen wahrgenommenen Qualität bibliothekarischer Online-Angebote sind vor allem zwei Aspekte von Bedeutung. So werden die Standards für einfach bedienbare Recherchewerkzeuge seit Jahren von den gängigen Internet-Suchmaschinen wie Google oder Bing gesetzt. Nutzer/-innen erwarten daher mittlerweile einen einfachen Suchschlitz sowie den Zugriff auf elektronische Volltexte, wie das bei diesen Angeboten der Fall ist. Erweiterte Suchformulare, Operatoren und ähnliches werden hingegen von typischen Nutzer/-innen nur selten verwendet. Um am Markt bestehen zu können, müssen sich Bibliotheken daher mit ihren Online-Angeboten in gewisser Weise Internet-Suchmaschinen anpassen. Ein Beleg hierfür sind die zahlreichen, in den letzten Jahren durchgeführten Projekte zur Umstellung auf sogenannte „Next Generation Library Catalogs“ beziehungsweise „Discovery Tools“. Bibliotheken stehen also zunehmend vor der Herausforderung, ihre Online-Angebote aus Nutzersicht zu optimieren. Dabei muss eine grosse Menge an Informationen in übersichtlicher und einfach handhabbarer Form bereitgestellt werden, was neben etablierten bibliothekarischen Tätigkeitsbereichen ein neues, technologie- und innovationsorientiertes Aufgabenportfolio mit sich bringt (vgl. Weinhold/Öttl/Bekavac 2011:11).

Entscheidend für die Qualität entsprechender Angebote ist, dass deren Konzeption und Entwicklung sich stets an den Fähigkeiten und Bedürfnissen der Benutzer/-innen orientieren sollte. Eine Herausforderung in diesem Kontext ist der Umstand, dass die Dienstleistungen von Bibliotheken und insbesondere jene von öffentlichen Bibliotheken auf heterogene Gruppen von Anwender/-innen mit unterschiedlichen Fähigkeiten und Kenntnissen ausgerichtet sind (vgl. Battleson/Booth/Weintrop 2001:190). Dabei kann sich

zudem der Informationsbedarf der einzelnen Nutzer/-innen deutlich unterscheiden. Je nach Informationsbedarf können in diesem Kontext grundsätzlich drei Recherchestrategien verfolgt werden (vgl. Weinhold/Hamann/Bekavac 2011:32):

- Gezielte Suche nach einem speziellen Objekt/einer speziellen Information;
- Erkunden, was zu einem bestimmten Thema/Autor/Inhalt et cetera vorhanden ist;
- Weiterverwertung von Rechercheergebnissen (Objekt/Trefferliste), zum Beispiel Speichern, Zitieren, Taggen, und so weiter;

Vor diesem Hintergrund gilt es für Bibliotheken, das Potenzial digitaler Technologien möglichst optimal auszuschöpfen, um Benutzer/-innen einen möglichst einfachen und nahtlosen Zugang zu ihren Beständen anbieten zu können (vgl. Norlin/Winters 2002:1).

3. Evaluation bibliothekarischer Online-Angebote

Idealerweise sollte bei der Evaluation bibliothekarischer Online-Angebote ein möglichst ganzheitlicher Ansatz verfolgt werden, der unterschiedliche Betrachtungsebenen umfasst. Hierbei kann man sich am von Tsakonias/Papatheodorou (2006) entwickelten Interaktions-Triptychon orientieren, das in Abbildung 1 dargestellt ist.

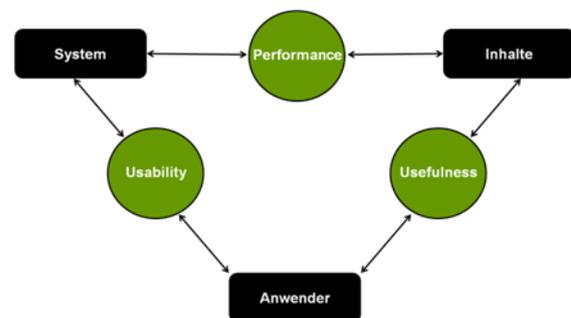


Abb. 1: Interaktions-Triptychon (vgl. Tsakonias/Papatheodorou 2006:402)

Dabei sind die drei Ebenen Usability, Usefulness und Performance von Bedeutung. Die Usefulness wird im Allgemeinen im Bereich der Human-Computer-Interaction (HCI) gemeinsam mit der Usability als integraler Bestandteil eines ganzheitlichen Ansatzes betrachtet. Im Bereich digitaler Bibliotheksangebote bezieht sich die Usefulness vor allem auf die bereitgestellten Inhalte und deren Relevanz für die Erfüllung der Informationsbedürfnisse der Anwender/-innen. Die Performance, die zwischen System und Inhalten an-

gesiedelt ist, beschreibt die Leistungsfähigkeit eines Systems. Sie ist stark von den verwendeten Datenformaten, inhaltlichen Strukturen und der Informationsrepräsentation abhängig (vgl. Fuhr et al. 2007:26).

Arten und Methoden der Usability-Evaluation

Mittlerweile existiert eine Vielzahl unterschiedlicher Methoden, die für Usability-Evaluationen herangezogen werden können. Eine mögliche Art der Unterscheidung der einzelnen Methoden basiert auf dem Zeitpunkt, zu dem die Evaluation durchgeführt wird. Dabei unterscheidet man zwischen formativen und summativen Verfahren. Die formative Evaluation findet bereits während der Entwicklung eines Produktes statt, wobei konkrete Verbesserungsmöglichkeiten im Vordergrund stehen. Summative Evaluationen erfolgen dagegen meist nach Abschluss der Produktentwicklung, da durch diese Verfahren hauptsächlich Urteile und Gesamtschätzungen ermittelt werden (vgl. Nielsen 1993:170).

Ein weiteres Unterscheidungskriterium basiert darauf, welche Art von Gutachter/-innen für die Evaluation herangezogen wird. Man unterscheidet zwischen den expertenorientierten oder analytischen Methoden und den benutzerorientierten beziehungsweise empirischen Methoden. Bei den benutzerorientierten Methoden der Usability-Evaluation werden künftige, reale Benutzer/-innen mit dem Produkt konfrontiert und hierzu befragt oder während der Interaktion mit diesem Produkt beobachtet (vgl. Schweibenz/Thissen 2003:43). Dagegen werden bei der Verwendung von expertenorientierten beziehungsweise analytischen Methoden Experten als Ersatz-Benutzer/-innen für die Untersuchung herangezogen. Diese untersuchen Benutzungsschnittstellen mit Hilfe vorgegebener Richtlinien und anhand ihrer Erfahrung auf mögliche Usability-Probleme (vgl. Heuer 2003:116).

Expertenbasierte Evaluationsverfahren

Das wohl bekannteste analytische Evaluationsverfahren ist die sogenannte heuristische Evaluation. Im Rahmen einer heuristischen Evaluation überprüft eine Reihe von Experten/-innen, inwiefern das User Interface eines Produktes mit bestimmten anerkannten Prinzipien des Usability-Engineerings übereinstimmt. Diese Prinzipien werden als Heuristiken bezeichnet, wobei es sich um Richtlinien für die benutzerfreundliche Gestaltung von Benutzungsschnittstellen handelt, die auf Basis empirischer Erkenntnisse entwickelt wurden (vgl. Schweibenz/Thissen 2003:100).

Einen weiteren Ansatz, der recht ähnlich zur eben beschriebenen heuristischen Evaluation ist, stellen Kriterienkataloge dar. Hierzu gibt es in der Literatur je nach Autor und teilweise auch nach spezifischem Einsatzzweck eine Reihe unterschiedlicher Bezeichnungen wie zum Beispiel Experten-Leitfäden, Design-Guidelines, Gestaltungsrichtlinien und so weiter. Für die Erstellung eines Kriterienkatalogs werden die für ein Produkt als relevant erachteten Merkmale zusammengetragen. Diese können gegebenenfalls je nach Relevanz unterschiedlich gewichtet werden. Im Unterschied zu den vorhin vorgestellten Heuristiken weisen solche Listen üblicherweise einen höheren Detaillierungsgrad auf.

Beim Cognitive Walkthrough handelt es sich um eine Methode, mit der aufgabenorientierte Abläufe evaluiert werden können. Dabei erfolgt mittels vorab definierter Leitfragen eine strukturierte, schrittweise Untersuchung der Aufgaben, die mit dem Produkt erfüllt werden sollen (vgl. Schweibenz 2011:13). Dieses Verfahren basiert auf der Beobachtung, dass viele Benutzer/-innen die Funktionen eines Produktes durch Ausprobieren erlernen, wobei das Explorieren eine zentrale Rolle spielt. Ziel ist es dabei, Produkte zu entwickeln, die von den Nutzer/-innen leicht erlernbar sind (vgl. Sarodnick/Brau 2011:152).

Benutzerorientierte Evaluationsverfahren

Die einfachste Form benutzerorientierter Usability-Evaluationsverfahren stellen Fragebögen dar. Beispielsweise können diese genutzt werden, um festzustellen, ob ein Produkt generell bestimmten Kriterien entspricht, wie beispielsweise den Dialogprinzipien der ISO 9241-110. Darüber hinaus werden Fragebögen aber auch im Anschluss an Usability-Tests eingesetzt, um ein Gesamturteil zu einem Produkt zu bekommen (vgl. Burmester 2003:107). Mittlerweile existieren für unterschiedliche Untersuchungszwecke eine Reihe verschiedener, vorgefertigter Fragebögen. Beispiele dafür sind ISONorm110, IsoMetrics, AttrakDiff oder auch der User Experience Questionnaire (UEQ)¹.

Eine weitere Form der Befragung stellen Focus Groups beziehungsweise Fokusgruppen-Interviews dar. Dabei handelt es sich um strukturierte Diskussionen zu vorgegebenen Themen, die von einem geschulten Moderator geleitet werden. Ein typisches Fokusgruppen-Interview wird in der Regel mit acht bis zwölf Teilnehmer/

¹ Eine Beschreibung der genannten Fragebögen ist zum Beispiel auf den Webseiten des Usability-Labors der HTW Chur unter <http://www.cheval-lab.ch/was-ist-usability/usabilitymethoden/fragebogen/> [Stand: 03.12.2013] verfügbar.

-innen durchgeführt und dauert üblicherweise etwa zwei Stunden, wobei die Teilnehmer/-innen der Fokusgruppe potentielle Benutzer/-innen des zu untersuchenden Produktes sein sollten (vgl. Dumas/Redish 1999:44). Von entscheidender Bedeutung für den Ablauf der Diskussion in einer Fokusgruppe ist, dass der thematische Schwerpunkt mit Hilfe eines Stimulus gesetzt wird. Durch diesen Stimulus sollte die Phantasie der Diskussionsteilnehmer/-innen angeregt werden. Zu diesem Zweck eignen sich zum Beispiel einführende Texte, Bilder oder auch Videos (vgl. Schweibenz/Thissen 2003:122).

Die bisher vorgestellten empirischen Verfahren dokumentieren üblicherweise Aussagen über ein potenzielles Verhalten der Benutzer/-innen. Hierbei besteht die Problematik, dass Aussagen von Benutzer/-innen über ihr eigenes Verhalten verzerrt sein können. Ausserdem finden sich Benutzer/-innen durchaus mit gewissen Schwachstellen ab, so dass sie diese gar nicht kommunizieren. Hinzu kommt, dass Probanden teilweise auch nicht richtig ausdrücken können, wo jetzt genau das Problem beziehungsweise die Schwachstelle eines Systems zu sehen ist.

Ein Verfahren, das diesen Aspekt berücksichtigt, ist das sogenannte formative Usability-Testing. Im Rahmen eines Usability-Tests bearbeiten Testpersonen mit dem zu untersuchenden Produkt realistische Aufgaben, wobei sie von Usability-Experten/-innen beobachtet werden. Auf Basis der Äusserungen der Probanden, Beobachtungen des Testleiters oder auch auf Grund von Messungen (zum Beispiel der Zeitdauer für die Aufgabenbearbeitung) können Probleme im Umgang mit dem System erkannt und Verbesserungsmöglichkeiten erarbeitet werden (vgl. Sarodnick/Brau 2011:163). Entscheidend für den Nutzen der Untersuchung ist die möglichst repräsentative Auswahl der Testpersonen. Zwar lassen sich bereits mit 4-6 Personen verlässliche Ergebnisse erzielen, üblicherweise werden aber 8-12 Personen als Probanden/-innen herangezogen. Usability-Tests können in verschiedenen Varianten und mit unterschiedlichem technischem und finanziellem Aufwand durchgeführt werden. In der Regel erfolgt eine Videoaufzeichnung der Interaktion der Testpersonen mit dem System anhand spezieller Software, bei der zum Beispiel Mausclicks aufgezeichnet werden und die Mimik der Probanden per Webcam mitgeschnitten wird. Üblicherweise werden die Nutzer/-innen zudem aufgefordert, gleichzeitig ihre Handlungen zu beschreiben (Methode des lauten Denkens), was für die Auswertung von Vorteil ist. Alternativ kann das aufgezeichnete Video (oder Ausschnitte davon) auch nach den Tests mit dem Probanden gemeinsam diskutiert werden – in Form von sogenannten Post-Video-Sessions beziehungsweise

mittels des retrospektiven lauten Denkens. Noch genauere Analysen ermöglicht der Einsatz von Eyetracking-Systemen. Diese Geräte zeichnen die exakten Blickpositionen der Benutzer/-innen auf dem Bildschirm auf. Dabei lässt sich zum Beispiel herausfinden, wohin die Anwender/-innen zuerst schauen, ob wichtige Bereiche eines User Interfaces schnell erkannt werden und welche Bereiche eines User Interface besonders intensiv wahrgenommen werden.

Modelle der Usefulness-Evaluation

Das Prinzip der Nützlichkeit wurde bereits 1789 durch den Philosophen Jeremy Bentham beschrieben: „Unter Nützlichkeit ist jene Eigenschaft an einem Objekt zu verstehen, durch die es dazu neigt, Gewinn, Vorteil, Freude, Gutes oder Glück hervorzubringen (...).“ (zitiert nach Hörz 2009)

Wendet man dieses Prinzip auf digitale Angebote von Bibliotheken an, dann sind diese dann nützlich, wenn sie den Benutzer/-innen das bieten, was sie benötigen. Das Angebot muss dem Informationsbedürfnis entsprechen, damit am Ende der Recherche Antworten auf die Fragen gefunden werden und diese „erfolgreich“ abgeschlossen werden kann. Die inhaltsorientierte Analyse der Usefulness ist damit für Bibliotheken ebenso wichtig, wie die eher oberflächen- und funktionsorientierte Analyse der Usability. Wünschenswert wären daher hybride Verfahren, mit denen beide Ebenen untersucht werden können. Leider existieren solche Ansätze in der Praxis jedoch noch nicht, und generell ist im Bereich der Evaluation der Usefulness aktuell noch eine methodologische Armut zu beobachten. Das Mittel der Wahl sind häufig Fragebögen – nicht zuletzt deshalb, weil die bislang entwickelten Modelle zur Evaluation der Usefulness digitaler Bibliotheksangebote wenig Spielraum für andere Methoden zulassen (vgl. Birri Blezon/ Hügi/Schneider 2011:55ff).

Einer der ersten Ansätze, der unmittelbar auf digitale Bibliotheksangebote ausgerichtet ist, wurde von Fuhr et al. (2001) entwickelt. Dieses Modell besitzt vier Dimensionen: Daten/Bestand, System/Technologie, Benutzer/-innen und Benutzung/ Umgebung. Die Definition der Benutzergruppen bestimmt den geeigneten Inhalt des Bestands. Der Bestand bestimmt wiederum die geeigneten Technologien. Die Attraktivität des Bestands für die Nutzer/-innen und die einfache Bedienung des Systems bestimmen schliesslich das Ausmass der effektiven Systemnutzung (vgl. Fuhr et al. 2001:191).

Die Interaktion zwischen Benutzer/-innen, Inhalten und System steht auch beim bereits vorgestellten Interaktions-Triptychon im Fokus. Die

Relation der Usefulness, die sich aus der Beziehung zwischen Nutzer/-innen und Inhalt ergibt, wird nachfolgend näher beschrieben. Sie kann anhand von verschiedenen Eigenschaften gemessen werden:

- der Relevanz der gefundenen Information, also der Angemessenheit in Bezug auf den geäußerten Informationsbedarf;
- der Brauchbarkeit, also der Wichtigkeit der Information für die zu lösende Aufgabe;
- des Formats, in dem die Information zur Verfügung gestellt wird;
- der vom/von der Benutzer/-in wahrgenommenen Glaubwürdigkeit in Bezug auf den beziehungsweise die Informationsvermittler/-innen;
- des sogenannten Levels, das heisst des Erschließungsgrads, mit dem die Information zur Verfügung gestellt wird: etwa als Abstract oder Volltext, mit bibliographischer Information und Metadaten und ähnliches;
- des zeitlichen Abdeckungsgrads, also der Aktualität der Quelle oder ihrer diachronen Vollständigkeit;

Daneben existieren noch einige andere Ansätze, wie beispielsweise der Framework von Sandusky (2002) oder das Modell von Saracevic und Covi (2000). Beide versuchen digitale Bibliotheken anhand von Attributen, die sich auf alle Typen von digitalen Bibliotheken anwenden lassen, zu evaluieren. Sandusky (2002:37ff) definiert zum Beispiel sechs Ebenen: Publikum, Institution, Zugang, Inhalt, Dienstleistungen und Design/Entwicklung. Saracevic und Covi (2000:348) definieren sieben Attribut-Levels, die sich in nutzerorientiert und systemorientiert einteilen lassen.

Es wird jedoch ersichtlich, dass es nicht einfach ist, die Usefulness zu evaluieren. Die Herausforderung besteht darin, für das subjektive Prinzip der Nützlichkeit objektive Bewertungskriterien zu finden, die allgemein gültig sind für alle Arten von digitalen Bibliotheksangeboten. Ebenso, wie jede/r Benutzer/-in eine eigene Vorstellung davon hat, was für sie/ihn nützlich ist, haben auch Evaluationsexperten/-innen ihre eigene Definition der Nützlichkeit und versuchen, diese mit unterschiedlichen Attributen und Ebenen zu definieren. Darüber hinaus wird noch zu selten darauf eingegangen, wie die Evaluation anhand dieser Modelle umgesetzt werden kann (vgl. Birri Bleton/Hügi/Schneider 2011:60).

4. Anwendung von Evaluationsmethoden in der Praxis

Für die Evaluation der Usability steht, wie im vorangegangenen Kapitel bereits dargestellt, ein breites Spektrum unterschiedlicher Methoden zur Verfügung. Neben der Durchführung von Usability-Tests mit Benutzer/-innen wird im Bibliotheksbereich dabei insbesondere häufig auf die heuristische Evaluation zurückgegriffen. Ein Grund hierfür ist der Umstand, dass Heuristiken praktisch über den gesamten Produktentwicklungszyklus hinweg eingesetzt werden können (George 2008:130).

Es ist jedoch anzumerken, dass die in gängigen Heuristiken enthaltenen Kriterien, wie zum Beispiel die 10 Bewertungsaspekte, die Nielsen (1994:30) in seinen Heuristiken präsentiert, recht abstrakt gehalten sind. Aus diesem Grund ist ein gewisses Mass an Hintergrundwissen aus den Bereichen Usability-Engineering und Web beziehungsweise Interaction Design erforderlich, um auf deren Basis eine detaillierte Analyse bestimmter Komponenten einer Website vornehmen zu können.

Um hingegen auch Personen ohne umfassende Erfahrung in der Durchführung von Usability-Evaluationen die Möglichkeit zu geben, eine Analyse der Schwachstellen bibliotheksspezifischer Online-Angebote vornehmen zu können, sind detailliertere Kriterienkataloge erforderlich, die möglichst optimal auf die Branchenspezifika im Bibliotheks- und Informationssektor zugeschnitten sein sollten. Ein solcher Kriterienkatalog, der im Rahmen des Kooperations- und Innovationsprojektes „e-lib.ch: Elektronische Bibliothek Schweiz“ entwickelt wurde, wird im folgenden Kapitel vorgestellt.

In Bezug auf die Evaluation der Usefulness überwiegt, wie bereits im vorangegangenen Kapitel angedeutet, aktuell noch die Entwicklung theoretischer Modelle, wenngleich in den vergangenen Jahren erste konkrete Evaluationen der Usefulness digitaler Bibliotheksangebote durchgeführt worden sind. In diesen wurden primär unterschiedlichste Methoden aus dem Bereich der Usability-Evaluation auf ihre Anwendbarkeit in Bezug auf die Analyse der Usefulness eines Angebots hin untersucht. Dazu zählen unter anderem Fragebögen, strukturierte Interviews, Fokusgruppen, Tests mit lautem Denken, Fallstudien und Logfile-Analysen (vgl. Birri Bleton/Hügi/Schneider 2011:60ff).

Den derzeit wohl verbreitetsten Ansatz zur Evaluation der Usefulness stellen, wie bereits erwähnt, Fragebögen dar. So haben beispielsweise Kani-Zabihi/Ghinea/Chen (2006) auf diese Weise die Angebote ScienceDirect und die Classical Music Library untersucht. Von Xie (2006, 2008)

wurden zwei Umfragen durchgeführt, deren Ziel darin bestand, zu ermitteln, welche Evaluationskriterien Benutzer/-innen für die Evaluation digitaler Bibliotheksangebote heranziehen würden. Tsakonas/Papathodorou (2006) führten eine Umfrage zur Evaluation von elektronischen Informationsdienstleistungen durch. Es sollte festgestellt werden, welche Funktionalitäten und Inhalte am wichtigsten sind und den Erfolg der Aufgabe beeinflussen. Darüber hinaus sollte auch die Beziehung zwischen Usability und Usefulness aufgezeigt werden.

Im Rahmen des Forschungsprojekts PECEI der Haute école de gestion de Genève wurde dazu eine Online-Plattform entwickelt, mit der Informationszentren die Usefulness ihres Online-Angebots selbstständig evaluieren können. Diese soll im weiteren Anschluss genauer vorgestellt werden.

5. BibEval – ein Kriterienkatalog für die Usability-Evaluation bibliothekarischer Online-Angebote

Das Schweizerische Institut für Informationswissenschaft (SII) in Chur hat im Rahmen von e-lib.ch einen spezifisch auf die Gegebenheiten und Inhalte von Bibliotheken zugeschnittenen Kriterienkatalog zur Evaluation entsprechender Online-Angebote entwickelt. Dieser soll Bibliotheken in die Lage versetzen, bis zu einem gewissen Grad selbstständig Usability-Evaluationen der eigenen Angebote durchführen zu können.

Die Entwicklung des Kriterienkatalogs „BibEval“ erfolgte in einem iterativen Prozess. Zu Beginn wurde mittels einer Literaturrecherche der State-of-the-Art zur Durchführung von Usability-Evaluationen im Bibliothekssektor erhoben. Im Rahmen einer Best-Practice-Analyse bei der eine Vielzahl von Bibliothekswebsites untersucht wurde, wurden ausserdem gängige Bestandteile entsprechender Angebote sowie Kriterien für deren Bewertung ermittelt. Bei der Ausarbeitung der Bewertungskriterien wurden zusätzlich diverse allgemeine Normen, Richtlinien und Heuristiken wie beispielsweise die ISO-Norm 9241 herangezogen (vgl. Weinhold/Hamann/Bekavac 2011:38). Auf dieser Basis wurde eine erste Version des Leitfadens erstellt, die im Rahmen einer Fokusgruppe mit Experten aus unterschiedlichen Bereichen (Bibliotheken, Web Design, Usability) diskutiert und weiter verfeinert wurde. Das Resultat ist eine modular verwendbare, hierarchisch strukturierte Liste von Evaluationskriterien.

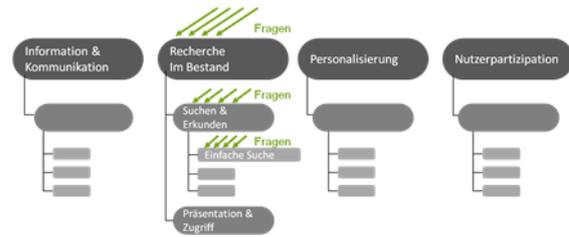


Abb. 2: Struktur des Kriterienkatalogs „BibEval“

Der Leitfaden ist in drei Ebenen gegliedert und deckt die Bereiche „Information und Kommunikation“ (Informationen & Dienstleistungen der Bibliothek), „Nutzerpartizipation“ (Web 2.0), „Personalisierung“ (individuelle Anpassungsmöglichkeiten der Webseite beziehungsweise –anwendung) und „Katalog“ (Recherche und Weiterverwertung der Treffer) ab (vgl. Abbildung 2). Entsprechend ihrer Funktionalität und ihrem Einsatzzweck sind diesen Bereichen einzelne, unterscheidbare Komponenten zugeordnet. Für alle Komponenten wird eine Einstufung vorgenommen, wobei je nach Relevanz und Verbreitungsgrad eine Komponente dabei entweder als optional oder obligatorisch klassifiziert werden kann. Dies ermöglicht den Nutzer/-innen des Leitfadens eine einfache Bewertung der relativen Wichtigkeit eines Features für den eigenen Web-auftritt vorzunehmen. Für jeden Bereich und jede Komponente wurden anschliessend detaillierte Fragen erarbeitet, die es den Anwender/-innen erlauben, die konkrete Umsetzung der einzelnen Bausteine zu bewerten, wobei ebenfalls wieder zwischen verbindlichen und optionalen Fragen unterschieden werden kann.

Um im Rahmen von Evaluationen eine Einschätzung der Auswirkungen der identifizierten Usability-Probleme vornehmen zu können, sollte ein sogenanntes Severity Rating vorgenommen werden. Ein solches Rating kann zum Beispiel genutzt werden, um die Prioritäten für ein etwaiges Re-Design festlegen zu können. Hierfür wird bei „BibEval“ in Anlehnung an Nielsen (1994) die nachfolgend dargestellte Bewertungsskala verwendet.

Rating	Beschreibung
0	Nicht zutreffend
1	Kein Usability-Problem
2	Kleines Usability-Problem
3	Mittleres Usability-Problem
4	Schweres Usability-Problem
5	Nicht umgesetzt, obwohl notwendig

Abb. 3: Kategorien des Severity-Ratings

Um die modulare Struktur des Kriterienkatalogs in der Praxis möglichst optimal nutzen zu können, steht dieser in Form einer interaktiven Webanwendung, die auf Basis von MySQL, PHP und JavaScript beziehungsweise AJAX entwickelt wurde, auf den Seiten des Usability-Labors der HTW Chur zur kostenlosen Nutzung zur Verfügung.²

Entsprechend der zuvor beschriebenen Struktur des Kriterienkatalogs ist auch die Webanwendung in drei Bereiche gegliedert. Anwender/-innen können somit Schritt für Schritt den Fokus der Evaluation des zu untersuchenden Online-Angebots festlegen und diese durchführen. Hierfür sind zunächst die Bereiche sowie die Komponenten beziehungsweise Funktionalitäten auszuwählen, die untersucht werden sollen. Die einzelnen Dialogschritte werden hierbei dynamisch angepasst, was es den Anwender/-innen erlaubt, Einstellungen jederzeit zu modifizieren. Sie erhalten somit stets unmittelbares Feedback über die von ihnen getroffenen Auswahlentscheidungen. Im dritten Bereich der Webanwendung bekommen Anwender/-innen entsprechend der getroffenen Auswahl ihren individuellen Fragenkatalog angezeigt, der dann schrittweise abgearbeitet werden kann.

Hierfür stellt die Webanwendung für jede Frage ein entsprechendes Auswahlfeld zur Verfügung, über das eine Bewertung gemäss der in Abbildung 3 dargestellten Kategorien des Severity Ratings vorgenommen werden kann. Zusätzlich kann bei Bedarf bei jeder Frage auch ein zusätzliches Kommentarfeld eingeblendet werden, über welches für die spätere Auswertung zusätzliche Informationen erfasst werden können. Am Ende einer Untersuchung kann ein Evaluationsbericht als PDF-Dokument und/oder CSV-Datei erzeugt werden, wobei alle getroffenen Einstellungen, Bewertungen und Kommentare mit übernommen werden.

Um auch Zwischenstände während einer Studie speichern zu können, wurde die Anwendung Ende 2012 um eine Projektverwaltungskomponente erweitert, mit der sich die Nutzung des Tools noch etwas komfortabler gestaltet. Dabei wird beispielsweise die Evaluation in Teams mit einer Gruppe von Gutachter/-innen unterstützt. Zudem können einmal zusammengestellte, individuelle Kriterienkataloge auch als Basis für zukünftige Evaluationsprojekte verwendet werden.³

² <http://www.cheval-lab.ch/usability-in-bibliotheken/leitfaden-bibeval/> [Stand: 04.12.2013].

³ Detaillierte Informationen zur Projektverwaltung von „BibEval“ sind unter <http://www.cheval-lab.ch/usability-in-bibliotheken/projektverwaltung-bibeval/> [Stand: 04.12.2013] verfügbar.

6. PECCI

Das Forschungsprojekt PECCI (Plateforme d'évaluation des centres d'information) hatte eine Online-Plattform zum Ziel, anhand derer Informationszentren die Usefulness ihrer Online-Angebote evaluieren können. Schritt für Schritt wird die Vorgehensweise einer Usefulness-Evaluation erklärt und somit jedem Laien ermöglicht, diese selbst durchzuführen. Die Plattform soll Interessierten durch den Prozess leiten und konkrete Methoden anbieten. Insbesondere ist damit die Auflistung von Fragen gemeint, welche nach Anpassung an den Kontext des jeweiligen Informationszentrums in einen Fragebogen integriert werden können (vgl. Birri Blezon/Hügi/Schneider 2011:67).

Basierend auf den von Tsakonas und Papa-theodorou (2006:402) entwickelten Attributen für die Beschreibung des Nützlichkeitsaspekts wurde eine allgemein verbindliche Vorgehensmethode erprobt. Bei zwei Fallstudien (für fonoteca.ch, die Schweizerische Nationalphonothek, und für infoclio.ch, das Fachportal für die Geschichtswissenschaften der Schweiz) wurden die Methoden des ethnographischen Interviews, der Logfile-Analyse und der Online-Umfrage eingesetzt, um die Nützlichkeit der Inhalte und Dienstleistungen, welche in beiden Fällen online angeboten wurden, zu evaluieren. Beide Portale besitzen die Eigenheit, dass ihnen der direkte Kontakt zu den Anwender/-innen fehlt, was wiederum zu fehlendem Feedback bezüglich der Benutzung und der Beurteilung geführt hat. Die angewendeten Methoden sollten darüber Aufschluss geben, welche Elemente als nützlich betrachtet werden, welche Aspekte ausgebaut und welche Komponenten abgeschafft werden können. Die Resultate der Studie bieten folglich eine Grundlage für die Entscheidungsfindung bezüglich zukünftiger Strategien und Ausrichtungen.

Eines der Hauptziele des Forschungsprojekts war es, die entwickelte Methode anderen so niederschwellig wie möglich anbieten zu können. Das Zielpublikum dieses Vorhabens sind vor allem Informationsvermittler/-innen im Bereich Bibliothek, Dokumentation und Archiv. Damit diese mit möglichst wenig Aufwand und mit geringem Wissen möglichst schnell zu aussagekräftigen Resultaten kommen, wurde entschieden, eine Online-Plattform für die Nützlichkeitsbewertung zu entwickeln. Diese Plattform hat einen Fragebogen als Kernstück, welcher Fragen auf mesoskopischer Ebene enthält. Die Fragen wurden den Umfragen der beiden Fallstudien entnommen und insoweit verallgemeinert, dass Benutzer/-innen der Plattform diese auf einfache Weise an ihren eigenen Kontext anpassen können. Für die technische Umsetzung dieser Platt-

form wurde auf die Open-Source Online-Umfrage-Applikation LimeSurvey zurückgegriffen. Zur Erhöhung der Benutzerfreundlichkeit wurde eine vereinfachte Schnittstelle für die Benutzung der Software erzeugt, da viele der angebotenen Funktionalitäten in diesem Kontext nicht gebraucht werden. Benutzer/-innen finden auf der Online-Plattform usefulness.ch Informationen zu den verschiedenen Methoden, welche für die Nützlichkeitsbewertung eines Online-Angebots eingesetzt werden können. Darüber hinaus haben sie Zugang zu einer Zusammenstellung vorgefertigter Fragen, die entsprechend der Bedürfnisse der durchzuführenden Evaluation ausgewählt, modifiziert und angereichert werden können. Die Umfrage kann direkt von der Plattform aus gestartet werden, wo die Antworten der Umfrageteilnehmer/-innen gesammelt werden. Dafür erhalten die Benutzer/-innen einen Link, welcher beispielsweise per E-Mail verschickt oder auf einer Website hinzugefügt werden kann. Alternativ gibt es auch die Möglichkeit, den Fragebogen zu exportieren, um ihn in einer offiziellen LimeSurvey-Umgebung zu nutzen. Die Resultate der gesammelten Antworten können in einer CSV-Datei heruntergeladen werden. Die Online-Plattform steht auf Deutsch und Französisch zur Verfügung und kann kostenlos genutzt werden.

7. Fazit und Ausblick

Bei der Evaluation bibliothekarischer Online-Angebote müssen verschiedene Aspekte berücksichtigt werden. Denn sowohl die zur Verfügung stehenden Funktionalitäten, deren Ausgestaltung sowie die generelle Gestaltung der Benutzeroberfläche als auch die angebotenen Inhalte sind für die von Anwender/-innen wahrgenommene Qualität von grosser Bedeutung.

Für die Evaluation der Usability kann hierbei auf ein umfangreiches und erprobtes Spektrum unterschiedlicher Methoden zurückgegriffen werden. Hinsichtlich der Evaluation der Usefulness ist dies leider nicht der Fall, vielmehr besteht hier noch Forschungsbedarf.

Ein Grund hierfür sind die besonderen Umstände im Bibliothekssektor. Denn hier stehen, insbesondere bei digitalen Bibliotheken, aus Sicht der Anwender die zur Verfügung gestellten Inhalte selbst im Fokus. In vielen anderen Branchen wie beispielsweise bei E-Commerce-Plattformen ist der Inhalt eines Webauftritts natürlich auch von Bedeutung – diese stellen jedoch in der Regel nicht das eigentlich gewünschte Gut dar, sondern sind nur Mittel zum Zweck, um die gewünschten Produkte zu finden. In diesem Fall können Usability und Usefulness besser gemeinsam betrachtet und evaluiert werden, da hier bei-

de Aspekte in einem ganzheitlichen Konzept zusammengefasst sind. Im Vordergrund stehen dabei die Funktionalitäten und deren Ausrichtung auf die Nutzerbedürfnisse.

Im Bibliothekssektor hingegen ist aus Sicht der Usefulness vor allem die Erfüllung der jeweiligen Informationsbedürfnisse von Bedeutung. Bezüglich der Bewertung liegen hier die gleichen Probleme vor, die aus dem Bereich des Information Retrieval hinsichtlich der Relevanzbewertung seit langem bekannt sind. Es stellt sich die Frage, wie der stark subjektiv geprägte Begriff der Relevanz objektiv und transparent bestimmt werden kann.

Wie in den vorangegangenen Kapiteln dargestellt, werden im Bereich digitaler Bibliotheken derzeit überwiegend Fragebögen zur Evaluation der Usefulness herangezogen. Hier muss kritisch hinterfragt werden, ob diese wirklich das einzige beziehungsweise das bestgeeignetste Mittel darstellen. Es bleibt daher spannend zu sehen, wie sich die in der Theorie vorhandenen Modelle zur Bestimmung der Nützlichkeits zukünftig in praktisch anwendbaren Methoden niederschlagen werden.

Literaturverzeichnis

- Battleon, Brenda / Booth, Austin / Weintrop, Jane (2001): Usability Testing of an Academic Library Web Site: A Case Study. In: *The Journal of Academic Librarianship* 27,3. S. 188-198. Online verfügbar unter [http://dx.doi.org/10.1016/S0099-1333\(01\)00180-X](http://dx.doi.org/10.1016/S0099-1333(01)00180-X).
- Birri Blezon, Rahel / Hügi, Jasmin / Schneider, René (2011): „Sieht gut aus, aber was bringt es mir?“ – Zur Evaluation der Nützlichkeit digitaler Inhalte. In: Bekavac, Bernard / Schneider, René / Schweibenz, Werner (Hg.): *Benutzerorientierte Bibliotheken im Web – Usability-Methoden, Umsetzung und Trends*. Berlin: De Gruyter. S. 55-73.
- Burmester, Michael (2003): Ist das wirklich gut? Bedeutung der Evaluation für die benutzerzentrierte Gestaltung. In: Machate, Joachim / Burmester, Michael (Hg.): *User Interface Tuning – Benutzungsschnittstellen menschlich gestalten*. Frankfurt: Software und Support. S. 97-119.
- Dumas, Joseph S. / Redish, Janice C. (1999): *A Practical Guide to Usability Testing*. 2. Auflage. Exeter: Intellect Books.
- Fuhr, Norbert et al. (2007): Evaluation of Digital Libraries. In: *International Journal On Digital Libraries* 8,1. S. 21-38. Online verfügbar unter <http://dx.doi.org/10.1007/s00799-007-0011-z>.
- Fuhr, Norbert et al. (2001): Digital Libraries: A Generic Classification and Evaluation Scheme. Research and Advanced Technology for Digital Libraries: 5th European Conference, ECDL 2001, Darmstadt, Deutschland, 4.-9. September, 2001. Berlin: Springer. S. 187-199. Online verfügbar unter <http://www.sics.se/~preben/papers/ecdl-2001.pdf> [Stand: 04.12.2013].
- Geis, Thomas (2010): Usability und User Experience unterscheiden. Online verfügbar unter <http://blog.procontext.com/2010/03/usability-und-user-experience-unterscheiden.html> [Stand: 10.07.2013].
- George, Carole A. (2008): *User-Centred Library Websites: Usability Evaluation Methods*. Oxford: Chandos Publishing.
- Heuer, Jens (2003): Expertevaluation. In: Heinsen, Sven / Vogt, Petra (Hg.): *Usability praktisch umsetzen*. München: Hanser Verlag. S. 116-135.
- Hörz, Manfred (2009): Utilitarismus. Online verfügbar unter <http://philmath.org/Utilitarismus.pdf> [Stand: 04.12.2013].
- International Organization for Standardization (2010): ISO 9241-210: Ergonomics of Human-System Interaction – Part 210: Human-Centered Design for Interactive Systems.
- International Organization for Standardization (2006): ISO 9241-110: Ergonomics of Human-System Interaction – Part 110: Dialogue Principles.
- International Organization for Standardization (1998): ISO 9241-11: Ergonomics of Human-System Interaction – Part 11: Guidance on Usability.
- Kani-Zabihi, Elahe / Ghinea, Gheorghita / Chen, Sherry Y. (2006): Digital Libraries: What do Users Want? In: *Online Information Review* 30,4. S. 395-412. Online verfügbar unter <http://dx.doi.org/10.1108/14684520610686292>.
- Nielsen, Jakob (2012): Usability 101: Introduction to Usability. Online verfügbar unter <http://www.nngroup.com/articles/usability-101-introduction-to-usability/> [Stand: 04.12.2013].
- Nielsen, Jakob (1994): Heuristic Evaluation. In: Nielsen, Jakob / Mack, Robert L. (Hg.): *Usability Inspection Methods*. New York: Wiley. S. 25-62.
- Nielsen, Jakob (1993): *Usability Engineering*. San Diego: Academic Press.
- Norlin, Elaina / Winters, CM! (2001): *Usability Testing for Library Web Sites: A Hands-On Guide*. Chicago: American Library Association.
- Sandusky, Robert J. (2002): Digital Library Attributes: Framing Usability Research. In: *Proceedings of the Workshop on Usability of Digital Libraries at the Second ACM/IEEE-CS Joint Conference on Digital Libraries, Portland, Oregon, USA, 14.-18. Juli 2002*, S. 35-38. Online verfügbar unter <http://www.ucl.ac.uk/annb/docs/Sandusky35.pdf> [Stand: 04.12.2013].
- Saracevic, Tefko / Covi, Lisa (2000): Challenges for Digital Library Evaluation. *Knowledge Innovations: Celebrating Our Heritage, Designing Our Future*. In: *Proceedings of the 63rd Annual Meeting of the American Society for Information Science, November 11-16, 2000, Chicago, IL*. S. 341-350.
- Sarodnick, Florian / Brau, Henning (2011): *Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung*. 2. Auflage. Bern: Hans Huber.
- Schweibenz, Werner (2011): Usability-Evaluation von Bibliothekswebsites. In: Bekavac, Bernard / Schneider, René / Schweibenz, Werner (Hg.): *Benutzerorientierte Bibliotheken im Web – Usability-Methoden, Umsetzung und Trends*. Berlin: De Gruyter. S. 9-30.

- Schweibenz, Werner / Thissen, Frank (2003): Qualität im Web – Benutzerfreundliche Webseiten durch Usability Evaluation. Berlin: Springer.
- Tsakonas, Giannis / Papatheodorou, Christos (2006): Analysing and Evaluating Usefulness and Usability in Electronic Information Services. In: Journal of Information Science 32,5. S. 400-419. Online verfügbar unter <http://dx.doi.org/10.1177/0165551506065934>.
- Weinhold, Thomas / Hamann, Sonja / Bekavac, Bernard (2011): Usability-Evaluation von Bibliothekswebsites. In: Bekavac, Bernard / Schneider, René / Schweibenz, Werner (Hg.): Benutzerorientierte Bibliotheken im Web – Usability-Methoden, Umsetzung und Trends. Berlin: De Gruyter. S. 31-53.
- Weinhold, Thomas / Öttl, Sonja / Bekavac, Bernard (2011): BibEval – Ein webbasierter Kriterienkatalog zur Usability-Evaluation von Bibliothekswebsites. In: IWP – Information Wissenschaft & Praxis 62,1. S. 11-18.
- Xie, Hong I. (2008): Users' Evaluation of Digital Libraries (DLs): Their Uses, their Criteria, and their Assessment. In: Information Processing and Management 44,3. S. 1346-1373. Online verfügbar unter <http://dx.doi.org/10.1016/j.ipm.2007.10.003>.
- Xie, Hong I. (2006): Evaluation of Digital Libraries: Criteria and Problems from Users' Perspectives. In: Library and Information Science Research 28,3. S. 433-452. Online verfügbar unter <http://dx.doi.org/10.1016/j.lisr.2006.06.002>.

Thomas Weinhold ist wissenschaftlicher Mitarbeiter am Schweizerischen Institut für Informationswissenschaft (SII), Hochschule für Technik und Wirtschaft (HTW) Chur, Ringstrasse/Pulvermühlestrasse 57, CH-7004 Chur, Tel.: ++41 81 286 24 26, E-Mail: thomas.weinhold@htwchur.ch.

Jasmin Hügi ist Forschungsassistentin am Studiengang in Informationswissenschaften, Haute école de gestion de Genève, Campus de Battelle, Route de Drize 7, CH-1227 Carouge, Tel. ++41 22 388 18 76, E-Mail: jasmin.hugi@hesge.ch.

Prof. Dr. René Schneider ist Koordinator des bilingualen Studienganges in Informationswissenschaften, Haute école de gestion de Genève, Campus de Battelle, Route de Drize 7, CH-1227 Carouge, Tel. ++41 22 388 18 66, E-Mail: rene.schneider@hesge.ch.

10

Prof. Dr. Bernard Bekavac ist Leiter des Bachelor Studienganges Information Science sowie des Forschungsschwerpunktes Information Engineering, Hochschule für Technik und Wirtschaft (HTW) Chur, Ringstrasse/Pulvermühlestrasse 57, CH-7004 Chur, Tel.: ++41 81 286 24 70, E-Mail: bernard.bekavac@htwchur.ch.