

UNIVERSITÄT SIEGEN

HUMAN COMPUTER INTERACTION

MASTERARBEIT

Akzeptanz als Erfolgsfaktor eines Empfehlungssystems

Autor:

Malte PASKUDA

Matrikelnr.: 970800

Betreuer:

Dr. Markus ROHDE

Martin STEIN

Erstprüfer:

Prof. Dr. Volker WULF

Zweitprüfer:

Prof. Dr. Volkmar PIPEK

24. Oktober 2013

Inhaltsverzeichnis

Abbildungsverzeichnis	4
Tabellenverzeichnis	5
1 Einleitung	6
1.1 Forschungsfrage	6
1.2 Erwartetes Ergebnis	7
1.3 Vorgehen	7
2 Verwandte Arbeiten	7
2.1 Technikakzeptanz	7
2.2 Empfehlungssysteme	11
2.2.1 Precision und Recall	12
2.2.2 Algorithmen	12
2.3 Akzeptanz von Empfehlungssystemen	16
2.4 Datenschutz und Privacy	23
3 Verwandte Systeme	30
3.1 Nutzung von Empfehlungen in Websystemen	30
3.2 Stand von Mitfahrzentralen	33
3.2.1 Kategorisierung von Mitfahrzentralen	34
4 Anforderungen an das Empfehlungssystem	37
4.1 Anforderungen durch Akzeptanzfaktoren	37
4.2 Anforderungen durch das Projekt	38
4.3 Anforderungen durch die Nutzer	39
5 Beschreibung des Empfehlungssystems	45
5.1 Diskussion der Anforderungen	48
6 Beschreibung der Evaluation	50
6.1 Workshopgruppenplan	50
6.1.1 Das Ziel des Workshops	50
6.1.2 Formales	51
6.1.3 Vorbereitung	51
6.1.4 Ablauf	51
6.2 Fragebogen	52
6.3 Ablauf	53
7 Auswertung	53
7.1 Workshop	53
7.1.1 Gruppe 1	53
7.1.2 Gruppe 2	60

7.1.3	Fazit der Workshops	63
7.2	Fragebögen	64
7.2.1	Verhältnis zwischen Workshop und Fragebogen	65
7.2.2	Vergleich mit der Bewertung der Gesamtplattform	66
7.2.3	Vergleich mit den Interviews	68
8	Fazit	69
	Literatur	71

Abbildungsverzeichnis

1	Die Komponenten der Theory of Reasoned Action, übernommen aus Hale et al. [2002].	8
2	Modell TAM2, übernommen aus Venkatesh and Davis [2000].	9
3	Modell TAM3, übernommen aus Venkatesh and Bala [2008].	10
4	Ein Beispiel einer Folksonomie, übernommen aus Bender [2011].	12
5	Beispiel für den k-nearest-neighbour-Algorithmus. Gilt der innere Kreis, würde das Fragezeichen ein rotes Dreieck. Gilt der äußere, würde es ein blaues Quadrat. Übernommen von https://en.wikipedia.org/wiki/File:KnnClassification.svg , Abrufdatum 24.09.2013	13
6	Korrelation zwischen Akzeptanz und PU und PEOU, übernommen aus Hu and Pu [2009a].	16
7	Empfehlung und ihre Erklärung bei der Suche nach Spas, aus Zanker [2012]	18
8	Das Modell hinter ResQue, übernommen aus Pu et al. [2011].	20
9	Das getestete strukturierte Empfehlungssystem, übernommen aus Chen and Pu [2008].	21
10	Das getestete listenartige Empfehlungssystem, übernommen aus Chen and Pu [2008].	22
11	Designmethode angelehnt an juristische Abwägungsprozesse, übernommen aus Iachello and Abowd [2005].	29
12	Beispiel für von Facebook genutzte Empfehlungsarten.	30
13	Beispiel für von Amazon genutzte Empfehlungen.	32
14	Von Flinc angezeigte vertrauensbildende Informationen: Bewertungen, gemeinsame Bekannte, Identifizierung der Plattform gegenüber.	35
15	Das Empfehlungssystem ist unten rechts zu sehen, links füllt die normale Veranstaltungsliste die Seite.	49
16	Von einem Teilnehmer sortierte Profilkarteikarten.	58
17	Als Beispiel: Graph des Zusammenhangs zwischen Zustimmung zur These, dass Technik bedrohlich ist, und der Bewertung des Empfehlungssystems.	68

Tabellenverzeichnis

1	Beispiel einer Originaltabelle, übernommen aus Li et al. [2007]	24
2	Beispiel einer problematischen Tabelle trotz einer l-diversity von 3, übernommen aus Li et al. [2007]	25
3	Bearbeitete Tabelle mit minimierter t-closeness, übernommen aus Li et al. [2007]	25
4	Ergebnis des Fragebogens beider Gruppen. Die mit einem * markierten Teilnehmer nahmen am zweiten Workshop teil. Die Likert-Skala reichte von 1-5. Siehe Abschnitt 6.2.	65
5	Durchschnittliche Gesamtbewertung.	65
6	Ergebnis des Fragebogens der Teilnehmer in der vorherigen Studie von Beil, siehe Beil [2013]. Die mit einem * markierten Teilnehmer nahmen am zweiten Workshop teil. Die Likert-Skala reichte von 1 bis 5.	66
7	Durchschnittliche Gesamtbewertung in der vorherigen Studie, siehe Beil [2013].	67

Ein Empfehlungssystem wurde entworfen, das in eine Mobilitätsplattform eines Mobilitätsprojektes (S-Mobil 100) für ältere Mitbürger integriert wurde. Anhand von Ex-Ante und Ex-Post Interviews bzw. einer Evaluation, die zu Beginn und nach Abschluss des Projektes durchgeführt wurden, konnte untersucht werden, ob das Empfehlungssystem akzeptiert wird, welche Faktoren dazu beitragen und ob dies anhand der Interviews bereits zu Projektbeginn hätte festgestellt werden können.

1 Einleitung

Empfehlungssysteme werden auf vielen großen Plattformen zu unterschiedlichen Zwecken eingesetzt. Auf Seiten wie Amazon empfehlen sie Produkte an Kunden und unterstützen den Verkauf (Hennig-Thurau et al. [2010]). Auf Seiten wie Facebook empfehlen sie Freunde und Aktivitäten und dienen daher dem weiteren Engagement auf der Seite (Geyer et al. [2008]). Es sind Systeme aus dem Bereich der Graphenalgorithmien und Big Data, die inzwischen seit Jahren erforscht und entwickelt werden, zu denen Daten zu ihrer Effektivität erhoben werden (Linden et al. [2003]) und die umfangreich evaluiert werden können (Bender [2011]).

S-Mobil 100 ist ein Projekt, das ebenfalls ein Empfehlungssystem benutzen könnte. Unter anderem für die integrierte Mitfahrzentrale bietet es sich an, inserierte Fahrten mindestens an dem Nutzer bekannte Mitglieder zu empfehlen. Die Nutzerschaft des Portals ist jedoch eine spezielle Zielgruppe: Es sind ältere Bürger, die in der Umgebung von Siegen teils schlecht angebunden an das öffentliche Verkehrsnetz leben. Schon generell sind wenige Daten zur Technikakzeptanz von Empfehlungssystemen vorhanden (Hu and Pu [2009a]), noch weniger wurde die Akzeptanz durch diese spezielle Zielgruppe untersucht.

Die Teilnehmer des Projektes dagegen wurden umfassend untersucht. Von der Uni Heidelberg wurde eine Studie [Beil, 2013] zur Technikakzeptanz im Gebiet Siegen durchgeführt, inklusive der Projektteilnehmer. Im Ergebnis wurde deutlich, dass die Teilnehmer überdurchschnittlich technikaffin und technikerfahren sind. Die Studie liefert weiter Hintergrundinformationen, mit denen bei der Bewertung der Reaktion der Teilnehmer auf die Empfehlungen der jeweilige Hintergrund der Teilnehmer mit einbezogen werden kann.

Eine Besonderheit im Vergleich zu normalen kollaborativen Empfehlungssystemen ist das offene Freundschaftsmodell im Projekt. Nutzer tragen keine Freundschaften ein, sondern eine etwaige Nähe zwischen den Mitgliedern muss aus den bisherigen Aktivitäten - beispielsweise mehreren gemeinsamen Fahrten oder gemeinsamen Interessen - abgeleitet werden.

1.1 Forschungsfrage

Gibt es vorher feststellbare Faktoren, welche die Akzeptanz eines Empfehlungssystems beeinflussen?

1.2 Erwartetes Ergebnis

Eine Reihe von Faktoren, die mit der Akzeptanz des Empfehlungssystems korrelieren, aber schon in einem generischen Interview vor Beginn eines solchen Projekts erfragbar sind. Ergebnis könnte aber auch sein, dass wenig technikaffine ältere Menschen einem Empfehlungssystem grundsätzlich misstrauen.

1.3 Vorgehen

Geplant war, dem Schema einer Design Case Study (Wulf et al. [2011]) zu folgen. Demnach wurde

1. die vorliegenden Interviews analysiert
2. mit den Erkenntnissen aus der Analyse und der Literatur ein Empfehlungssystem implementiert
3. abschließend die Akzeptanz des Empfehlungssystems evaluiert

Im Rahmen dieser Arbeit sollte also ein solches Empfehlungssystem entworfen und implementiert werden, das dem einzelnen Nutzer passende Fahrten anderer Mitglieder des Systems empfiehlt. Daraufhin sollte untersucht werden, wie dieses System von den Mitgliedern angenommen wird - z.B. durch TAM.

2 Verwandte Arbeiten

Es gibt eine Vielzahl von Arbeiten, die sich mit Empfehlungssystemen als solche beschäftigen, beispielsweise mit den verwendeten Algorithmen. Zusammenfassend werden Ansätze davon in Abschnitt 2.2 erklärt. Im Folgenden wird der Stand der wissenschaftlichen Forschung in relevanten Themenbereichen dargestellt.

2.1 Technikakzeptanz

Mit Technikakzeptanz können zwei verschiedene Konzepte bezeichnet werden. Eine Interpretation geht in Richtung Technikangst: Ist ein Individuum generell bereit, technische Apparaturen zu akzeptieren? Diese Sichtweise ist in dieser Arbeit jedoch nicht gemeint. Stattdessen geht es um die Akzeptanz eines bestimmten technischen Systems, ob es vom Nutzer akzeptiert wird, was bedeutet, dass er mindestens an dessen Existenz keinen Anstoß findet; besser noch, dass er es in der Zukunft selbst benutzen wird.

Das Technikakzeptanzmodell *TAM* ist ein Modell, um angesichts der Nutzerwahrnehmung der Einfachheit (PEOU – Perceived Ease Of Use) und der Nützlichkeit (PU – Perceived Usefulness) des Systems grob vorherzusagen, ob es von dem Nutzer verwendet werden wird (Davis [1989]). Es wird in dieser Arbeit genutzt werden, um - neben der qualitativen Evaluation - per Fragebogen die Akzeptanz des Systems zu bestimmen.

TAM basiert auf der Theorie des überlegten Handelns (TRA) (Benbasat and Barki [2007]), einer sozialpsychologischen Theorie. Nicht nur trennt diese Theorie die Intention

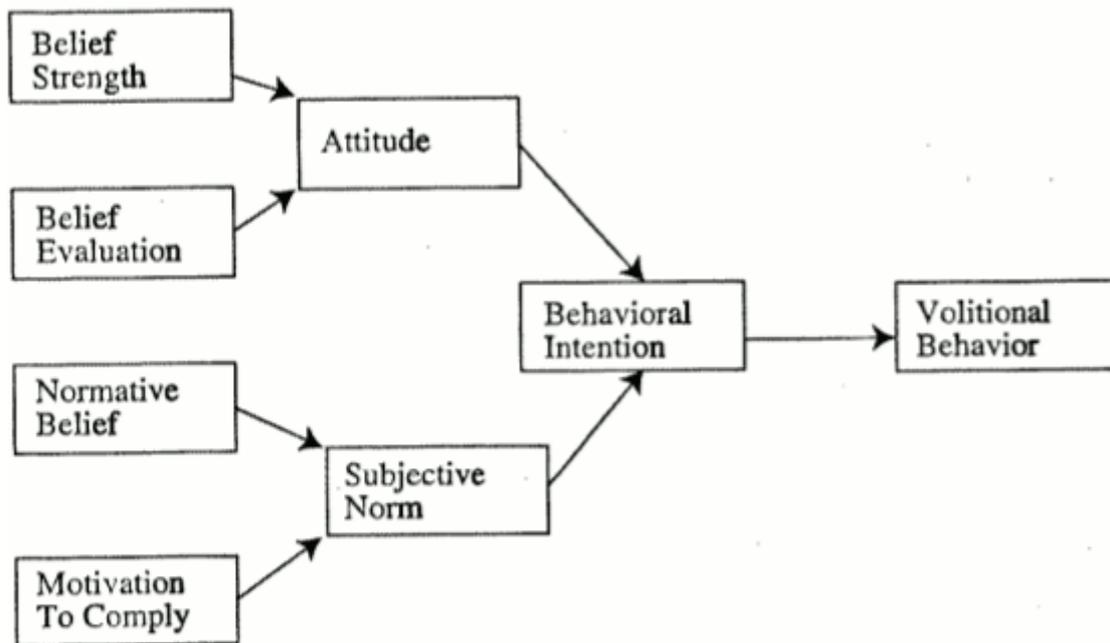


Abbildung 1: Die Komponenten der Theory of Reasoned Action, übernommen aus Hale et al. [2002].

zur Handlung vom eigentlichen Handeln in einzelne Schritte, sondern sie argumentiert, dass die Intention einer Person, etwas bestimmtes zu tun, davon abhängt, wie die Person diese Handlung wahrnimmt und wie sie vermutet, dass sie selbst nach dieser Handlung wahrgenommen werden würde (Fishbein and Ajzen [1975], siehe auch Abbildung 1). Das kann runtergebrochen werden auf die einfache Formel $BI = A_b W_1 + (SN)W_2$ wobei BI die Handlungsintention, A_b die Einstellung zur Handlung, SN die subjektive Norm und W_n empirisch gewonnene Gewichtungen sind (Hale et al. [2002]).

TAM ist nicht ohne Kritik. So wurde kritisiert, dass (Benbasat and Barki [2007]):

1. die Grundlage des Modells nicht ausreichend hinterfragt werden würden, IT-Artefakt-Design und ihre Evaluation, und die Konsequenzen der IT-Adaption unter dem Modell nicht ausreichend untersucht werden würden.
2. TAM-basierte Forschung nur zu einer Illusion der Wissensansammlung führe
3. angesichts des fortwährenden Wandel der tatsächlichen Adaption von IT-Systemen die Probleme bei der Erweiterung des Kernmodells von TAM dazu führten, dass es im Laufe der Zeit immer weniger geeignet würde
4. die Versuche, TAM durch Erweiterungen und Anpassungen zu korrigieren, nicht ausreichend wissenschaftlich fundiert seien

Der Kritik zum Trotz gilt TAM als erfolgreichstes Modell im Bereich der IT-Systeme (Lee et al.). Tatsächlich aber wurden einige Versuche unternommen, TAM zu erweitern, die alle den Referenzfragebogen erweitern.

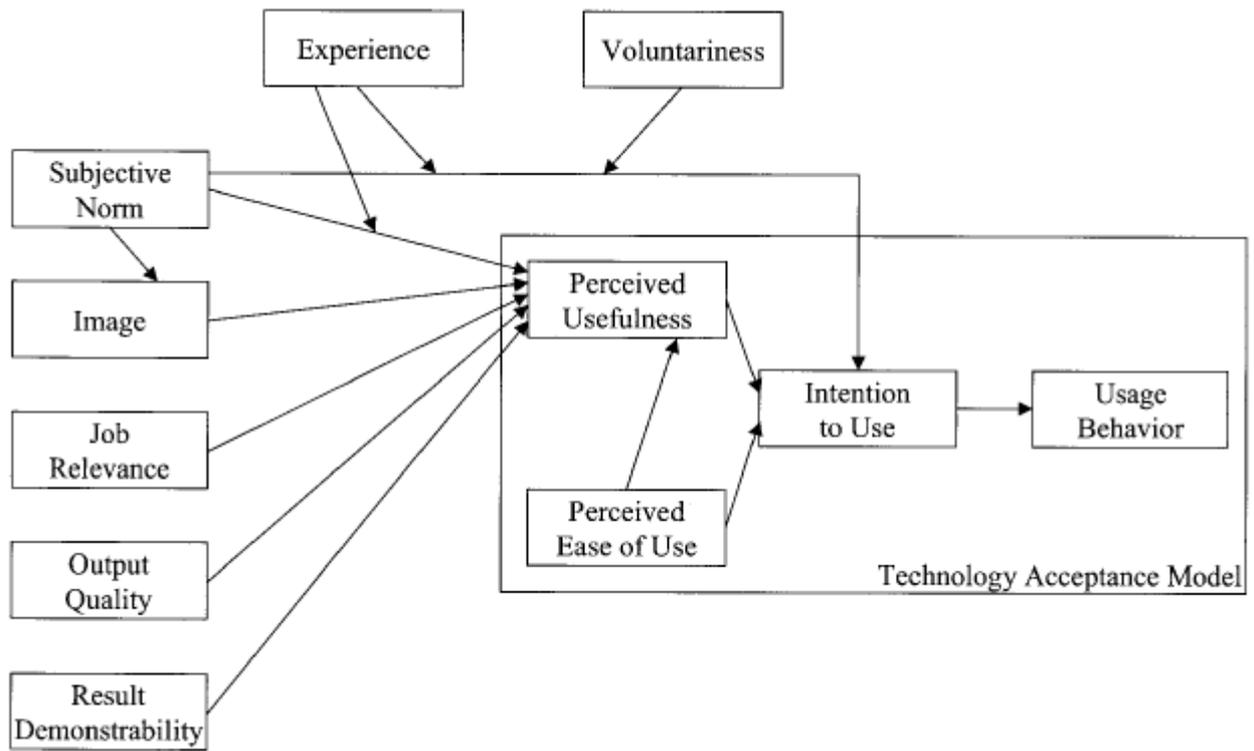


Abbildung 2: Modell TAM2, übernommen aus Venkatesh and Davis [2000].

TAM2 war ein solcher Versuch. Es wurde 2000 in einer Studie von vier Systemen in vier Organisationen vorgestellt. *TAM2* erweitert *TAM* und damit den Fragebogen um Faktoren zum sozialen Einfluss und der erweiterten Funktionalität, also um Faktoren wie der wahrgenommenen sozialen Erwartungshaltung des Umfelds oder der Qualität des Ergebnisses (Venkatesh and Davis [2000], siehe auch Abbildung 2).

Die *Unified theory of acceptance and use of technology* (UTAUT) war der nächste. Venkatesh et al. verglich acht verschiedene Modelle mit ihren Faktoren, inklusive *TAM*, um darauf aufbauend ein kombiniertes zu entwickeln. Es benennt vier Faktoren: Performance Expectancy, Effort Expectancy, Social Influence und Facilitating Conditions, die wiederum von vier Modifikatoren, Alter, Geschlecht, Erfahrung und Freiwilligkeit, beeinflusst würden. (Venkatesh et al. [2003]).

Schließlich wurde 2008 *TAM3* vorgestellt. *TAM3* hat als Modell wie *TAM* die wahrgenommene Nützlichkeit und Einfachheit im Fokus, benennt aber eine Reihe von Faktoren, die diese determinieren und andere, die diese Determinierung beeinflussen würden. Diese Modifikatoren sind Erfahrung und Freiwilligkeit. Faktoren für PE sind Image, Job

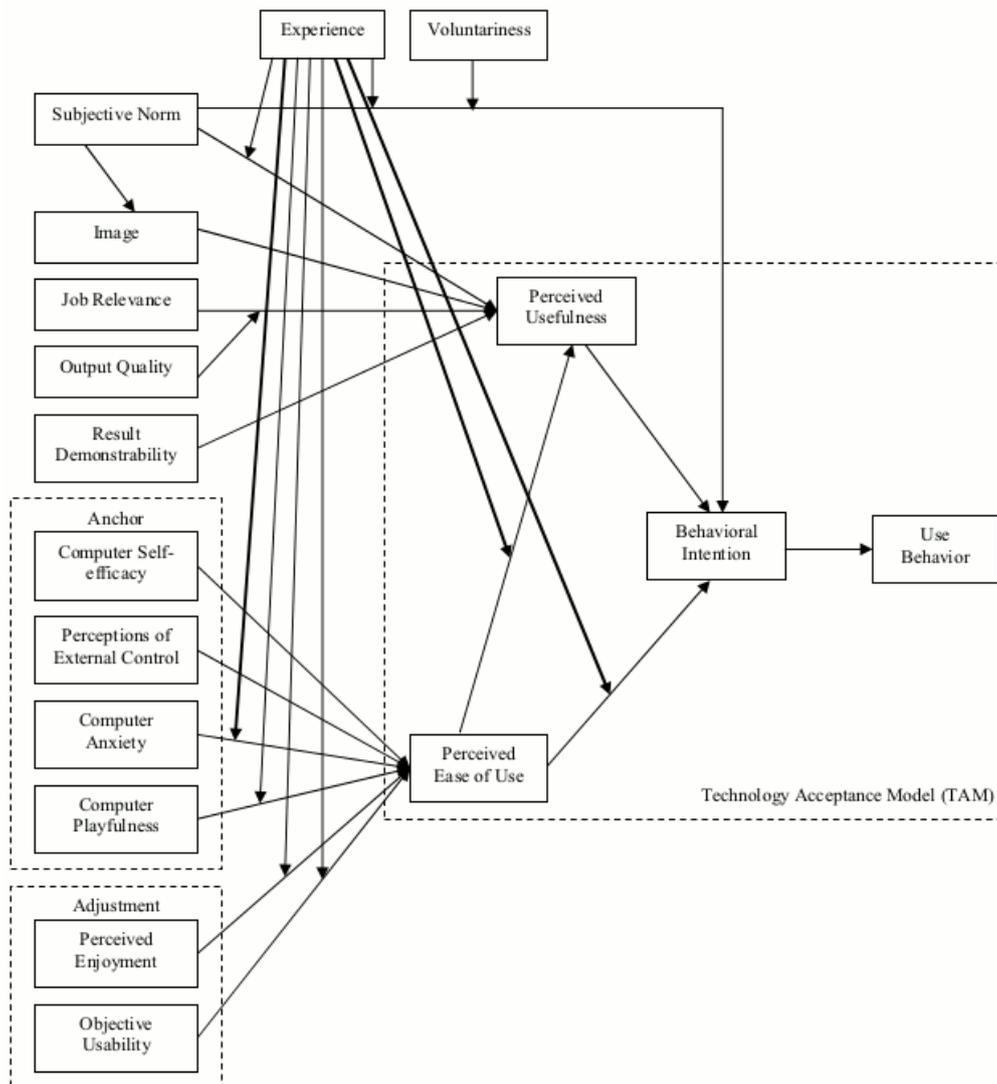


Abbildung 3: Modell TAM3, übernommen aus Venkatesh and Bala [2008].

Relevance, Output Quality und Result Demonstrability, Faktoren für PEOU Computer Self-efficacy, Perceptions of External Control, Computer Anxiety, Computer Playfulness, Perceived Enjoyment und Objective Usability. Dazu kommt der Faktor Subjective Norm, definiert als Grad der Wahrnehmung, inwiefern das geschätzte soziale Umfeld denkt, dass das System genutzt werden sollte, das nicht nur indirekt über PU, sondern als einziger Faktor auch direkt das beabsichtigte Verhalten beeinflusse (Venkatesh and Bala [2008], siehe auch Abbildung 3).

Diese Technikakzeptanzmodelle werden desöfteren angepasst und erweitert. So wurde in [Pu et al., 2011] ein Modell speziell für die Messung der Akzeptanz von Empfeh-

lungssystemen vorgestellt. Und [Porter and Donthu, 2006] stellte fest, dass Faktoren wie Alter, Herkunft, Bildung und Einkommen für die Akzeptanz des Internets relevant sind und erweiterten TAM entsprechend, und führten zusätzlich neben PU und PEOU die Variable AB (perceived Access Barriers) für die wahrgenommenen Nutzungsbarrieren ein.

Natürlich gibt es auch andere Methoden als TAM und darauf basierende. Definiert man Technikakzeptanz wie oben relativ zum technischen Artefakt, braucht eine Bestimmung der Technikakzeptanz eine Bestimmung der Nutzereinschätzung dieses Artefakts, also sowohl ihrer Funktionalität, Usability und User Experience. Eine solche Bestimmung leistet der Fragebogen Attrakdiff. Dieser versucht, gleichzeitig die pragmatischen und hedonistischen sogenannten Qualitäten eines Artefakts zu bestimmen. Er bestimmt dafür die Bewertung des Produkts in drei Bereichen mit: der pragmatischen, der hedonistischen im Bereich Stimulation und der hedonistischen im Bereich Identität (Hassenzahl et al. [2003]). Dabei sind es dann auch Faktoren wie die Schönheit des Artefakts, welche im Rahmen von *User Experience* die Akzeptanz beeinflussen, möglicherweise ohne Änderung der Bewertung von Nützlichkeit und Usability (Hassenzahl [2008]).

Zuätzlich geht Technikakzeptanz weiter als die Frage, ob ein System als Ganzes genutzt wird oder nicht. So kann Vertrauen in die korrekte Funktionsweise, gekoppelt mit der Wahrnehmung, das System ausreichend verstanden zu haben, beeinflussen, ob kritische Teile des Systems genutzt werden, selbst wenn das Gesamtsystem durchaus genutzt wird (Cramer et al. [2009]).

2.2 Empfehlungssysteme

Ein Empfehlungssystem im allgemeinen Sinne zeigt jemanden aufgrund irgendwelcher Zusammenhänge aus einer Datenmenge eine bestimmte Submenge. Besonders eignen sich dafür alle Datenmengen, die als Graphen abbildbar sind, was bei jeder Datenmenge funktionieren sollte, aus der man üblicherweise Empfehlungen gewinnen will.

Eine gern gewählte Datenstruktur ist die von Vander Wal so benannte Folksonomie (Wal [2007]). Eine Folksonomie ist eine Datenstruktur, in der Nutzer Ressourcen mit Tags versehen, wobei Tags Ressourcen mehrfach von verschiedenen Nutzern zugewiesen werden können (siehe Abbildung 4). Auf dieser Datenstruktur können entsprechend gestaltete Algorithmen ohne weitere Anpassungen genutzt werden.

Die formale Definition (übernommen von Bender [2011]) beschreibt Folksonomien eindeutig: Eine Folksonomie F ist definiert als Tupel

$$F := (U, T, R, Y) \quad (1)$$

wobei U die Menge der Nutzer, T die Menge der Tags, R die Menge der Ressourcen und Y die Relation ist, die Tags von Nutzern an Ressourcen zuweist, also

$$Y \subseteq U \times T \times R \quad (2)$$

Eine Folksonomie kann als ungerichteter Graph betrachtet werden. Die genutzte Datengrundlage für das hier benutzte Empfehlungssystem kann als reduzierte beziehungsweise umgedrehte Folksonomie betrachtet werden ¹.

¹siehe Abschnitt 5

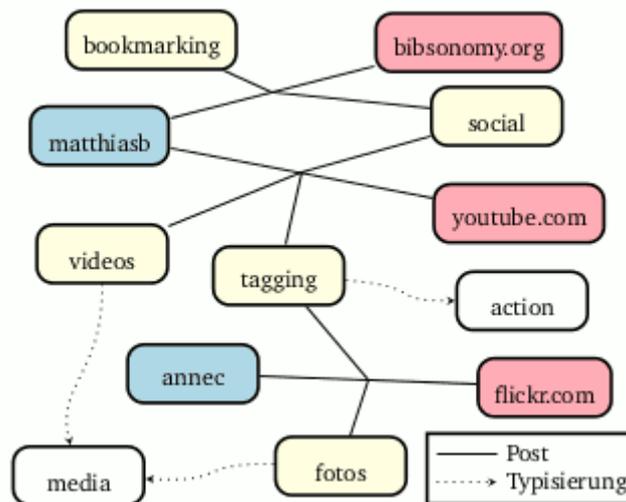


Abbildung 4: Ein Beispiel einer Folksonomie, übernommen aus Bender [2011].

2.2.1 Precision und Recall

Um verschiedene Empfehlungssysteme zu vergleichen oder auch nur um eines zu bewerten braucht es eine Qualitätsmetrik. Die Evaluation verschiedener Algorithmen ist ein eigenes Thema, weil hier auch die Beispieldaten gleich sein müssen. Unabhängig davon existieren die Qualitätsmetriken Precision und Recall. Precision beschreibt die Wahrscheinlichkeit, dass die erhaltenen Ergebnisse relevant sind, und Recall die Wahrscheinlichkeit, dass die relevanten Teil der erhaltenen Ergebnisse sind. Oder anders: Die Precision ist dann gut, wenn wirklich nur relevante Ergebnisse gezeigt werden, und Recall ist dann gut, wenn auch wirklich alle relevanten Ergebnisse enthalten sind (Han and Kamber [2006]).

Ein Empfehlungssystem, das einfach alle möglichen Optionen auflistet, hätte also einen hohen Recall, aber eine schlechte Precision. Daher müssen beide Werte zusammen betrachtet werden, wofür der F-Score genutzt wird (Han and Kamber [2006]).

Bei einem realen System ist vorab nicht bekannt, welche Ergebnisse für den Nutzer relevant sind, weshalb ein iterativer Entwicklungsprozess mit Nutzerbeteiligung sinnvoll erscheint. Die vorherige technische Evaluation kann nur eine grobe Schätzung der späteren Nutzerzufriedenheit erreichen, indem geprüft wird, ob die Empfehlungen in den Beispielszenarien sinnvoll erscheinen. Denn die Metriken sind nur begrenzt dazu geeignet Empfehlungssysteme zu bewerten, die Dinge empfehlen sollen, die dem Nutzer bisher unbekannt sind (Konstan and Riedl [2012]).

2.2.2 Algorithmen

Die Komplexität des genutzten Algorithmus ist bei Empfehlungssystemen besonders wichtig, weil die typischen Datenmengen sehr groß sein können. Nicht umsonst wird das

Themenfeld auch als *Big Data* bezeichnet.

In diesem speziellen Fall war die Komplexität unproblematisch, weil die Nutzermenge der Plattform auf $n=19$ (plus einigen wenigen Entwickleraccounts) begrenzt war. Trotzdem sollen im folgenden zwei generelle Ansätze beschrieben werden, zum einen weil die Grundidee des k -nearest-neighbours-Algorithmus nah an der umgesetzten Implementation ist, zum anderen, weil der FolkRank-Algorithmus eine mögliche Weiterentwicklung bei steigender Nutzerzahl sein könnte.

k-nearest neighbour

Die Nutzer des Systems stelle sich als Knoten eines Graphen vor, die auf einer Fläche verteilt sind. Ihre Entfernung zueinander wird von einer Entfernungsfunktion bestimmt, welche genau die Eigenschaften enthält, die für die Empfehlung genutzt werden sollen.

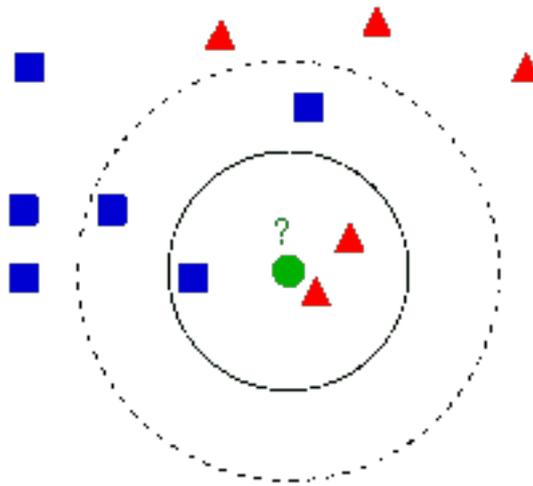


Abbildung 5: Beispiel für den k -nearest-neighbour-Algorithmus. Gilt der innere Kreis, würde das Fragezeichen ein rotes Dreieck. Gilt der äußere, würde es ein blaues Quadrat. Übernommen von <https://en.wikipedia.org/wiki/File:KnnClassification.svg>, Abrufdatum 24.09.2013

Nun werden die Knoten in Ähnlichkeitsgruppen eingeteilt, wobei immer Mitglieder der Gruppe einander empfohlen werden sollen. Für den bestimmten Node würde der k -nearest-Algorithmus dafür die k Nachbarn mit der geringsten Distanz betrachten und dem Node die Gruppe zuweisen, welche die meisten dieser benachbarten Nodes haben (Cover and Hart [1967], siehe auch Abbildung 5)

Allerdings wurde schon 1999 gezeigt, dass Nearest-Neighbour-Aussagen bei multi-dimensionalen Distanzen, als die man Ähnlichkeitsbeziehungen in Sozialen Netzwerken auch betrachten kann, schnell aussagelos werden (Beyer et al. [1999]).

FolkRank

Folkrank ist ein Kategorisierungs- und Suchalgorithmus für Folksonomien, eine Anpassung des Pagerank-Algorithmus. Die Idee ist, dass eine Ressource, die von wichtigen Nutzern mit wichtigen Tags versehen wurde, selbst wichtig wird.

Hierfür wird die Folksonomie in einen ungerichteten gewichteten Graph umgewandelt, wobei die Knoten die Ressourcen, Tags und Nutzer abbilden und die Kanten das gemeinsame Vorkommen von Nutzern mit Tags, Nutzern mit Ressourcen und Tags mit Ressourcen. Dieser Graph kann dann mit dem *Random Surfer Model* gefolgt und so den Knoten ein Pagerank gegeben werden. Für Folkrank dagegen wird nun statt diesem Modell ein angepasster Vorliebenvektor \vec{p} gewählt, der dem Themenwunsch des gerade gewählten Suchers entspricht (Hotho et al. [2006]).

Interessant sind Berichte basierend aus dem Betrieb eines großen Systems. So wurde die technische Grundlage von Amazons Empfehlungssystem ausführlich vorgestellt (Linden et al. [2003]). Dort wird beschrieben, dass Amazon nicht auf die (damals) üblichen Algorithmen vertraute, sondern ein eigenes System entwickelte. Anstatt Nutzer mit Nutzern zu vergleichen, werden für alle gekauften oder bewerteten Items eines Nutzers ähnliche Items gewählt, und aus diesen ähnlichen Items eine Empfehlungsliste gebildet. Ähnlichkeit von Items wird dadurch bestimmt, dass solche als ähnlich deklariert werden, die zusammen gekauft wurden. Während dieser Schritt aufwändig ist, ist die dadurch mögliche Empfehlung von Items nur abhängig von den bisher gekauften Items eines Nutzers und daher entsprechend schnell.

Das Beispiel von Amazon deutet es an: Es gibt mehr als eine Möglichkeit, Empfehlungssysteme aufzubauen. Dementsprechend findet sich eine relativ genormte englischsprachige Klassifizierung in der Literatur (Pu et al. [2012]):

Rating-based

Bei bewertungsbasierten Empfehlungssystemen bewerten Nutzer Ressourcen, wobei diese Bewertungen Teil ihres Profils werden, anhand dessen sich dann weitere Empfehlungen finden lassen. Solche Systeme kann man noch einmal unterteilen: Als *content-based* bezeichnete Systeme empfehlen Ressourcen, die den favorisierten ähnlich sind, während Systeme mit *collaborative filtering* für diese Empfehlung die Favoriten anderer Nutzer mit einer ähnlicher (aber eben nicht identischen) Bewertungshistorie benutzen. Bei letzteren Systemen entfällt der Schritt, eine Ähnlichkeitsbeziehung zwischen den Items aufgrund von wie auch immer gearteten objektiven Kriterien bestimmen zu müssen.

Feature-based

Empfehlungssysteme, die statt mit allen früheren Bewertungen ein allgemeines Profil zu erstellen, versuchen herauszufinden, welche speziellen Features ein Nutzer braucht. Dies sei besonders gut für Produkte, die nicht oft gekauft werden, und muss wiederum unterschieden werden:

1. **Case-based**

Case-based Empfehlungssysteme betrachten das Problem als ein Problemfall. Sie versuchen, ein früheres ähnliches Problem samt Lösung zu finden, diese zu verallgemeinern und auf das momentane Problem anzuwenden. Dies entspricht frühen Algorithmen und Denkweisen der schwachen KI.

2. Utility-based

Hier wird versucht, Bedürfnisse des Nutzers zu bestimmen und die möglichen Empfehlungen daraus abzuleiten, wodurch diese Bedürfnisse am besten bedient werden. Das führt zu multidimensionalen Problemen, die durch eine Bewertungsfunktion abgebildet werden müssen.

3. Knowledge-based

Der Empfehler beachtet funktionales Wissen im Hinblick auf das Bedürfnis der Nutzer. Weiß er, dass ein bestimmtes Item etwas produziert, was ein bestimmtes Bedürfnis eines Nutzers abdeckt, kann dieses Item empfohlen werden.

4. Critiquing-based

In einem virtuellen Verkaufsgespräch werden Items empfohlen. Der Nutzer hat dann in einem (möglicherweise zyklischen) Prozess die Möglichkeit, bestimmte Features des Items zu kritisieren, beispielsweise ein billigeres und schnelleres Produkt zu fordern, wodurch dann die Empfehlung und das Profil des Nutzers angepasst werden kann. Ein solches System wurde z.B. in [Pu et al., 2009] benutzt, um Empfehlungen für Parfums auszusprechen, wobei diese *editorial picked critiques* durch einer Kombination von ausgewählten Magazinkritiken, Beliebtheitsinformationen und mit durch Data Mining gewonnene Faktoren erstellt wurden.

Personality-based

Bei einem solchen System wird versucht, die Persönlichkeit des Nutzers zu bestimmen. Dafür gibt es mehrere Methoden, von stiller Beobachtung bis explizitem Fragebogen. Empfehlungen können dann durch das Wissen über diese Persönlichkeit ausgesprochen werden - eventuell wieder über Parallelen mit den Vorlieben anderer Nutzern mit ähnlichem Profil, wie bei collaborative filtering.

Problematisch ist die Evaluation von solchen Empfehlungssystemen, selbst rein auf der technischen Seite und unter Auslassung von Faktoren wie Usability und Nutzerakzeptanz. Da ist zum einen die Komplexität des Algorithmus selbst und damit die benötigte Rechenzeit, aber auch der benötigte Speicherplatz. Dazu kommt aber auch die Qualität der Empfehlungen selbst. Daher benötigt die Evaluation von Empfehlungssystemen eigentlich nicht nur den Vergleich verschiedener Algorithmen des Forschers im jeweiligen Testfall, sondern eine genormte Datenmenge bzw ein Evaluationsframework (Bender [2011]).

2.3 Akzeptanz von Empfehlungssystemen

Auch die Technikakzeptanz von Empfehlungssystemen selbst ist bereits Thema gewesen. So wurde untersucht, welche Eigenschaften der Software die Akzeptanz steigern. Denn es reicht nicht, nur auf die Genauigkeit der Empfehlung zu achten, sondern es muss auch beachtet werden, welche Anforderungen der Nutzer an das Empfehlungssystem hat, der das System beispielsweise auch zur Entdeckung entfernter verwandter Artikel nutzen können wollte (Shani and Gunawardana [2011]). Diese Erkenntnis - die Subjektivität der Qualität eines Empfehlungssystems - wird noch verstärkt durch die Erkenntnis, dass der Wunsch von Nutzern nach Diversität von Empfehlungsergebnissen von Persönlichkeitsfaktoren der Nutzer abhängt (Chen et al. [2013]).

Ein Ergebnis dieser Kenntnis kann sein, nutzerorientierte Messfaktoren wie die Stabilität eines Algorithmus einzuführen - also ob eine Empfehlungssystem die gleichen Dinge empfiehlt, auch nach einem längeren Zeitraum, wenn keine neuen Erkenntnisse vorliegen, weil ein Fehlen dieser Stabilität für den Nutzer verwirrend sei - um die Nutzerwahrnehmung direkt im Kern des System zu beachten (Adomavicius and Zhang [2012]).

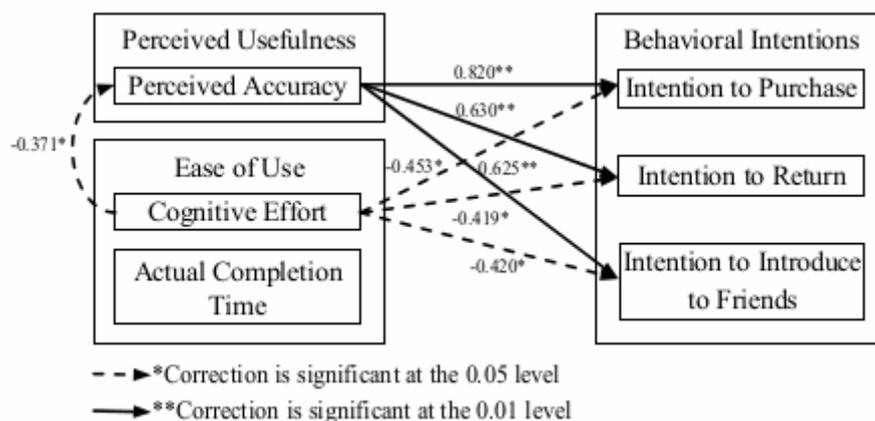


Abbildung 6: Korrelation zwischen Akzeptanz und PU und PEOU, übernommen aus Hu and Pu [2009a].

In [Hu and Pu, 2009a] dagegen wurden zwei Empfehlungssysteme verglichen, wobei eines davon ein System war, das Empfehlung aufgrund mittels eines Fragebogens ermittelter Charakterähnlichkeit zwischen Nutzern aussprach (PBR), und das andere ein klassisches bewertungsbasiertes Empfehlungssystem war. Die Akzeptanz der Systeme wurde mittels TAM bestimmt. Aus ihren Ergebnissen (siehe Abbildung 6) leiteten sie drei Richtlinien ab (Hu and Pu [2009a]):

1. Versuche, die wahrgenommene Genauigkeit zu verbessern, um Nutzer zum Kauf zu motivieren
2. Gestalte unterhaltsame Persönlichkeits-Fragebögen (für PBR)
3. Erhöhe Transparenz und Nutzerkontrolle

Über die Relevanz dieser Richtlinien kann man streiten, sind sie doch eher Allgemeinplätze, die grundsätzlich für jedes Softwaredesign gelten sollten. Doch kann man immerhin versuchen, aus der Studie die Wichtigkeit von Transparenz bei der Darstellung der Empfehlungen abzuleiten, was nicht unbedingt selbstverständlich ist und von bestehenden Systemen nicht immer befolgt wird.

In einer anderen Studie mit 64 Teilnehmern wurden zwei Musikempfehlungssysteme verglichen, das von Pandora (content-based) und das von Last.fm (bewertungsbasiert), wobei alle Teilnehmer erst das eine, dann das andere System testeten. Fokus der Studie lag auf der Einschätzung des Systems und seines Interfaces nach einer relativ kurzen Probezeit, also der Nutzeradoption des Systems. Ergebnis war, dass das System von Pandora stark bevorzugt wurde, sowohl bezüglich des Interfaces als auch wegen der Qualität der Empfehlungen. Wieder wurden Richtlinien zum Gestalten eines Empfehlungssystems abgeleitet (Jones and Pu [2007]):

1. Minimiere den Nutzeraufwand wie die nötige Zeit um sich zu registrieren und Empfehlungen zu erhalten
2. Maximiere die Qualität der Empfehlungen insofern, dass sie einem Richtwert wie dem Geschmack eines Freundes entsprechen und sie genossen werden können und tatsächlich neue Songs empfehlen
3. Maximiere die Einfachheit des Interfaces

Wieder kommen diese Richtlinien in den Bereich normaler Software-Usability und User-Experience-Empfehlungen, und doch ist wieder ein Hinweis auf die Wichtigkeit der Transparenz der Empfehlungen zu finden.

Dies unterstreicht das Ergebnis eines früheren Experiments. Ein Empfehlungssystem wurde simuliert, das 100 Teilnehmern Empfehlungen für Filme aussprach. Dabei musste jeder Teilnehmer vier Phasen durchlaufen: Zuerst musste er sich registrieren und dabei ein Profil ausfüllen. Dann musste er 20 bis 30 Filme bewerten. Nächster Schritt war das Wählen jeweils einer Empfehlung aus 48 Paaren von Empfehlungen. Schließlich wurden einige der vorher genutzten Profile bewertet. Ergebnis des Experiments war, dass zwei Faktoren Einfluss darauf hatten welcher Empfehlung geglaubt wurde: Ähnlichkeit des Profils zum eigenen und die als Variable ausgedrückte Übereinstimmung bei der Bewertung von Filmen, also wiederum Ähnlichkeit, nämlich die Ähnlichkeit des Filmgeschmacks (Bonhard et al. [2006]). Dies zeigt wiederum, dass Transparenz über die Herkunft der Empfehlung, sowie die Beachtung persönlichkeitsgebundenen Faktoren, ein Faktor für ein als gut empfundenenes Empfehlungssystem sein kann, wenn dieses auf die spezifischen Empfehlungen einzelner Nutzer basiert.

Transparenz bedeutet, dass den Nutzern die Herkunft der Empfehlung erklärt wird. Diese Erklärungen mit den Empfehlungen anzuzeigen kann verschiedene Ziele haben (Tintarev and Masthoff [2007]):

- Transparenz: Die Funktionsweise des Systems erklären
- Diagnostizierbarkeit: Fehler des Systems erkennbar machen

- Vertrauen: Nutzer stärker dem System glauben lassen
- Effektivität: Nutzern helfen, bessere Entscheidungen zu treffen
- Überzeugung: Nutzern zum Kauf oder zu einer Entscheidung verleiten
- Effizienz: Nutzer Entscheidungen schneller treffen lassen
- Zufriedenheit: Nutzerzufriedenheit erhöhen, durch bessere Usability

In [Tintarev and Masthoff, 2007] wurde in zwei Fokusgruppen mit 11 akademischen Teilnehmern versucht herauszufinden, welche Informationen die Teilnehmer über ihre Filmempfehlungen erhalten wollen, um darüber zu konstruieren, welche Informationen in den Erklärungen eines Empfehlungssystems enthalten sein sollten. Ihr Fazit war, dass Erklärungen je nach Kontext und Nutzer unterschiedliche Faktoren beinhalten sollten, die genannte Quelle Auswirkungen hat (falls eine Erklärung beispielsweise auf der Bewertung eines Freundes basiert), aber die Anzahl an nötigen Features der Erklärungen relativ gering war.

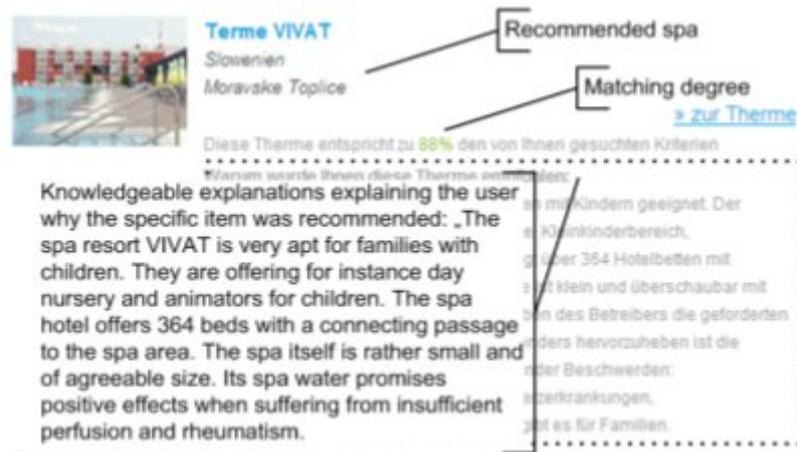


Abbildung 7: Empfehlung und ihre Erklärung bei der Suche nach Spas, aus Zanker [2012]

Der Wert von Erklärungen für Empfehlungssysteme wurde durch bestätigt durch die Ergebnisse von [Zanker, 2012]. In einem Online-Experiment wurden Nutzer einer Suche nach Spas Ergebnisse empfohlen (siehe Abbildung7) und diese Nutzer in zwei Gruppen unterteilt, wovon eine Erklärungen präsentiert bekam und die andere nicht. Anschließend wurden die Teilnehmer in einer an TAM angelehnten Online-Umfrage befragt. 134 Ergebnisse wurden ausgewertet, Ergebnis war, dass Erklärungen die wahrgenommene Nützlichkeit des Systems deutlich erhöhten. Gleiches gilt für die wahrgenommene Verständnis und Kontrolle des Systems, zumindest bei sozialen Empfehlungssystemen (Knijnenburg et al. [2012]). Ähnliches bestätigt Cramer et al. [2008], wobei sich dort

zeigte, dass gesteigertes Vertrauen in die Empfehlungen nicht automatisch auf das Gesamtsystem übertragen wird.

Fraglich blieb bisher, wie diese Erklärungen genau aussehen sollten. Eine Antwort ist, dass sie erklären sollten, warum genau diese Empfehlung abgegeben wurde (Lim et al. [2009]). Das ist nicht völlig selbstverständlich - so könnte eine Erklärung auch erklären, was genau gerade passiert ist oder warum nicht etwas anderes empfohlen wurde, was es unter bestimmten Umständen empfehlen würde oder wie etwas bestimmtes erreicht werden würde. Das folgt aus [Lim et al., 2009] insofern, als dass in dieser Arbeit ein kontextbasierendes intelligentes System mit 158 Teilnehmern getestet wurde, das Erklärungen folgend der fünf beschriebenen Kategorien präsentierte. Dann wurde das Verständnis der Nutzer getestet. Erklärungen, die das warum erklärten, schnitten mit am besten ab. Die Übertragbarkeit auf Empfehlungssysteme muss aber nicht als gegeben angesehen werden.

Dies ist bei [Symeonidis et al., 2009] anders: Dort wurde anhand eines Filmempfehlungssystem mit 42 Studenten getestet, welche Art von Erklärungen (Keywords, KSE - z.B. der Schauspieler, Einfluss, ISE - z.B. Anzahl bewerteter Filme mit einem Feature, oder KISE, eine Kombination) den Testern es am besten ermöglichte, die ausgeblendete Bewertung des empfohlenen Films vorherzusagen. Ergebnis war die kombinierte Wertung, was zeigen sollte, dass Nutzer aus ihr den meisten Nutzen ziehen konnten.

Wie aus dem bisher präsentierten Arbeiten ersichtlich, ist Kernidee einiger Forschung an Empfehlungssystemen im HCI-Bereich der Gedanke, dass Nutzer nicht automatisch das nach statistischen Methoden bestimmten beste System bevorzugen. Sonst wären Überlegungen zur Präsentation der Ergebnisse und Diskussionen über die Auswirkungen von Erklärungen unnötig. Dies gilt nochmal mehr, wenn man [Guy et al., 2009] in Betracht zieht, die zu dem Schluß kamen, dass Empfehlungen basierend auf dem Sozialen Netzwerk des Nutzers - statt auf solchen von Nutzern mit gleichen Interessen - überlegen waren.

Doch auch ohne solche Überlegungen zeigt sich, dass algorithmische Qualität alleine nicht genügt. In einer Studie mit 210 Teilnehmern und sieben verschiedenen Empfehlungssystemen (jeweils auf der gleichen Datenbasis) wurde diese Annahme untersucht. Dafür wurden sieben qualitativ unterschiedliche Empfehlungssysteme getestet:

- Correlation Neighborhood
- Non-Normalized cosine Neighborhood
- Asymmetric SVD
- TopPop
- Latent Semantic Analysis

Die Teilnehmer sollten Filme aussuchen und bekamen aufgrund ihrer Auswahl Filme empfohlen, jeweils mit verschiedenen Empfehlungssystemen. Ergebnis war, dass von den verwendeten Algorithmen der simpelste, TopPop, als am relevantesten eingestuft wurde, während bei der statistischen Auswertung dieser die höchste Fehlerrate hatte. Das Fazit

der Autoren war, dass die statistischen Evaluierungsmethoden nicht immer ausreichen, um die Qualität eines Empfehlungssystems für Nutzer zu antizipieren (Cremonesi et al. [2011]).

Diese HCI-Perspektive auf Empfehlungssysteme wurde auch von [Swearingen and Sinha, 2001] eingenommen. In dieser Arbeit wurde mit 19 Teilnehmern untersucht, welche Faktoren die Akzeptanz verschiedener Empfehlungssysteme für Bücher und Filme beeinflussen. Ergebnis waren Hinweise darauf, dass - außer dass wieder Erklärungen Auswirkungen haben sollten - auch Faktoren wie die Interfacegestaltung und die vorhandene Beschreibung des empfohlenen Produktes Auswirkungen auf die Bewertung des Systems haben. Dazu käme der individuelle Hintergrund des Nutzers, da es vertrauensfördernd sein könne, wenn Nutzer Dinge empfohlen bekommen, die sie selbst als gut einschätzen.

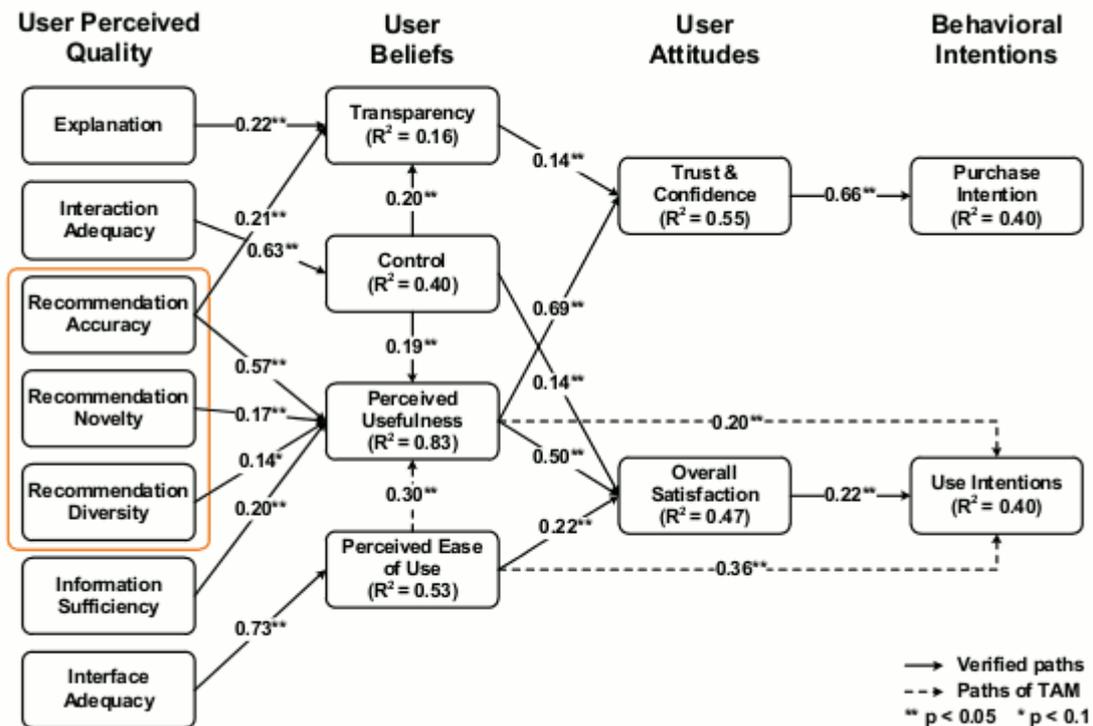


Abbildung 8: Das Modell hinter ResQue, übernommen aus Pu et al. [2011].

Dementsprechend und aufbauend auf anderen Arbeiten, welche die Bedeutung von Faktoren abseits der Empfehlungsgenauigkeit zeigen, wurde das Evaluationsframework *ResQue* entworfen und getestet. Dieses Framework evaluiert mithilfe von Methoden, die auf TAM und ähnlichen Modellen aufbauen, nicht nur die Einschätzung des Nutzers zur Qualität der Empfehlungen, sondern auch weiterer Empfehlungen wie der Qualität des Interfaces oder den aus TAM bekannten Kriterien wahrgenommener Nützlichkeit und Einfachheit der Nutzung (siehe Abbildung 8). Die Methodik der Evaluation ist wieder ein Fragebogen, mit 32 Fragen, wobei eine kürzere Version vorgeschlagen wurde. Die

Evaluation der Autoren des eigenen Modells bestätigte das Modell (Pu et al. [2011]). Es wäre eine Möglichkeit gewesen, mit diesem Evaluationssystem das Empfehlungssystem dieser Arbeit zu evaluieren.

Doch die Akzeptanz von Empfehlungssystemen kann auch schlicht davon abhängen, welcher Ansatz für das Empfehlungssystem genutzt wurde. In einer Studie mit 30 Teilnehmern wurden zwei Filmempfehlungssysteme verglichen, wobei das eine System einen Quiz-basierenden Persönlichkeitsansatz hatte, und das andere auf Bewertungen basierte. Für die Studie wurden die unterschiedlichen Systeme angeglichen. Ergebnis war, dass das Quiz-basierte System wesentlich bevorzugt wurde, nur bei der Genauigkeit der Empfehlungen hatte das Bewertungsbasierte einen kleinen Vorsprung (Hu and Pu [2009b]).

The product best matching your preferences		
	Lenovo ThinkPad T61 Notebook Add to saved list \$ 1028 Lenovo, Microsoft Windows XP Pro, 3.6 hours battery, 14.1 in display size, 160 GB HD, 1 GB memory, Intel Core 2 Duo Mobile, 2.2 GHz processor speed, 2.265 kg weight. detail	Specify your own criteria for "Better Features"
Recommended other products with some Better Features		
These products have Different Manufacturer and Longer Battery Life, although they have slightly Heavier Weight		
	Sony VAIO VGN-NR260E/S Notebook Add to saved list \$ 755 Sony, Microsoft Windows Vista Home Premium, 5 hours battery, 15.4 in display size, 200 GB HD, 2 GB memory, Intel Pentium Dual-Core Mobile, 1.6 GHz processor speed, 2.6 kg weight. detail	Better Features
	Sony VAIO VGN-FZ240E/B Notebook Add to saved list \$ 949 Sony, Microsoft Windows Vista Home Premium, 5.5 hours battery, 15.4 in display size, 250 GB HD, 2 GB memory, Intel Core 2 Duo Mobile, 2 GHz processor speed, 2.673 kg weight. detail	Better Features
6 products Show All		
These products have Cheaper Price and Different Operating System, although they have slightly Smaller Hard Drive Capacity		
	Apple 12.1 iBook G4 Notebook Add to saved list \$ 659 Apple, Mac OS X, 6 hours battery, 12.1 in display size, 40 GB HD, 0.5 GB memory, PowerPC G4, 1.33 GHz processor speed, 2.22 kg weight. detail	Better Features
	Sony VAIO VGN-N365E/B Notebook Add to saved list \$ 656.14 Sony, Microsoft Windows Vista Home Premium, 4 hours battery, 15.4 in display size, 120 GB HD, 1 GB memory, Intel Core Duo, 2 GHz processor speed, 2.945 kg weight. detail	Better Features

Abbildung 9: Das getestete strukturierte Empfehlungssystem, übernommen aus Chen and Pu [2008].

Von Reichling et al. wurde die Einführung eines Experten-Empfehlungssystems beschrieben. Das Szenario dort war etwas anders als hier: Zum einen wurde das System in eine Organisation eingeführt. Dementsprechend wurde die Selbstdarstellung der Angestellten eine Problematik, die wollten, dass sie zu ihrem Kerngebiet als Experten geführt würden (Reichling et al. [2007], Reichling and Wulf [2009]). Außerdem ist ein Experten-Empfehlungssystem etwas anderes als ein gewöhnliches Empfehlungssystem, geht es doch bei diesem schlicht nicht darum einen Experten für ein Gebiet zu bestimmen, sondern um die Antizipation des Wunsches nach einer bestimmten Ressource, mit der der Nutzer dann interagieren würde. Trotzdem gibt es natürlich Parallelen: Die Einführung eines

Artefakts in ein bestehendes System und die Beobachtung, ob der Nutzer dieses Artefakt akzeptiert, wobei das Artefakt in beiden Fällen ein (wenn auch eben nicht gleichartiges) Empfehlungssystem ist. Dies zeigt die Bedeutung von Faktoren neben der Qualität der Empfehlungen selbst.

Einer dieser Faktoren scheint der kulturelle Hintergrund zu sein. In [Chen and Pu, 2008] wurden zwei Interfaces eines Empfehlungssystems verglichen, eine strukturierte Variante mit mehreren Kategorien von anderen empfohlenen Produkten (siehe Abbildung 9), die jeweils genauer erklärt werden, und eine listenbasierte Variante mit einer kürzeren Begründung der Empfehlung (siehe Abbildung 10). Dieser Vergleich wurde mit zwei Gruppen durchgeführt, jeweils überwiegend Studenten, einmal mit 60 Teilnehmern aus China und einmal mit 60 Teilnehmern aus Europa. Beide Gruppen bevorzugten die strukturierte Variante, aber die chinesische Gruppe bevorzugte die strukturierte Variante stärker als die europäische.

The product best matching your preferences			
Best Fit! why?		Lenovo ThinkPad T61 Notebook Add to saved list \$ 1028 Lenovo, Microsoft Windows XP Pro, 3.6 hours battery, 14.1 in display size, 160 GB HD, 1 GB memory, intel Core 2 Duo Mobile, 2.2 GHz processor speed, 2.265 kg weight. detail	Specify your own criteria for "Better Features"
Recommended other products with some Better Features			
2nd Best why?		Lenovo ThinkPad X61 Notebook Add to saved list \$ 899.99 Lenovo, Microsoft Windows Vista Business, 3.9 hours battery, 12.1 in display size, 120 GB HD, 1 GB memory, Intel Core 2 Duo Mobile, 2 GHz processor speed, 1.4 kg weight. detail	Better Features
3rd Best why?		Lenovo ThinkPad T42 Notebook Add to saved list \$ 419 Lenovo, Microsoft Windows XP Pro, 4.5 hours battery, 14.1 in display size, 40 GB HD, 0.5 GB memory, Intel Pentium M, 1.7 GHz processor speed, 2.22 kg weight. detail	Better Features
4th Best why?		Lenovo Thinkpad R61i Notebook Add to saved list \$ 609.95 Lenovo, Microsoft Windows XP Pro, 3.5 hours battery, 15.4 in display size, 80 GB HD, 1 GB memory, Intel Core 2 Duo Mobile, 1.5 GHz processor speed, 2.673 kg weight. detail	Better Features
5th Best why?		Lenovo ThinkPad T23 Notebook Add to saved list \$ 299 Lenovo, Microsoft Windows 2000 Professional, 3.5 hours battery, 14.1 in display size, 30 GB HD, 0.125 GB memory, Intel Pentium III, 1.13 GHz processor speed, 2.5 kg weight. detail	Better Features

Abbildung 10: Das getestete listenartige Empfehlungssystem, übernommen aus Chen and Pu [2008].

Die Zielgruppe kann aber auch anderer Hinsicht relevant werden. Beispielsweise, wenn ein Empfehlungssystem nicht etwas für ein Individuum empfehlen soll, sondern eine Gruppe adressiert wird. Dann werden Faktoren wie der soziale Einfluss oder die Bereitschaft, eigene Interessen denen der anderen unterzuordnen, relevant. Außerdem könnte dann nicht nur die Einzelinteressen relevant sein, sondern auch die Eignung der Empfehlung, eine langfristig funktionierende Gruppe zu formen (Herr et al. [2012]).

Datenschutzbedenken können dazukommen. In [Hu and Pu, 2010] zeigte sich, dass die gemessene Intention der Nutzer, ein - diese per Fragebögen evaluierend - persönlich-

keitsbasierendes Empfehlungssystem erhebliche Bedenken bei den Nutzern verursachte, sodass die Autoren vermuteten, dass dies die Bewertung des Systems negativ beeinflusste. Für Empfehlungssysteme, die solche Bedenken berechtigt zerstreuen können, z.B. durch das glaubwürdige Einsetzen von Verschlüsselung und Hashverfahren zur Unkenntlichmachung der Daten, besteht Bedarf (Konstan and Riedl [2012]).

2.4 Datenschutz und Privacy

Teil der Grundfrage und Annahme dieser Arbeit ist, dass die Teilnehmer erwarten, dass ihre Daten in einem gewissen Maße geschützt sind. Selbstverständlich wissen wir inzwischen, dass alle im Internet übertragenen Daten von Geheimdiensten abgehört und Verschlüsselungen gebrochen werden wo möglich. Doch das Vorhandensein dieser Überwachung bedeutet nicht, dass ein spezifisches System nicht mehr auf Datenschutz achten müsste. Ein Empfehlungssystem wie das in dieser Arbeit entwickelte muss mit negativen Reaktionen rechnen, wenn es private Daten im Übermaß verwendet und anderen anzeigt (wie an Hu and Pu [2010] zu sehen), das Wissen ob des Vorhandenseins der geheimdienstlichen Überwachung dürfte diese Wahrnehmung nicht ändern, insbesondere nicht instantan.

Rechtlicher Leitlinie dieses Themenkomplexes ist das Recht auf informationelle Selbstbestimmung. Es wird definiert als (Bundesverfassungsgericht [1983]):

... die Befugnis des Einzelnen, grundsätzlich selbst über die Preisgabe und Verwendung seiner persönlichen Daten zu bestimmen.

Grundsätzlich sind alle privaten Daten entsprechend geschützt sind, auch weil durch das Vorhandensein von Computersystemen nicht eindeutig absehbar sei, welche Daten durch Verknüpfung eine höheren Stellenwert bekommen. Das Vorhandensein dieser Argumentation zeigt jedoch, dass es private Daten persönlicher Natur geben kann, die von einem Nutzer zumindest initial stärker als schutzwürdig betrachtet werden als andere, und dies sind wahrscheinlich solche, die den Kernbereich der Persönlichkeitssphäre betreffen.

Ein Nutzer in Konfrontation mit einem Empfehlungssystem könnte also eine Erwartung von Anonymität haben. Anonymität in einem IT-System würde bedeuten, dass (Pfitzmann and Hansen [2010]):

... the subject is not identifiable within a set of subjects, the anonymity set.

Entsprechend dieser Definition gibt es Modelle, um dieses Anonymitätsset für Datensätze in Datenbanken zu formalisieren. Dies ist nicht völlig das Szenario dieser Arbeit, zeigt jedoch, wie im akademischen Umfeld über Anonymity und damit Datenschutz gedacht wird.

Eines dieser Modelle lautet k-anonymity. In einem Datenbankmodell mit einer k-anonymity von drei würden die Nutzerdaten soweit gekürzt, dass zwischen immer drei Nutzern nicht unterschieden werden kann, weil alle ihre zur Identifizierung geeigneten Daten entsprechend gekürzt wurden. Es gäbe in der Datenbank also immer mindestens

drei Nutzer mit gleicher Postleitzahl, Alter und Krankheit, wenn das die einzigen zur Identifizierung nutzbaren Daten wären (SWEENEY [2002]).

Ein solches Set kann aber immer noch Informationen transportieren, die nicht gewünscht sind. Konkret sind zwei Attacken auf solche Datensätze bekannt:

1. Bei einer *homogeneity attack* werden trotz Gruppierung kritische Daten nicht verschleiert. Sind Grunddaten über ein Individuum bekannt, kann möglicherweise über diese Grunddaten die Menge erkannt werden, in der die Person enthalten ist, und herausgelesen werden, was beispielsweise die Krankheit ist, die alle Personen in dieser Gruppe haben
2. Bei einer *Background Knowledge Attack* wird ebenso wie bei der vorherigen Attacke mit vorhandenen Grunddaten das Anonymitätsset identifiziert. Angenommen, die drei Personen hätten zwei verschiedene Krankheiten. Aufgrund von weiteren vorhandenen Informationen (beispielsweise: Es ist bekannt, dass es sich um eine schwere Krankheit handelt, eine der Krankheiten ist jedoch eine leichte) kann die konkrete Person identifiziert werden

Gegen solche Angriffe soll *l-diversity* schützen. Anonymitätssets werden nach diesem Prinzip so zusammengesetzt, dass die sensitive Information (in dem Beispiel der Angriffe: die Krankheit) in allen Anonymitätssets ausreichend häufig vorkommt, um beide beschriebene Angriffe unmöglich zu machen (Machanavajjhala et al. [2007]), indem in jedem Set mindestens l Vorkommen des zu schützenden Attributes auftauchen. l -diversity kann demnach als Maßzahl dafür gesehen werden, wie wenig der Datensatz dafür geeignet ist, Unsicherheit (im Sinne von Unwissen) bei einem Betrachter des Datensatzes zu reduzieren (Bezzi [2010]).

	ZIP Code	Age	Salary	Disease
1	47677	29	3K	gastric ulcer
2	47602	22	4K	gastritis
3	47678	27	5K	stomach cancer
4	47905	43	6K	gastritis
5	47909	52	11K	flu
6	47906	47	8K	bronchitis
7	47605	30	7K	bronchitis
8	47673	36	9K	pneumonia
9	47607	32	10K	stomach cancer

Tabelle 1: Beispiel einer Originaltabelle, übernommen aus Li et al. [2007]

Doch auch l -diversity gilt nicht als perfekt. Es wurde argumentiert, dass es als Maßzahl weder ausreichend noch notwendig sei, um Daten zu schützen (siehe Tabellen 1, 2 und 3). Stattdessen wurde t -closeness vorgeschlagen. t -closeness ist ein Prinzip, demnach in jedem Anonymitätsset die Verteilung des sensiblen Attributes bis auf den Faktor t der Verteilung des Attributes im Gesamtdatensatz entsprechen soll (Li et al. [2007]). t -closeness kann also als Maßzahl dafür betrachtet werden, wie wenig der Datensatz dafür

	ZIP Code	Age	Salary	Disease
1	476**	2*	3K	gastric ulcer
2	476**	2*	4K	gastritis
3	476**	2*	5K	stomach cancer
4	4790*	≥ 40	6K	gastritis
5	4790*	≥ 40	11K	flu
6	4790*	≥ 40	8K	bronchitis
7	476**	3*	7K	bronchitis
8	476**	3*	9K	pneumonia
9	476**	3*	10K	stomach cancer

Tabelle 2: Beispiel einer problematischen Tabelle trotz einer l-diversity von 3, übernommen aus Li et al. [2007]

	ZIP Code	Age	Salary	Disease
1	4767*	≤ 40	3K	gastric ulcer
3	4767*	≤ 40	5K	stomach cancer
8	4767*	≤ 40	9K	pneumonia
4	4790*	≥ 40	6K	gastritis
5	4790*	≥ 40	11K	flu
6	4790*	≥ 40	9K	bronchitis
2	4760*	≤ 40	4K	gastritis
7	4760*	≤ 40	7K	bronchitis
9	4760*	≤ 40	10K	stomach cancer

Tabelle 3: Bearbeitete Tabelle mit minimierter t-closeness, übernommen aus Li et al. [2007]

geeignet ist, dass ein Betrachter des Datensatzes aus ihm neue Informationen gewinnt (Bezzi [2010]).

In [Dwork, 2006] wird argumentiert, dass es unmöglich ist, eine Datenbank so zu gestalten, dass ein Angreifer (der über weitere Informationen verfügen kann) unter keinen Umständen Informationen aus der Datenbank gewinnen kann. Deswegen schlägt Dwork eine neue Betrachtungsweise für Privacyverletzungen vor: Das Risiko für negative Konsequenzen, das ein Individuum durch die Veröffentlichung des Datensatzes trägt, sollte durch die Einbezogenheit der eigenen Daten in einer Datenbank nicht wesentlich steigen. Um diesem Risiko zu begegnen wird *differential privacy* definiert: Bei einer Abfrage der Datenbank werden immer zufällig generierte Stördaten mit ausgegeben. Diese sind so generiert, dass sie die Definition von ϵ -*differential privacy* erfüllen. Dies ist eine Definition, demnach durch das Entfernen einzelner Datensätze die Wahrscheinlichkeit für bestimmte Ausgaben sich nicht wesentlich erhöht:

A randomized function K gives ϵ -differential privacy if for all data sets $D1$ and $D2$ differing on at most one element, and all $S \subseteq \text{Range}(K)$,

$$\Pr[K(D1) \in S] \leq \exp(\epsilon) \times \Pr[K(D2) \in S] \quad (3)$$

Doch in dieser Arbeit wird gar keine Datenbank gestaltet oder gepflegt, auch wenn Datensätze einer entnommen werden. Stattdessen werden Informationen Nutzern angezeigt und es geht um die Frage, welche Informationen Nutzer über sich wann wem preisgeben wollen. Statt um informationelle Selbstbestimmung gegenüber einer Organisation oder Staat geht es also primär um persönliche Privacy, um das Preisgeben von Daten und das Bewahren der Privatsphäre gegenüber anderen Individuen (Iachello and Hong [2007]). Wenn es stimmt, dass Individuen sich kontinuierlich in einem Prozess befinden, in dem sie zwischen ihrem Bedürfnis nach Privacy und ihren Mitteilungs- und Kommunikationsbedürfnis abwägen (Westin [2001]), ist Privacy in einem Empfehlungssystem auf einer Webplattform, durch welches das eine Bedürfnis zugunsten des anderen verletzt wird, für die Nutzer ein relevantes Thema.

Dementsprechend gibt es Arbeiten über konkrete Lösungsversuche bei Systemen, die ebenfalls zwischen Privacy und Datenpreisgabe abwägen müssen. Eine davon ist [Neustaedter and Greenberg, 2003]. In einem sogenannten home media space, durch das zuhause arbeitende Mitarbeiter über Videoübertragung mit Kollegen in Kontakt treten können und selbst erreichbar sein sollten, wurden verschiedene Lösungen ausprobiert. Dafür wurde zuerst eine vorher vorgeschlagene Lösung verworfen, das blurren des Videobildes, um unerwünschte Informationen zu verdecken. Vorgeschlagen wurden stattdessen explizite und implizite Mechanismen, die den Anwender schützen sollen: Beispielweise die Einstellung, die Kamera zu deaktivieren, sobald eine weitere Person ins Bild kommt. Oder wenn 5 Minuten keine Person im Bild war. Das wurde kombiniert mit Schiebereglern, um die Kamera und das Bild zu steuern und so im Zweifel zu deaktivieren.

Grundlegendere Erkenntnisse über die Erwartungen von Nutzern an IT-Systeme kommen aus einer Interviewanalyse. Ihrzufolge sei eine wichtige Motivation des Wunsches nach Privacy von Nutzern von Chatsoftware der Wunsch der Nutzer, die Selbstdarstellung gegenüber den anderen Nutzern zu steuern (Dearden and Watts [2004]).

Doch die obige Aussage zeigt eines der Probleme solcher Überlegungen: Aus einem konkreten Anwendungsfall zu versuchen, allgemeingültige Nutzererwartungen zu extrahieren, ist eine fragwürdige Vorgehensweise. Dies steht in Verbindung mit der eingangs beschriebenen Unterscheidung zwischen Privacy als Datenschutz, bei der Daten als Teil einer Datenbank betrachtet werden, für die dann auch bestimmte allgemeine Richtlinien wie eine *k-anonymity* gelten können, und Privacy als persönliches Schutzbedürfnis eines einzelnen Nutzers während der Interaktion mit einem bestimmten System, oder eben auch mehreren verschiedenen Systemen. Denn während für die Datenschutzsichtweise die beschriebenen allgemeingültigen Regeln definiert werden können, ist das bei dem Konzept persönlicher Privacy nicht so einfach anzunehmen. Also würde es nicht gelingen, in diesem Bereich einfache Regeln wie *Anwender möchten keine Fotos von sich teilen* aufzustellen. Stattdessen müsste bei jedem System in jedem Kontext immer wieder neu ausgehandelt werden, was erlaubt ist (lose aus Iachello and Hong [2007]).

Entsprechend argumentiert [Barkhuus, 2012] und bezieht sich dabei auf Nissenbaums Contextual Integrity (Nissenbaum [2004]), demnach es bei Privacyüberlegungen um die Angemessenheit des gesamten Informationsflusses ginge. Weiterführend wird so argumentiert:

Third and finally, the issues we as researchers should address regarding privacy are more complex than just concern. Privacy is also a way for us to maintain distance to weak-tie acquaintances and it is something we maintain by not ‘bothering’ others, as illustrated above. There are many reasons for hesitancy to share personal information, from courtesy and modesty to uncertainty about the audience.

Der Wunsch nach Datenschutz sei also nicht einfach ein persönliches Bedenken, sondern könne aus verschiedenen Motivationen wie beispielsweise dem Wunsch nach Aufrechterhaltung angemessener gesellschaftlicher Distanz entstammen. Forschung müsse die Motivation konkreter Forderungen und Bedenken untersuchen.

Dies ist auch ein Gedanke der *Collective Information Practice* (Dourish and Anderson [2006]). Diesem Modell zufolge soll Privacy nicht als eigenständiges Konstrukt begriffen werden, bzw. sei so nicht begreifbar. Stattdessen sei Privacy eingebettet in eine kollektive Praxis des Umgangs mit Informationen, und ein davon betroffenes System könne nur unter Beachtung dieser kollektiven Praxis mit ihren fortschreitenden Änderungen und Neudefinitionen (die Parallele zu Foucaults Diskursbegriff ist auffällig) sinnvoll gestaltet werden kann.

All das harmonisiert mit Vertrauen als eigenständigem Thema, das in diesen Themenkomplex hineinwirkt. Patrick et al. formuliert:

A lack of trust will result in systems being ill-used at best, and not used at all at worst. A lack of understanding of trust, in both user and system, will result in the wrong decision being made in security contexts or no decision at all.

Desweiteren wird beschrieben, dass Vertrauen als mehrstufiges Modell beschrieben werden könnte, wobei in verschiedenen Phasen verschiedene Faktoren unterschiedliche Aus-

wirkungen haben könnten. Beispielsweise sei das professionelle Aussehen einer Webseite besonders bei der ersten Bewertung wichtig. Entsprechend könnte die Einhaltung von Guidelines wie diese das Vertrauen in das System stärken, und damit die Abwägung zwischen Privacy und anderen Interessen zugunsten der Nutzung des Systems beeinflussen (Patrick et al. [2005]):

1. Ensure good ease of use.
2. Use attractive design.
3. Create a professional image—avoiding spelling mistakes and other simple errors.
4. Don't mix advertising and content—avoid sales pitches and banner adverts.
5. Convey a 'real world' look and feel, for example with the use of high quality photographs of real places and people.
6. Maximise the consistency, familiarity, or predictability of an interaction both in terms of process and visually.
7. Include seals of approval such as TRUSTe.
8. Provide explanations, justifying the advice or information given.
9. Include independent peer evaluation such as references from past and current users and independent message boards.
10. Provide clearly stated security and privacy statements, and also rights to compensation and returns.
11. Include alternative views, including good links to independent sites within the same business area.
12. Include background information such as indicators of expertise and patterns of past performance.
13. Clearly assign responsibilities (to the vendor and the customer).
14. Ensure that communication remains open and responsive and offer order tracking or alternative means of getting in touch.
15. Offer a personalized service which takes account of each client's needs and preferences and reflects their social identity.

Hier sollte man allerdings beachten, dass einige dieser Punkte Empfehlungen sind, die man auch aus generellen Usability-Richtlinien wie [Norman, 2002] ableiten könnte.

Eine Methode zum Erreichen eines Designs, das dem Geist dieser Empfehlungen entspricht, könnte in [Iachello and Abowd, 2005] beschrieben worden sein. Dort wird versucht, auf die Fundierung der Rechtssprechung zurückzugreifen, indem ein Design unter

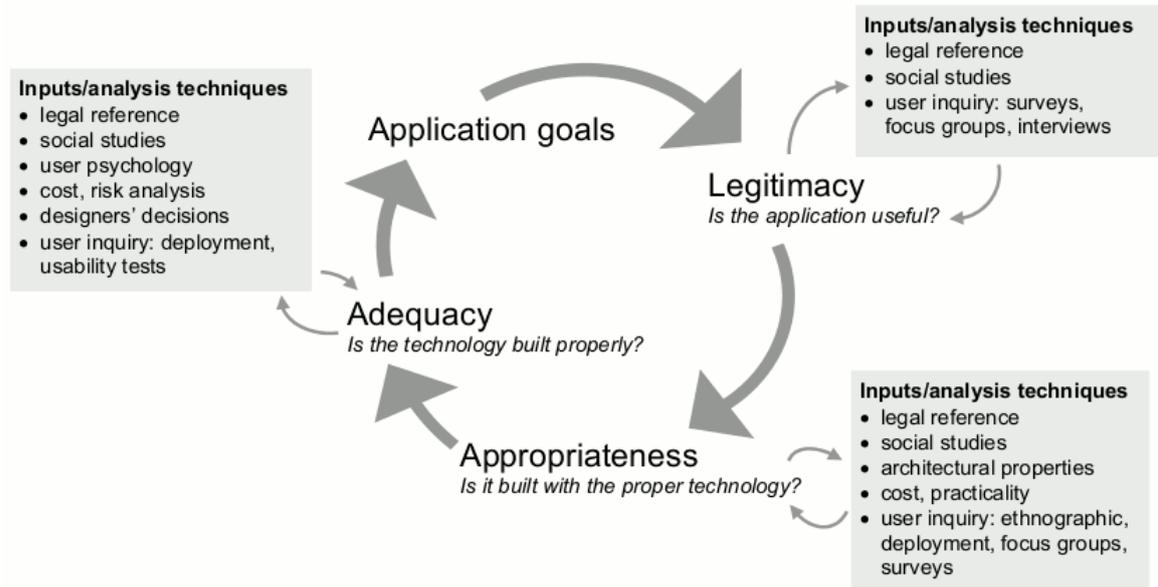


Abbildung 11: Designmethode angelehnt an juristische Abwägungsprozesse, übernommen aus Iachello and Abowd [2005].

Abwägung von Legitimität, Angemessenheit und Eignung (legitimacy, appropriateness, adequacy, siehe auch Abbildung 11) im Hinblick auf Datenschutzbedenken motiviert wird.

Insgesamt zeigt sich, dass Datenschutz, Anonymität und Privacy ein kompliziertes Thema mit vielen Ausprägungen ist, insbesondere unter dem Blickwinkel der HCI. Ein weiteres Beispiel für die Problematik, Privacythemen mit allgemeingültigen Regeln zu lösen, soll diese Aussage untermauern. In [Peerce and Shneiderman, 2009] wird ein Wandel in der Einstellung zu Privacy beschrieben, der sich bei der Beobachtung verschiedener Generationen zeige. Jüngere Nutzer zeigten generell wenig Problembewusstsein, wenn Bilder von Feiern in entsprechend spärlicher Kleidung im Internet auftauche und dort auf unbestimmte Zeit verfügbar bleibe. Es wird suggeriert, dass dies bei älteren Nutzern generell anders sei - aber es wird auch behauptet, dass die Einstellung von einzelnen Nutzern einer gemeinsamen Generation stark variere. Hier sei beachtet, dass dieser Beobachtung die Prämisse zugrunde liegt, dass Fotos von leichtbekleideten Menschen auf Feiern etwas problematisches sei, etwas, was später unter Nichtbeachtung des Kontextes negativ interpretiert würde, was wahrscheinlich ebenfalls eine generationen- oder klassenabhängige Einstellung sein dürfte, oder einfach auf einer bestimmten Moralvorstellung beruht. Nichtsdestotrotz, die Überlegung, dass dies ein Faktor bei der Bewertung Sozialer Systeme sei, stimmt mit Aussagen der bisher angeführten Arbeiten überein.

3 Verwandte Systeme

3.1 Nutzung von Empfehlungen in Websystemen

Empfehlungssysteme haben zwei typische Anwendungsfälle: Soziale Netzwerke wie Facebook und Online-Shops wie Amazon.

Auf Seiten wie Facebook erfüllen Empfehlungssysteme primär den Zweck, Nutzern Informationen über andere Teilnehmer anzuzeigen. So zeigt Facebook eine Seitenleiste mit möglichen Bekanntschaften, wobei für diese Information eine Reihe von Informationen genutzt werden könnten, wie die Anzahl gemeinsamer Freunde oder das Vorhandensein der eigenen Adresse im Telefonbuch des anderen. Dies soll das Engagement der Nutzer auf der Seite erhöhen und funktioniert von Anfang an, selbst wenn solche Empfehlungen direkt bei der Registrierung des neuen Nutzers angezeigt werden (Freyne et al. [2009]).



Abbildung 12: Beispiel für von Facebook genutzte Empfehlungsarten.

Aber es gibt weitere Empfehlungsarten, wie das Präsentieren von Prominenten (siehe Abbildung 12), von Gruppen oder von Veranstaltungen, wobei diese Anzeige teilweise kontextsensitiv zur gerade aufgerufenen Seite ist.

Facebook stellt hier meist keinerlei Transparenz her, warum die Empfehlung präsentiert wird. Dies könnte die Akzeptanz dieser Empfehlungen verschlechtern, was besonders in Deutschland angesichts der Datenschutzdiskussion über Facebook von Bedeutung sein

könnte. Doch [Iachello and Abowd, 2005] folgend könnte man zumindest versuchen, die beiden beschriebenen Empfehlungsarten auf Nutzen und Problematik für die Nutzer hin zu analysieren:

Prominentenempfehlung

Zuerst zur Empfehlung von Prominenten als Submenge von Unterhaltungsseiten. Das Ziel einer solchen Empfehlung dürfte die Stärkung der Facebookplattform sein: Für Facebook ist es von Vorteil, wenn die Kommunikation zwischen Prominenten und Fans über Facebook läuft, um Nutzer zu binden und die Attraktivität der Plattform zu erhöhen. Für Nutzer könnte es interessant sein, über diesen Weg mit ihren Stars in Kontakt zu treten.

Die Empfehlung der Prominenten selbst dürfte unproblematisch sein. Sie sind sowieso Teil des öffentlichen Lebens und ihre Bilder scheinbar aus offiziellen Quellen, ihre Privatsphäre bleibt unbeschädigt. Andererseits könnte es ein rechtliches Problem sein, wenn hier der Eindruck einer offiziellen Seite erweckt würde, diese aber automatisch oder von Fans erstellt wäre.

Aber zusätzlich wird angezeigt, welche Freunde die Seite und damit den Prominenten mit *Gefällt mir* markiert haben. Dies erfüllt zwei Ziele:

1. Die Signifikanz der Empfehlung wird erhöht
2. Eine Erklärungsmöglichkeit, warum die Seite empfohlen wird, wird angeboten

Es ist aber auch etwas, was einen Nutzer davon abhalten könnte, eine Seite zu markieren, nämlich wenn er nicht wünscht, dass seine Zuneigung zu einem bestimmten Prominenten seinen Facebook-Kontakten mitgeteilt wird.

Freundesempfehlungen

Hier ist das Ziel klar: Die Aktivität der Nutzer auf der Plattform zu erhöhen. Die Effektivität dieser Empfehlungen wird belegt durch Studien wie [Freyne et al., 2009].

Nutzer könnten Anstoß daran nehmen, dass ihr Profilfoto Fremden gezeigt wird. Dementsprechend ist das Foto nicht nur selbst auswählbar, sondern es gibt Sichtbarkeitseinstellungen. Weiterhin könnten Empfehlungen von vermeintlichen Freunden, die von dem Nutzer nicht gesehen werden wollen (Erzfeinde, ehemalige Lebensgefährten). Dagegen existiert eine Ausblendungsfunktion.

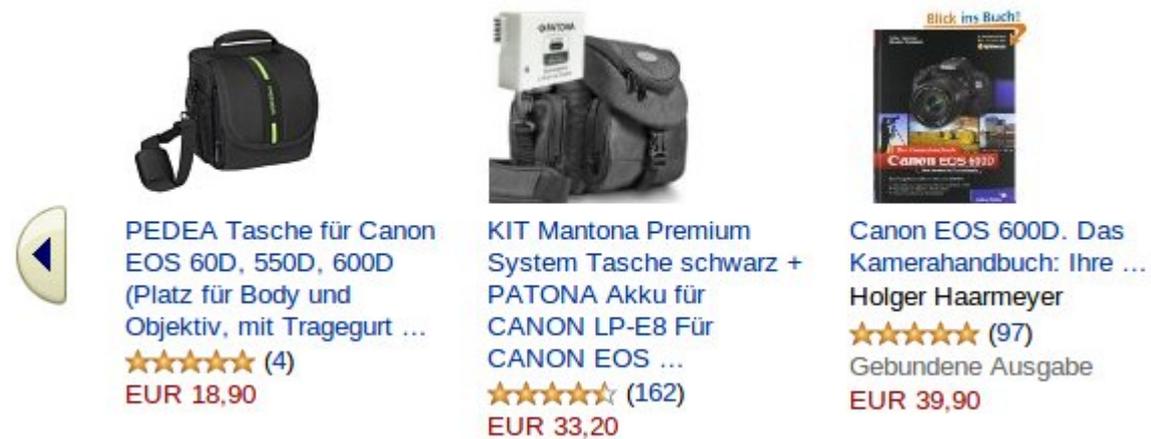
Die Anzeige der gemeinsamen Freunde dürfte die gleiche Funktion erfüllen wie die Anzeige der empfehlenden Freunde bei den Prominentenempfehlungen.

In einem Sozialen Netzwerk wurde ein Empfehlungssystem mit 2000 Nutzern getestet, das Empfehlungen für die Selbstbeschreibung in dem Profil aussprach. Grundlage war die Überlegung, dass solche Freitextfelder nur selten genutzt und mit qualitativ nicht hochwertigen Inhalten gefüllt werden. Das Empfehlungssystem empfahl daher eine Reihe von Fragen, die der Nutzer in diesem Feld beantworten sollte. Ergebnis war, dass Nutzer nicht nur relativ häufig mit den Empfehlungen interagierten (41%), sondern

sich auch die Zeit zwischen Profilaktualisierungen verringerte (Geyer et al. [2008]). Dies unterstreicht die Wichtigkeit solcher Systeme für Soziale Netzwerke.

Amazon dagegen nutzt ein Empfehlungssystem für das Präsentieren von Produkten. "Kunden, die diesen Artikel gekauft haben, kauften auch" (siehe Abbildung 13) darf wohl als archetypisch für Empfehlungssysteme gelten. Gleichzeitig ist das ein Beispiel für Transparenz, denn der Titel sagt klar, wie die Empfehlung zustande kam - oder gibt dem Nutzer zumindest ein klares Modell, wie sie zustande gekommen sein könnte.

Kunden, die diesen Artikel gekauft haben, kauften auch



Product	Rating	Price
PEDEA Tasche für Canon EOS 60D, 550D, 600D (Platz für Body und Objektiv, mit Tragegurt ...)	★★★★★ (4)	EUR 18,90
KIT Mantona Premium System Tasche schwarz + PATONA Akku für CANON LP-E8 Für CANON EOS ...	★★★★☆ (162)	EUR 33,20
Canon EOS 600D. Das Kamerahandbuch: Ihre ... Holger Haarmeyer	★★★★★ (97)	EUR 39,90

Abbildung 13: Beispiel für von Amazon genutzte Empfehlungen.

Es ist aber nicht die einzige Art von Empfehlung, die Amazon anzeigt. Ganz im Gegenteil scheinen Empfehlungssysteme ein Kernkonzept der Plattform zu sein, wobei die Empfehlungen sich auf frühere Einkäufe und betrachtete Produkte stützen und auf der Startseite und im Profil des Nutzers angezeigt werden. Teilweise wird diese Profilbildung des Kunden auch dezenter benutzt, beispielsweise wenn die Startseite allgemein beliebte Produkte anzeigt, aber die Auswahl der angezeigten Produktkategorien von Nutzer zu Nutzer variiert.

Die Ziele der Empfehlungen wurden bereits beschrieben und der tatsächlich vorhandene Nutzen wird im nächsten Absatz belegt. Doch zuerst sollen die Privacy-Bedenken dargelegt werden, gerade im Vergleich zu den Empfehlungen von Facebook. Entschärfend dürfte wirken, dass hier keine Nutzer oder Menschen empfohlen werden und die Empfehlungen nur dem Nutzer selbst angezeigt werden, hierüber erwachsen also keine Privacy-Bedenken. Diese sind abstrakter: Durch das Anzeigen maßgeschneiderter Empfehlungen wird deutlich, wieviel der Konzern über den Nutzer weiß, was diesen erschrecken und zum Nachdenken über Mißbrauchsmöglichkeiten bringen könnte, insbesondere wenn gesellschaftlich brisante Produkte gekauft wurden.

Dass Amazon solche Empfehlungssysteme nutzt ist nicht überraschend. Nicht nur nach

eigene Aussage (Linden et al. [2003]) überzeugen sie Nutzer, Produkte zu kaufen. Auch [Castagnos et al., 2010] kommt zu solchen Schlüssen: Dort wurde auf einer Parfumverkaufsseite mittels eines Eyetrackers untersucht, inwiefern Empfehlungssystem beim Einkaufsvorgang genutzt werden. Ergebnis war, dass das Empfehlungssystem in bis zu 40% der Fälle genutzt wurde, und 50% mehr Produkte in den Warenkorb gelegt wurden als bei Verzicht auf ein solches System. Solche Daten sind für Online-Shops ein guter Grund, ein Empfehlungssystem zu benutzen.

Eine parallele geschäftstüchtige Betrachtungsweise erklärt auch, warum solche E-Commerce-Seiten Erklärungen zu ihren Empfehlungen anzeigen sollten, gibt es doch außer den schon erwähnten Arbeiten zum positiven Effekt Hinweise darauf, dass Erklärungen die Interaktion der Nutzer mit den empfohlenen Produkt erhöhen (Guy et al. [2009]).

Welche Bedeutung solche Empfehlungssysteme, die sie als *Social Recommender Systems* bezeichnen, für Seiten wie Amazon haben, wird in [Kim and Srivastava, 2007] formuliert. Die Autoren sprechen dort von einem Wandel von einer transaktionsbasierten Gesellschaft hin zu einer auf Beziehungen basierenden. Entsprechend wichtig könnten Empfehlungssysteme werden, die diese soziale Komponente beinhalten, wobei die Autoren noch einige offene Fragen sehen:

- Was ist der Zusammenhang zwischen Nutzergeschmack und Vertrauen?
- Sollte eine System Produkte empfehlen, an denen ein Nutzer angeblich nicht interessiert ist, obwohl sein soziales Umfeld interessiert ist?
- Wie stark ist der Einfluss von Vertrauenspersonen auf den Nutzergeschmack?

Die Rolle von Empfehlungssystemen in Websystemen wird auch durch *SuperMusic* deutlich. Diese Anwendung für Mobilgeräte sollte Anwendern das Abspielen von Musik unterwegs ermöglichen. Zentraler Teil des Konzepts war ein Empfehlungssystem, das Musik je nach Situation oder je nach Ähnlichkeit zu einem bestimmten Lied empfehlen sollte, was in einer Evaluation von 75% der Nutzern als wichtige Funktion bewertet wurde (Lehtiniemi [2008]). Gerade auf Interfaces, bei denen durch kleine Displays das klassische Stöbern unhandlich wird, scheinen Anwendungen zu existieren, bei denen das Empfehlungssystem zum Kern des Konzepts gehört.

Doch auch in anderen Bereichen als Musik und bei mobilen Anwendungen werden Empfehlungssysteme eingesetzt. Bei [Wang et al., 2008] wurde ein Empfehlungssystem (*CHIP*) gebaut, das Kunstwerke aus einer semantisch erweiterten Datenbasis eines Museums empfiehlt. Die Arbeit zeigt auch, wie im Laufe der Evaluation die Zielgruppe bestimmt und sich daraus Fragen wie die geeigneten Interaktionsformen für die erkannten Kleingruppen ergaben, und dass FOAF/RDF für die Modellierung des Nutzermodells des Empfehlungssystems erweitert werden könnten.

3.2 Stand von Mitfahrzentralen

Mit Mitfahrzentralen (Ridesharing-Systems) sind im folgenden Webseiten und Software gemeint, die es ermöglichen, bei einer Fahrt zu einem bestimmten Ziel jemanden mitzu-

nehmen oder bei jemanden mitzufahren.

Klassische Systeme wie mitfahrzentrale.de² benutzen keine Empfehlungssysteme. Stattdessen können einzelne Fahrten bewertet werden und die Mitglieder der Seite benutzen ihr Profil, um die Durchschnittsbewertung und rudimentäre Informationen anzugeben.

Einen Schritt weiter gehen Seiten wie [blablacar](http://blablacar.de)³. Das Profil ist hier prominenter und wird im Vergleich zu gewöhnlichen Sozialen Netzwerken um einige kontextspezifische Informationen ergänzt, wie die Einschätzung der eigenen Gesprächigkeit oder ob während der Fahrt gerne Musik gehört wird. Diese Informationen werden bei Suchergebnissen angezeigt und ermöglichen es den Nutzern, mit mehr Informationen eine Fahrt auszuwählen. Aber ein Empfehlungssystem ist das nicht.

Ein Empfehlungssystem als Grundkonzept hat die Smartphoneanwendung [flinc](https://flinc.org/)⁴. Als Smartphoneanwendung wird die Lokalisierungsfunktion genutzt und es besteht die Möglichkeit, vom jetzigen Standpunkt eine Fahrt zu einem Ziel zu suchen. Bei dieser Suche wird aber nicht zwingend eine lange Liste von möglichen Fahrern präsentiert, sondern einige oder sogar der dem System nach beste Fahrer empfohlen. Dafür greift das System sowohl auf eine Bewertungsfunktion als auch auf eine Abbildung der Freundschaftsbeziehungen zurück (siehe Abbildung 14). Das Empfehlungssystem ist also direkt als Kernkonzept eingebaut, auch wenn die Anwendung auch ohne dieses funktionieren würde.

3.2.1 Kategorisierung von Mitfahrzentralen

Flinc kann dabei in Tradition einer gewissen Art von softwarebasierten Mitfahrzentralen gesehen werden. Denn eine der Entwicklungslinien solcher Plattformen läuft entlang des Vorhandenseins technischer Plattformen: Da wären Seiten wie [zimride](http://www.zimride.com)⁵, das Facebook benutzt, um vorab Profilinformationen möglicher Mitfahrer anzuzeigen. Und dann wären Anwendungen wie Piggyback und Avego, für Android bzw iOS, welche die Möglichkeiten dieser mobilen Systeme nutzen. Dies wird kritisiert, weil es wenig erfolgsversprechend sei, auf technische Plattformen zu setzen, statt zu versuchen, lokale Netzwerke zu schaffen, die für bestimmte Gebiete oder Städte passende Fahrgemeinschaften bilden können (Brereton et al. [2009]).

Doch Mitfahrzentralen können auch nach ihrem Konzept kategorisiert werden. So gibt es eine ganze Reihe von Mitfahrzentralen, wenn auch weniger in Deutschland als in den USA, die ein bestimmtes Konzept verfolgen, das über das bilden von Fahrgemeinschaften von A nach B hinausgeht. Solche Konzepte lauten (Ozenc et al. [2011]):

Routine

Bei solchen Mitfahrzentralen wird versucht, bestehende Routinen zu beachten. Beispielsweise könnten Mitfahrer mit Fahrern zusammengebracht werden, die regelmäßig eine bestimmte Strecke fahren.

²mitfahrzentrale.de

³<http://blablacar.de>

⁴<https://flinc.org/>

⁵<http://www.zimride.com>

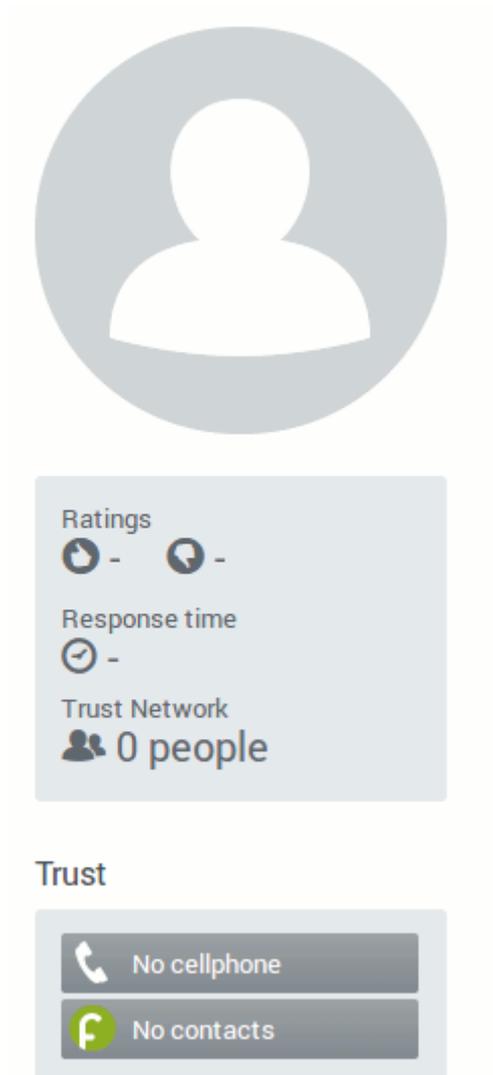


Abbildung 14: Von FlixBus angezeigte vertrauensbildende Informationen: Bewertungen, gemeinsame Bekannte, Identifizierung der Plattform gegenüber.

Systeme wie mitfahrzentrale.de bieten üblicherweise eine Funktion an, um wiederkehrende Fahrten einzutragen und so diese Kategorie abzudecken. Das Nutzungsszenario ist unterschiedlich für die Teilnehmer: Solchen mit bestehenden Routinen wird angeboten, mit ihren sowieso stattfindenden Fahrten etwas Geld einzunehmen. Und Mitfahrer finden leichter eine Fahrt, wenn sie beispielweise zu Beginn des Berufsverkehrs in Ballungszentren wollen und so von einem Pendler mitgenommen werden können.

Problematisch könnte ein Fokus auf diese Kategorie werden, wenn die Teilnehmer so in verschiedene Klassen eingeteilt werden und sich so etwa nur Fahrer mit bestehenden Routinen anmelden.

Zusätzliches Interesse (role performance)

Hier werden Leute in Fahrgemeinschaften zusammengebracht, die zusätzliche Interessen haben. [Ozenc et al., 2011] nennt als Beispiel ein System, welches die Datingvorlieben der Teilnehmer vergleicht und entsprechend die Fahrgemeinschaften zusammenstellt.

Die Eingrenzung der Zielgruppe dürfte hier ein Problem sein. Wer einfach nur an einen Ort gelangen will, möchte sich auf ein solches System wohl eher nicht einlassen. Und eine Autofahrt bietet wenig Möglichkeiten, bei Nichtgefallen das Date abzubrechen (was äquivalent auch für andere denkbaren Szenarien gilt).

Event

Fahrgemeinschaften mit Menschen, die zum gleichen Event gelangen wollen.

Das ist vergleichbar zur vorherigen Kategorie, sowohl beim Nutzungsszenario als auch bei den Nachteilen. Der zusätzliche Vorteil liegt in der Möglichkeit, Leute kennenzulernen, mit denen man direkt an der Veranstaltung teilnehmen kann, und dass so eventuell leichter die Rückfahrt ebenfalls gemeinsam bewältigt werden kann.

Doch die Unzuverlässigkeit, ob die Mitfahrer dafür geeignet sind, und die Eingrenzung der Zielgruppe, bleiben als Nachteile bestehen.

Bestehende Strukturen

Solche Mitfahrzentralen bauen auf bestehende Strukturen. Beispielsweise könnte versucht werden, den bestehenden Freundschaftsgraphen aus Facebook zu benutzen, um Mitfahrer und Fahrer zusammenzubringen.

Für Nutzer hat dies den Vorteil, dass sie direkt Kontakte zur Verfügung haben und einfacher mit auf der anderen Plattform markierten Bekannten kooperieren können.

Zwei Nachteile ergeben sich: Für die Mitfahrzentrale entsteht - je nach Umsetzung - eine Abhängigkeit von der Plattform, die den Freundschaftsgraphen liefern soll. Und die Nutzer könnte es stören, wenn der Freundschaftsgraph übertragen

wird, obwohl für diese neue Plattform andere Kontakte ausgewählt worden wären - beispielsweise nur solche, die in der Nähe wohnen.

Zusätzlich kann unterschieden werden zwischen Mitfahrzentralen, die einen lokalen Bezug haben - beispielsweise nur eine bestimmte Stadt ansprechen - oder unbeschränkt beziehungsweise zumindest landesweit funktionieren.

4 Anforderungen an das Empfehlungssystem

Da das Empfehlungssystem nicht als komplett eigenständiges Konzept entwickelt wurde, war eine Vielzahl an Anforderungen zu beachten. Inwiefern diese berücksichtigt wurden, wird in Abschnitt 5 beschrieben.

4.1 Anforderungen durch Akzeptanzfaktoren

In Abschnitt 2 wurde eine Vielzahl an Arbeiten vorgestellt, die sich mit Empfehlungssystemen im Allgemeinen, aber auch speziell mit der Akzeptanz solcher Systeme beschäftigen. Dazu kommen Überlegungen zu Datenschutz beziehungsweise generell zum Umgang mit personenbezogenen Informationen, da dies als für dieses System relevant erachtet wurde, sowie Arbeiten zu Technikakzeptanz. Aus diesem Abschnitt ergeben sich daher eine Reihe von Anforderungen, die beim Design eines Empfehlungssystems beachtet werden sollten.

Die wahrgenommene Nützlichkeit (PU) und die wahrgenommene Einfachheit der Nutzung (PEOU) sind in Modellen wie TAM die bestimmenden Faktoren, ob ein beliebiges System von seiner Zielgruppe genutzt werden wird (Davis [1989]). Daraus leitet sich die Anforderung ab, ein ebenso nützliches wie einfach bedienbares System zu gestalten. Dies war auch Kern der Empfehlungen von [Jones and Pu, 2007], die Einfachheit des Interfaces mit einem minimalen Nutzungsaufwand zu verbinden. Das Ziel, Nutzungsbarrieren zu vermeiden, wird zusätzlich durch [Porter and Donthu, 2006] gestützt.

Gleichzeitig sollte von den Teilnehmern nicht verlangt werden, sich direkt auf das Empfehlungssystem verlassen zu müssen. [Cramer et al., 2009] zeigte, dass Vertrauen in die korrekte Funktionsweise und in das korrekte Verständnis des Systems nötig ist, bevor Nutzer für sie kritische Aktionen von einem System erledigen lassen wollen. Demzufolge sollte das System als Awareness schaffendes oder als zusätzliches Element des Gesamtsystems bestehen, das zunächst keine kritische Aufgabe autark übernimmt.

Solches Vertrauen und generell die Akzeptanz eines Empfehlungssystems kann durch das Vorhandensein von Erklärungen geschaffen beziehungsweise verbessert werden (Cramer et al. [2008], Zanker [2012], Hu and Pu [2009a]). Dabei sollten die Erklärungen deutlich machen, warum gerade zu einem bestimmten Zeitpunkt eine Erklärung gezeigt wird (Lim et al. [2009]).

Als Anforderung kann auch der beschriebene Wandel von Empfehlungssystemen weg von der präzisen Vorhersagefunktion verstanden werden. Denn statt einer präzisen Vorhersage könnte auch Nutzeranforderung sein, mithilfe des Empfehlungssystems aus einer Reihe von Vorschlägen auszuwählen, was eine Diversität der Empfehlungen benötigt.

Doch hängt das Ausmaß dieser Anforderung speziell von den Nutzern ab und kann daher vorher schwer antizipiert werden (Shani and Gunawardana [2011], Chen et al. [2013]).

Die Stabilität der Empfehlungen ist eine weitere Anforderung, soll sie doch das Vertrauen der Nutzer in das Empfehlungssystem stärken (Adomavicius and Zhang [2012]).

Empfehlungen sollten auf dem Netzwerk der Nutzer basieren (Guy et al. [2009]). Statt also einfach einen Fahrer aufgrund gemeinsamer Interessen zu empfehlen wäre ein System zu bevorzugen, das einen Fahrer oder Mitfahrer aufgrund gemeinsamer Bekannter oder idealerweise aufgrund einer bestehenden Bekanntschaft empfiehlt.

Zusätzlich zeigt [Patrick et al., 2005], dass die negative Konsequenzen für das Gesamtsystem gravierend sein könnten, wenn ein Empfehlungssystem als grundsätzlich negativ wahrgenommen würde und deshalb bei Implementierung eines solchen das Vertrauen in die Plattform schwinden würde. Diese Herausforderung wird erschwert durch den Umstand, dass in einem IT-System Vertrauen anders funktioniert als in einem rein zwischenmenschlichen Umfeld. Gleichzeitig sind die Nutzer der Plattform regelmäßig mit Mitarbeitern des Projekts in Kontakt, also spielt auch dieses Vertrauen eine mögliche Rolle in der Bewertung der Plattform und des Empfehlungssystem. Daher folgt daraus die weitere, nicht-funktionale Anforderung, dass die Nutzer das Empfehlungssystem nicht als vertrauensminimierendes System wahrnehmen dürfen. Dies wird untermauert durch ?, worin beschrieben wird, wie leicht soziale Interaktionen die Nutzung und Bewertung eines Empfehlungssystems beeinflussen.

Datenschutzbedenken auf Seiten der Nutzer sollten minimiert werden. Dies folgt aus [Hu and Pu, 2010], da dort das System scheinbar aufgrund der Datenschutzbedenken der Nutzer kaum akzeptiert wurde. Dazu kommt der Wunsch der Nutzer, ihre Außendarstellung zu steuern (Dearden and Watts [2004]) Eine Empfehlung von Fahrern könnte von diesen als Beeinflussung dieses Selbstbildes verstanden werden, was vermieden werden sollte.

4.2 Anforderungen durch das Projekt

Eine Vielzahl an Anforderungen an das Empfehlungssystem wurde durch das S-Mobil-100-Projekt⁶ vorgegeben.

Die Zielgruppe des Projektes war eng umrissen. Für das Projekt gesucht wurden Teilnehmer aus der Region Siegen-Wittgenstein, die älter als 60 Jahre sind. Denn das Ziel der Plattform ist, für Menschen ab diesem Alter eine altersgerechte und auch generationsübergreifende Mobilitätsplattform zu entwickeln. Entsprechend sollte auch das Empfehlungssystem für diese Zielgruppe zugeschnitten sein.

Desweiteren verfolgt das Projekt das Ziel, diese Plattform bei gemeinsamer Datenbasis auf drei Geräteklassen anzupassen: Auf Smartphones via einer App, auf PCs via einer Webseite, dazu auf internetfähigen Fernsehern. Zum Zeitpunkt dieser Arbeit war die Webseite im Fokus der Entwicklung. Dennoch sollte das Empfehlungssystem potentiell auf allen diesen Geräten nutzbar sein, wobei der erste Test des Systems auf der Website erfolgen sollte.

⁶Beschrieben auf <http://www.sehr-mobil.de/>

Außerdem beeinflusste die Zielgruppenangepasstheit der Plattform die Form des Empfehlungssystems. So wurde aufgrund von Bedenken, dass dies für die Zielgruppe unangemessen sei, darauf verzichtet, den Sozialen Graph der Nutzer abbildbar zu machen, es gibt also keine Freundschaftsfunktion. Ebensowenig sollte es geschlossene und moderierbare Gruppen geben, denen die Nutzer beitreten können, um keinen Aufwand auf Seiten der Nutzer zu verursachen. Stattdessen waren die Gruppen eine lose Selbstzuweisung von Tags. Eine Bewertungsfunktion von Fahrten und Fahrern wurde aufgrund der befürchteten negativen sozialen Implikationen, insbesondere bei der kleinen Zahl von Testnutzern, ebenfalls nicht implementiert.

Bewertungsfunktion und Sozialer Graph sind Elemente, die auf einem Sozialen Netzwerk (Sozialer Graph) und auf E-Commerce-Plattformen (Bewertungsfunktion) Kernpfeiler eines Empfehlungssystem sein können. Es war also Anforderung, diese Einschränkungen zu umgehen. Für ein Empfehlungssystem blieben also die Profilinformatoren der Nutzer inklusive ihrer Selbstkategorisierung.

4.3 Anforderungen durch die Nutzer

Zu den Anforderungen durch das Projekt kamen die Anforderungen der Nutzer.

Zu Beginn des S-Mobil-100-Projekts wurde von Mitarbeitern des Projekts mit den Teilnehmern vorbereitende Interviews geführt. Bei diesen Interviews ging es nicht speziell um Empfehlungen, sondern um den Hintergrund der Teilnehmer, allgemeine Anforderungen an die Plattform und darum, ihre Wünsche und Fragen aufzunehmen und zu beantworten.

In diesen Interviews wurden auch Fragen gestellt, die Rückschlüsse über die Anforderungen an ein Empfehlungssystem zulassen. So wurde beispielweise gefragt, mit wem üblicherweise bisher gemeinsam gefahren wird. Oder ob man auch Fremde mitnehmen würde. Ein Empfehlungssystem, das entsprechend diesen Vorstellungen Empfehlungen ausspricht, müsste den Erwartungen der Teilnehmer an ein solches System entsprechen - und wenn nicht, wäre das eines der Ergebnisse dieser Arbeit.

Wenig überraschend war, dass die Teilnehmer natürlich schon ihr eigenes Netzwerk nutzen. Fast alle berichteten von Fahrgemeinschaften mit Familie, Freunden und Bekannten.

Erste Erkenntnis aus der Analyse der Interviews war, dass viele der Teilnehmer bereit sind, auch ihnen im Grunde unbekannt Personen mitzunehmen, wenn sie sich ihnen durch irgendeine Form der Gemeinschaft verbunden fühlen. Und ebenso sich auch von ihnen mitnehmen lassen würden (gestärkt wurde diese Vermutung durch [Guy et al., 2009] insofern, als dass Empfehlungen basierend auf dem Social Graph generell vorteilhaft sein können). Die Form der Gemeinschaft kann das Leben in der gemeinsamen Nachbarschaft sein, aber mehr noch die gemeinsame Teilnahme an einem Gruppentreffen, wie der Gruppe *Alter Aktiv* oder der Kirchengemeinde.

Hier ein Beispiel eines Teilnehmers, der erst verneinte, von Fremden aus der Nachbarschaft sich mitnehmen zu lassen, um dann zu erzählen, dass genau dies passiert ist:

- 1: Könnten Sie sich dann vorstellen bei anderen Personen aus dem Ort – ...
– mitzufahren, die auch dort hinfahren?

T9: Also ich kenne die Leute hier noch nicht so, ich meine. . .

2: Ist es denn eine Voraussetzung dafür, dass Sie die kennen? Angenommen jetzt würde zufällig jemand anhalten und sagen, ich kann Sie auch mitnehmen bis runter zum

T9: Nein, nein. Dann würde ich nicht mitfahren [Lachen] wenn da jemand anhält und sagt ich könne mitfahren. Nein [Lachen] Aber man muss halt dann auch mal – ich meine mich hat dann neulich die Tochter von der Nachbar Leuten unten in der Stadt eingeladen mit hochzufahren.

. . . .

T9: Die kannte ich auch nicht . . . das sind Polen . . . aber sie hatte mich wohl schon gesehen und gesagt „na Sie wollen doch auch da hoch fahren – gehen“ Ich sage: „ja ich will auch da hoch“ „dann können Sie auch bei uns mitfahren“.

2: Sind Sie dann auch mitgefahren?

T9: Dann bin ich auch mitgefahren. Und sie haben mich bis hierher gebracht.

Es gab auch Teilnehmer, die deutlich sagten, auch ihnen Unbekannte mitzunehmen, doch blieb das die Ausnahme. Wenn der Unbekannte doch mitgenommen werden sollte, müsste über eine gemeinsame Bekanntschaft eine gemeinsame Basis bestehen. Alternativ die zaghafte Überlegung, man könnte ja vorher Kontakt aufnehmen. Es lässt sich vermuten, dass bei einem unbekanntem Fahrer durch dessen Machtposition dieses Zögern noch ausgeprägter wäre.

Zweites wichtiges Thema war der Wunsch, anderen nicht zur Last zu fallen und die eigene Souveränität zu bewahren. Dies unterschied sich bei den Teilnehmern jedoch sehr stark, wurde mal kaum, dafür andermal sehr deutlich gemacht. Zum einen wurde dafür die Last verteilt: Ein Teilnehmer berichtete beispielsweise davon, gelegentlich entweder Bekannte zu fragen oder den Weg doch zu laufen, anstatt schon wieder die Tochter zu bitten, ihn zu fahren. Diese brauche ja auch ihren Freiraum. Zum anderen wurde auch deutlich, dass die Teilnehmer zwar andere mitnehmen würden, aber nur ungern andere nach einer Mitfahrgelegenheit fragen würden, wenn sie selbst eine bräuchten. Ein autobesitzender Teilnehmer formulierte das so:

1: Können Sie sich denn auch vorstellen in einer Gruppe von Personen so eine Anfrage zu schicken. Also zum Beispiel möchte in die Siegener Innenstadt oder wo auch immer hin. Fährt jemand?

S: Da müsste ich mich erst dran gewöhnen. Das ist so - im Moment eher schwierig. Also ich würde da eher so sagen, ich muss das Ganze, ich muss das ganze so planen, ok gut nachschauen fährt da zufällig jemand ja aber

1: Also wenn eh jemand fährt dann nachzufragen

S: Mitfahren, ja. Ob ich da nun. Also ich [unverständlich 1 Wort] weiß ich nicht aber kann sein, dass das

1: Ja das ist ungewohnt.

S: Ja, das ist ein absolut ungewohnt, an so einen Gedanken müsste man sich erstmals daran gewöhnen.

In anderen Fällen war es aber schon gängige Praxis, mitgenommen zu werden, wieder insbesondere bei Fahrten zu einer gemeinsamen Gruppenveranstaltung:

T9: Aber könnten Sie sich das vorstellen, dass das etwas verändert, wenn Leute ihre Routen einstellen, dass sich dann die Hemmschwelle verändert Leute zu kontaktieren ob man mitfährt?

P: Also ich habe keine Probleme. Von der Gemeinde aus oder so nehmen sie mich ja auch manchmal mit. Die FEG ist ja da ganz hinten am anderen Ende der Stadt. Da fahre ich dann öfter mal mit ..., die hier unten den Laden haben. Die bringen mich dann meistens bis hier her oder auch andere Gemeindemitglieder.

Wie oben zusammengefasst findet sich das in den Interviews häufiger, dass bereits Fahrgemeinschaften von Gruppen zu gemeinsamen Veranstaltungen bestehen. Zwar lässt sich herauslesen, dass manche Teilnehmer generell Fahrer sind, aber in anderen Konstellationen wird sich abgewechselt:

1: Okay. Ja gibt es denn sonst so Situationen, wo man äähm, also wo sie mitfahren, oder wo sie andere mitnehmen?

T12: Ja, wenn wir nach Weidenau zum, wenn ich nach Weidenau fahre, zum Turnen. Dann ääh wechseln wir uns ab

1: Ah-ja

T12: Mal, ja. Das geht dann im Rund und dann ist, bin ich alle 14 Tage dran

Zwar wurde von solchen Beinahe-Automatismen berichtet, aber generell erfordert die Fahrgemeinschaft den Schilderungen zufolge fast immer eine zusätzliche konkrete Absprache für genau den Anlass. Die Kommunikation dafür erfolgte bisher per Telefon, im direkten Gespräch, oder wurde nicht weiter beschrieben. T02 beschrieb eine ausgeprägte Fahrpraxis:

1: Ehm wie sprechen Sie sich denn mit Frau ... eh dann ab wenn Sie oder eh ja also wenn Sie gemeinsame Unternehmungen auch hier im Siegerland haben?

T02: Joa dahingehend ist ganz klar eh sagen wir mal wie sieht es aus, haste Zeit, wollen wir wat weiß ich Sonntag ...

1: Gehen Sie dann runter? Oder eh rufen Sie an?

T02: Nö, ruf ich an. Ja.

T02: Also wie gesagt, eh wir gehen auch gegenseitig und das muss auch net immer so abgesprochen sein, ja aber so ein jemand so, ja jeden Tag überfallmäßig klingelt, ach hier weißte wat ich muss da wat erzählen. Das ist es nicht, ja. Auch da ist eine gewisse sagen wir mal Distanz, nein also wie gesagt ehm, nicht auf auf die Pelle rücken, weder Sie bei mir noch ich bei ihr, ja. Es ist ein eh sehr angenehmes Verhältnis, ja eh was aber auch ne gewisse, ne gewissen jeden einen gewissen Freiraum gibt. Ja. Sie hat noch Bekannte, ich, wie gesagt, und sagen eh wir auch wenn wir so stundenlang aufeinander hocken oder stundenlang Telefongespräche, mach ich sowieso nicht, ja. Also so ein Gelaber, ja ist ja, gar nix.

1: Ja und fahren Sie dann auch zum Beispiel zusammen zur Tafel oder so?

T02: Ja immer wechselnd, ich bin vorherige Woche gefahren, diese Woche fährt Sie, also wie gesagt, das geht im Wechsel. Also fährt immer nur einer, ist ganz klar

Hier findet sich auch das Abwechseln, wahrscheinlich als Versuch, die Last zu verteilen - oder auch, um die Rolle des Fahrers nicht einer Person zu überlassen. Und man erkennt den Versuch, gleichzeitig Distanz und Autonomie zu bewahren, indem die Telefongespräche auf kurze Absprachen beschränkt werden.

Doch ist die Bekanntschaft durch Nachbarschaft vorhanden und das gemeinsame Ziel offensichtlich, kann eine Fahrgemeinschaft auch spontan entstehen:

1: Also das heißt, wenn sie jetzt also zum Beispiel wüssten, dass eh der Herr. der Herr eh oder die Frau ehm ja in ihrer Gegend wohnt, also das wüssten sie dann mit Sicherheit, dann wäre das zum Beispiel auch so ne Vertrauens, dass würde so Vertrauen stiften.

T02: Ja, also das was hier im Umfeld ist, die Leute kenne ich.

1/2: mhm.

T02: Wenn ich sie auch nicht alle persönlich, hier im Haus schon. Nicht alle per Name oder so ja oder so.

2: Aber wenn man so weiß, der wohnt da unten.

T02: Ich weiß das, genau. Da sehe ich die da unten eh oder wenn ich runterfahren will und die stehen hier an der Bushaltestelle ja dann frage ich auch, wollen sie runter in den Ort oder wollen sie.

2: Ach dann halten sie da an und fragen ja?

T02: Ja sicher. Ja ich weiß ja nicht ob der Bus so schnell kommt. Nö, halt ich an, mach die Scheibe runter, ich sag, ich fahr runter in den Ort, ne der bus kommt jetzt. Okay.

T02: Wenn er nicht kommt, mach ich die Türe auf, gut. Manche wollen zum Arzt, sind n bisschen knapp, sind vielleicht ganz froh

2: genau das.

T02: Wenn ich sie schnell dort unten absetzte, also wie gesagt, das eh das ist kein Problem, also wie gesagt mit ganz Fremden, haben sie ja gemerkt, n bisschen zögerlich, ich weiß es nicht, ich hab diese Situation so noch nie gehabt. Ich hab selbstverständlich Leute die bei der Tafel bei uns einkaufen, wenn ich die dann mit ihren Sachen unterwegs, wenn ich nach Hause fuhr, getroffen hab, dann hab ich gehalten hab sie mitgenommen.

Diese Praxis zu unterstützen sollte eine Anforderung zumindest an die S-Mobil-100-Plattform, wenn nicht sogar speziell an das Empfehlungssystem sein: Awareness zu schaffen, dass solche gemeinsame Fahrten aufgrund der gemeinsamen Strecke möglich sind. Aber gleichzeitig das soziale Umfeld zuerst anzusprechen, wenn das gewünscht ist, denn das war desöfteren ein klar formulierter Wunsch:

1: Ja, ähm ach so... ja vielleicht nochmal.. also, ja man hat dann ja bei der Plattform auch die Möglichkeiten zum Beispiel Angebote ähm einzustellen oder auch zu fragen, ob man mitgenommen werden könnte. Das wäre, wäre das dann für sie ein Unterschied, also wäre das dann was anderes, wenn zum Beispiel ihr Nachbar, sie sehen, dass ihr Nachbar äh nach Siegen fährt und sie möchten auch nach Siegen... wäre das dann was anderes ähm da, dem Nachbar Bescheid zu geben, dass man mitfährt? Als zum Beispiel die Kinder zu fragen ob die ähm... einen mitnehmen?

T12: Also... äh da muss ich sagen, da will ich lieber erst die Kinder fragen... äh als die Nachbarn... es gibt äh Leute, die sagen, oh wir haben ja eine tolle Nachbarschaft bei uns, da is, hilft einer dem anderen. Das gibt es schon. Nur, man will eigentlich nicht so ähm so gerne... ich weiß nicht, das ist in einem so drin, dass man einfach nicht, den Nachbarn nicht so gerne äh anspricht.

1: und äh wenn sie sehen würden, der fährt sowieso?

T12: Naja, erst weiß ich ja nicht, dass er nach Siegen fährt... wenn einer jetzt ins Auto steigt, weiß ich ja nicht... der kann da lang fahren oder da lang fahren ne? Oder nach Hilchenbach... ne? Ist ja, die Strecke geht nach Hilchenbach... unten diese Straße hier... und äh... also das ist wirklich... ähm ein kleines Problem, dass, dass man immer nicht gerne die Nachbarn anspricht, das höre ich von anderen auch.

1: Ka... kann ich auch gut nachvollziehen, also auch mh..

T12: Wenn man zufällig äh in eingekauft hat, man hat eine Tasche voll Zeug, und es ist auch gerade ein Nachbar in der Nähe... gerade ne? Ja, mein Gott, wie soll ich jetzt heimkommen mit den großen, mit der großen Tasche... dann, dann ist das was anderes, dann sagt man, willst du denn jetzt auch Heimfahren? Kann ich mal mitfahren... da ist man ja heilfroh, und da spricht man schon jemand an, aber von hier, wenn man noch keine Tasche voll hat... ne? sozusagen (lacht)

Gleichzeitig spielt die konkrete Situation, warum jemand mitgenommen werden sollte, eine Rolle bei der Bewertung:

T4: Also hier unsere Mieterin ist ja unmobil.

T6: Die kann, die kann, die ist, die hat nen ganz krummen Rücken, kann kaum noch Laufen, die hat aber ne gute Bekannte die sie mitnimmt.

T4: Die fährt sie, also das ist ja auch gut so, sind wir auch froh drüber und wenn sie irgendwohin muss, brauch sie nur zu sagen, fahren wir sie.

Wenn Bekanntschaft nicht bereits besteht, kann diese vielleicht im Vorfeld hergestellt werden. Dazu zeigten sich zumindest ein paar Teilnehmer bereit, auch wenn das teilweise zögerlicher klang als im Folgenden:

T3: Und dann könnte man ja vielleicht auch mal mit demjenigen oder der da bereit ist einen mitzunehmen, vielleicht mal telefonieren und nicht dadurch, erstmal Kontakt aufnehmen - das wär ja das äh das sagen wir mal das wichtigste, erstmal mit der Person äh mal Kontakt aufnimmt. Vielleicht ergibt sich ja da auch ne Interessengemeinschaft..xx alle viertel Jahr, gehen wir mal ins Apollo Theater oder was ne? aber jetzt sag ich mal wild, das wär mir nu aber ein bisschen (lacht)

Das könnte eine Rechtfertigung sein, Empfehlungen zuzulassen, bei denen sich das System der Bekanntschaft nicht sicher ist.

Selbstautonomie war schon ein Thema, und wiederholte sich. Bei manchen Arten von Fahrten würden sich Teilnehmer in eingeschränkt fühlen, wenn sie jemanden mitnehmen sollten. Es wird deutlich, dass Mitfahrgemeinschaften für diese Zielgruppe nur selten ad-hoc geschaffen werden können:

1: Wäre das für sie denn zum Beispiel eine Option mit Bekannten oder Freunden hier aus ... nach Siegen zu fahren zum Einkaufen zu fahren und dort ein Kaffee zu trinken?

I: Eher nein. Aus einem ganz einfachen Grund weil in den meisten Fällen, wenn ich schon darunter fahren, dann hat man bestimmte Ziele wo man hin will, die müssen nicht unbedingt für unseren täglichen Bedarf nach Siegen fahren wir nicht, dann sind das, dann haben wir ganz bestimmte Ziele, die ich in Siegen kaufen will, da stören mich Fremde dann bei und wenn ich zum Beispiel dran denke, ich brauche jetzt irgendwas für den Computer oder so, wobei das macht meistens mein Sohn, dann rufe ich den an, aber manchmal sagt der auch "Du fahr mal eben da runter, fahr in den und den Laden, besorg Dir dasüber das sind so, das sind so ganz bestimmte Sache, meistens sind das dann irgendwelche, wenn wir so was machen, dass es irgendwelche Sache sind, die hier fürs Haus wichtig sind. Nicht nur Baumarktsachen aber überwiegend Baumarktsachen. Und da stören mich andere, das ist kein, ich sage immer, das ist kein Genuss zum Einkaufen.

Das heißt aber nicht, dass schnelle, aber selbstständige Mobilität nicht geschätzt würde. Eine Teilnehmerin schlug beispielsweise ein System des öffentlichen Nahverkehrs

vor, nur mit für die Gegend angemesseneren kleinen Fahrzeugen, aber mit so häufigen Fahrten, dass Menschen immer mobil wären:

I: Wäre doch schön, wenn hier ständig so kleine Autos durch die Gegend flitzen würde, wissen Sie so alle zehn Minuten, hier und vielleicht für sechs Personen drin, die ständig durch die Gegend fahren und die Leute aufladen. Ja? Und man so überall mitfahren könnte zum Beispiel.

1: Hätten die denn konkrete Ziele, die kleinen Flitzer?

I: Das Zentrum - da wo etwas los ist wahrscheinlich. Menschen in Bewegung bringen, aus der Isolation holen.

Es zeigt sich, dass viele Aussagen den anfänglichen Eindruck stützen, dass Teilnehmer in ihrer bisherigen Praxis auf die bisherigen Bekanntschaften zurückgreifen. Das wäre auch nur in wenigen denkbaren Szenarien anders, beispielsweise bei spontanen Fahrgemeinschaften über bestehende Mitfahrzentralen, aber solche Praxen finden sich nicht. Doch wurde deutlich, dass die Teilnehmer sich auch genau das bisherige wünschen - denn Fremde einfach mitzunehmen wurde meistens als kritisch bewertet.

Gleichzeitig zeigt sich die Mitfahrpraxis als ausgeprägte soziale Praxis, bei der Elemente wie Distanzwahrung, Bewertung der Bekanntschaft, Bewahrung von Autonomie und Respekt gegenüber dem Lebensalltag des anderen, den man nicht unbedingt stören will, eine Rolle spielen. Gleichzeitig gibt es Teilnehmer, die nicht bei allen Fahrten Mitfahrer haben wollen; auch, weil die Vorstellung anzufinden ist, dass nach der Fahrt - wie bei einem Mitfahrssystem der Event-Kategorie - gemeinsam beispielsweise der Einkauf bewältigt wird, dies aber nicht immer gewünscht ist.

Es wäre daher von Vorteil, wenn sowohl die bestehende Routine als auch die bestehenden Strukturen beachtet werden könnten, sowohl vom Empfehlungssystem als auch von S-Mobil-100 selbst.

Insgesamt lassen sich aus den Interviews zwei konkrete Nutzeranforderungen an die Empfehlungen erschließen:

1. Statt einfach nur positiv bewertete Fahrer zu empfehlen, sollten solche empfohlen werden, die entweder in der Nähe wohnen oder Teil einer gemeinsamen Gruppe sind.
2. Die Empfehlungen sollten nicht so präsentiert werden, dass die Empfohlenen als Mitfahrtsuchende empfunden werden, um die sozialen Implikationen zu minimieren.

5 Beschreibung des Empfehlungssystems

Einige Unterschiede zu üblichen Sozialen Netzwerken, und auch Mitfahrzentralen, beeinflussten das Design des Empfehlungssystems. Zum einen war und ist kein Freundschaftssystem implementiert, S-Mobil-100 versucht nicht, die Freundschaftsbeziehungen auf der eigenen Plattform abzubilden. Außerdem ist bisher keine Bewertungsfunktion vorgesehen, sondern es wird sich stattdessen auf die sozialen Reputationsfaktoren innerhalb der kleinen Teilnehmergruppe verlassen.

Die Teilnehmer sind jedoch üblicherweise in Gruppen organisiert, die sich unabhängig dieser Plattform regelmäßig treffen. Dementsprechend können die Nutzer auf der Plattform in Gruppen eintreten und diese auch erstellen. Damit verbunden ist kein Gruppenmanagement, keine geschlossene Diskussionen oder ähnliches, sondern Gruppen auf der Plattform sind mehr eine selbstzugewiesene Kategorie, oder im Sinne einer Folksonomie eben ein Tag.

Über diese Gruppentags lassen sich nun Beziehungen zwischen den Nutzern ablesen und der erste Prototyp nutzt sie, um Empfehlungen auszusprechen. Um für einen Nutzer A Empfehlungen zu bestimmen sammelt der Prototyp eine Liste von Gruppen, in denen Nutzer A Mitglied bzw. getagged ist. Mit dieser Liste von Gruppen lässt sich eine Auswahl von Nutzern bestimmen, die mit Nutzer A verbunden ist. Diese können nun entweder direkt empfohlen werden, oder wiederum genutzt werden, um weitere Ressourcen zu finden: Zum einen Gruppen, in denen Nutzer A noch nicht ist, aber in denen mit ihm verbundene Nutzer sind, oder ebensolche Events.

Implementiert wurde also ein externes System mit vollem Zugriff auf die Datenbank von S-Mobil-100. Es handelt sich um eine Ruby/Sinatra Anwendung. Diese Anwendung hat im weitesten Sinne eine REST-API mit einer Funktion:

```
GET /recommendation
```

Diese erwartet das Vorhandensein des Parameters

```
userid
```

Für diesen spezifizierten Nutzer wird eine Empfehlung ausgesprochen. Oder eher drei: Es wird eine JSON-Objekt zurückgegeben, das drei Arrays beinhaltet, nämlich Nutzer-, Gruppen- und Event-IDs. Das Empfehlungssystem kümmert sich also in keiner Weise um die Präsentation der Empfehlungen, es definiert weder, was zusätzlich noch wo die Empfehlung angezeigt wird.

Das Empfehlungssystem sollte auf folgendem Datenbankschema arbeiten:

```
CREATE TABLE user_activity
(
  id bigint NOT NULL,
  version integer,
  connection_id bigint,
  event_id bigint,
  user_id bigint,
  CONSTRAINT user_activity_pkey PRIMARY KEY (id),
  CONSTRAINT fka04f4d033060ae90 FOREIGN KEY (user_id)
  REFERENCES users (id) MATCH SIMPLE
  ON UPDATE NO ACTION ON DELETE NO ACTION,
  CONSTRAINT fka04f4d03e718458a FOREIGN KEY (connection_id)
  REFERENCES smconnection (id) MATCH SIMPLE
  ON UPDATE NO ACTION ON DELETE NO ACTION,
  CONSTRAINT fka04f4d03fb45f584 FOREIGN KEY (event_id)
```

```
REFERENCES event (id) MATCH SIMPLE
ON UPDATE NO ACTION ON DELETE NO ACTION
)
```

```
CREATE TABLE event
(
id bigint NOT NULL,
address character varying(2048),
allday boolean NOT NULL,
description character varying(2048),
end_time timestamp without time zone,
ext_id integer NOT NULL,
start_time timestamp without time zone,
tag character varying(255) NOT NULL,
title character varying(255),
version integer,
CONSTRAINT event_pkey PRIMARY KEY (id)
)
```

```
CREATE TABLE users
(
id bigint NOT NULL,
birthday date,
email character varying(255) NOT NULL,
enabled boolean NOT NULL,
firstname character varying(255),
is_admin boolean,
is_testuser boolean,
lastname character varying(255) NOT NULL,
car_type integer,
number_of_seats integer,
pick_tolerance_meter integer,
pick_tolerance_time integer,
smoker boolean,
password character varying(255) NOT NULL,
phone character varying(255),
sex integer,
username character varying(255) NOT NULL,
version integer,
city character varying(255),
houenumber character varying(255),
street character varying(255),
postalcode character varying(255),
```

```
CONSTRAINT users_pkey PRIMARY KEY (id)
);
```

```
CREATE TABLE groups
(
id bigint NOT NULL,
name character varying(255),
description character varying(2048),
version integer,
CONSTRAINT groups_pkey PRIMARY KEY (id)
);
```

```
CREATE TABLE users_groups
(
users bigint NOT NULL,
groups bigint NOT NULL,
CONSTRAINT users_groups_pkey PRIMARY KEY (users, groups),
CONSTRAINT fkd034efeb3fcc4e89 FOREIGN KEY (users)
REFERENCES users (id) MATCH SIMPLE
ON UPDATE NO ACTION ON DELETE NO ACTION,
CONSTRAINT fkd034efeba10a925d FOREIGN KEY (groups)
REFERENCES groups (id) MATCH SIMPLE
ON UPDATE NO ACTION ON DELETE NO ACTION
);
```

Man erkennt die verbindende Funktion der Gruppentabelle über die Kombinationstabellen *users_groups* und *user_activity*.

Dieses naive Vorgehen führt zu einer Komplexität im Worst Case von $O(n^3)$, wenn vom einem Nutzer ausgehend alle Nutzer aus allen Gruppen alle Events empfohlen werden (auch wenn nur die ersten drei angezeigt würden). Doch da die Nutzeranzahl, wie oben beschrieben, gering ist, erwies sich das für den Prototypen nicht als problematisch. Außerdem können solche Empfehlungen sinnvoll gecached werden.

Das Design dieser Empfehlungsdarstellung lag in den Händen des Entwicklerteams der S-Mobil-Plattform und wurde lose abgesprochen. In der ersten Version des Prototyps werden nur die so empfohlenen Events in der Seitenleiste unter dem Titel *Interessantes* angezeigt (siehe Abbildung 15. Events können gebucht und für sie auf der Plattform eine Fahrt gesucht werden.

5.1 Diskussion der Anforderungen

Diese Implementation erfüllt die Anforderungen durch die Nutzer:

1. Indem statt Personen Veranstaltungen empfohlen werden, wird die soziale Implikation einer solchen Empfehlung minimiert. Kein Nutzer kann stolz oder pikiert

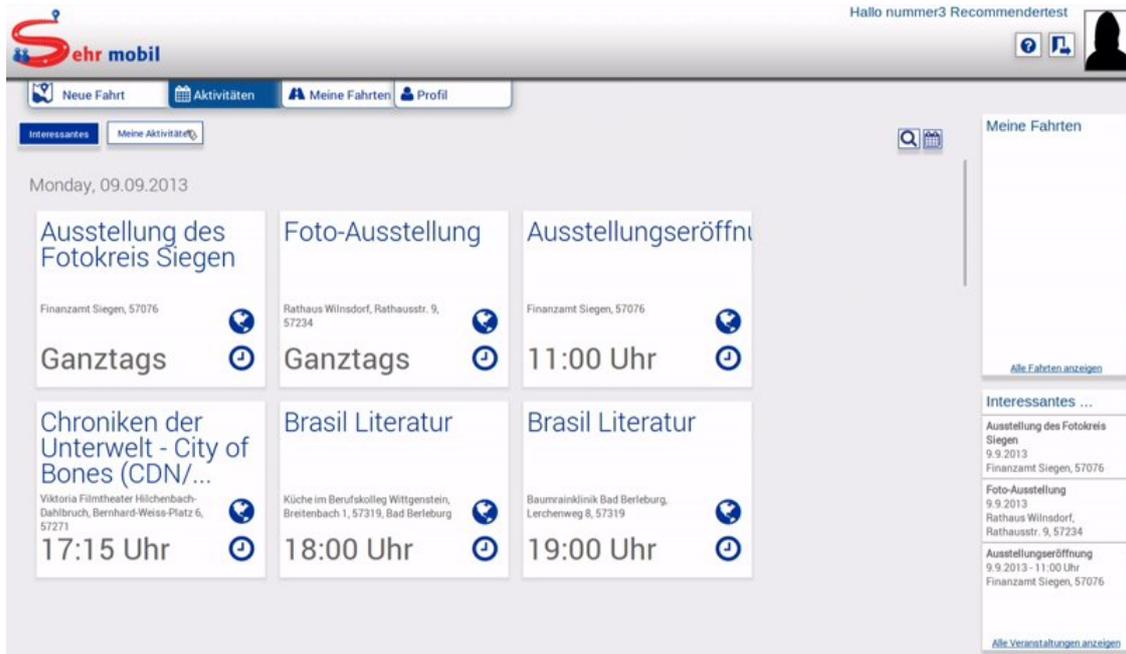


Abbildung 15: Das Empfehlungssystem ist unten rechts zu sehen, links füllt die normale Veranstaltungsliste die Seite.

sein, anderen empfohlen zu werden, und die Empfehlung kann nicht als Mitfahrge- such missinterpretiert werden. Eine solche implizite Hinführung zu einer sozialen Interaktion findet sich beispielsweise auch in [Svensson et al., 2001].

2. Impliziert werden dann doch Nutzer empfohlen, aber nur solche, die dem Nutzer schon bekannt sein dürften. Die Überlegung ist die: Veranstaltungen werden nur empfohlen, wenn mindestens ein anderer Nutzer die Veranstaltung gebucht hat, der mit dem Nutzer in einer gemeinsamen Gruppe ist. Sucht der Nutzer also nach einer Fahrt, besteht die Möglichkeit, diese beiden Nutzer in Kontakt zu bringen, und sie haben schon eine gemeinsame Basis, was mehrheitlich als notwendig angesehen wurde.

Gleichzeitig minimiert dieses System Datenschutzbedenken, da Daten von Anwen- dern nicht angezeigt werden, denn die Daten der Nutzer werden zur Empfehlung von Veranstaltungen genutzt, nicht von Nutzern. Ebenso wird ihre Selbstdarstellung nicht beeinflusst.

Problematisch sind die Erklärungen. Denn bei der gewählten Darstellungsform sind diese nicht vorgesehen. Es muss sich auch erst zeigen, ob die Empfehlungen als solche er- kannt werden. Das Problem mit den Empfehlungen liegt bei der für sie nötigen Preisgabe von Daten: Um wirklich zu erklären, warum eine Veranstaltung empfohlen wird, müsste eigentlich preisgegeben werden, welcher Nutzer in welcher Gruppe an dieser Veranstal- tung teilnimmt. Vielleicht wären auch Erklärungen mit weniger Gehalt ausreichend, aber

das müsste getestet werden. Gleichzeitig ist der Verzicht auf die Erklärungen im Hinblick auf die Forschungsfrage nicht unbedingt verkehrt: Wird ein System akzeptiert, das seine Akzeptanz nicht völlig perfekt beeinflusst, reicht das wahrscheinlich trotzdem, um diese zu im positiven oder negativen zu beantworten. Es könnte sogar helfen, wenn so eine Differenz zur Bewertung der Gesamtplattform offenbar wird.

Angesichts des verwendeten Vorgehens wäre es immerhin einfach, die Daten für solche Erklärungen auszugeben.

Die wahrgenommene Nützlichkeit muss evaluiert werden. Es scheint jedoch vertretbar, zu argumentieren, dass durch das simple Auflisten der Veranstaltungen die Anforderung erfüllt wurde, ein einfach bedienbares System ohne Nutzungsbarrieren zu gestalten.

Es wird von den Teilnehmern nicht verlangt, sich auf das System in einer kritischen Interaktion zu verlassen. Die Stabilität ist diskutierbar: Die empfohlenen Events müssen sich ohne neue Daten nicht ändern, andererseits ist der Platz für ihre Darstellung begrenzt und es könnte daher sinnvoll sein, sie häufig auszuwechseln.

Die Liste ist ein zusätzliches Element des Systems, die Awareness für bestimmte Veranstaltungen schafft; dies war eine der Anforderungen. Zusätzlich werden mehrere Veranstaltungen angezeigt und es ist einfach, weitere Veranstaltung zu durchstöbern. Die Anforderungen der Nutzer nach diversen Empfehlungen sollte erfüllt werden können.

Als externe JSON-API mit Datenbankzugriff eignet sich das System, um in alle möglichen Geräteklassen, in alle möglichen Darstellungsformen eingebunden zu werden. Die Darstellung wird nicht vorgegeben und es könnte eine Vielzahl weiterer Einsatzzwecke gefunden werden. Beispielsweise wäre es ein Leichtes, Veranstaltungen nicht nur an dieser Stelle, sondern zielgerichtet bei einer Veranstaltungssuche anzuzeigen, gegebenenfalls auch in völlig anderer Darstellungsform.

6 Beschreibung der Evaluation

Für die Evaluation des Prototypen wurde ein Participatory Design Workshop durchgeführt. Normalerweise dienen solche als Entwicklungsinstrument im frühen Entwicklungsprozess, doch sind sie in diesem iterativen Model auch später noch anwendbar (Wulf et al. [2011]). In diesem Workshop sollte den Teilnehmern zuerst das Empfehlungssystem gezeigt und erklärt werden, damit eine informierte Rückmeldung über die Akzeptanz eines solches Systems gegeben werden kann. Abschließend wurde ein an TAM orientierter Fragebogen ausgefüllt, um die Akzeptanz der Empfehlungen zu quantifizieren.

Im folgenden wird der Ablaufplan wiedergegeben:

6.1 Workshopgruppenplan

6.1.1 Das Ziel des Workshops

- Feedback zum bestehenden Empfehlungssystem bekommen
- Welche Funktionen können sie sich vorstellen, wo sehen sie Grenzen?

6.1.2 Formales

2 Gruppentreffen sollten durchgeführt werden, auch um mit den Erfahrungen der ersten die zweite anzupassen, falls nötig.

- Teilnehmer: 3-5 Teilnehmer des S-Mobil-Projekts
- Dauer: 60 Minuten

6.1.3 Vorbereitung

Zwei Profile in der gleichen Gruppe. A wählt eine Reihe von Events aus. Bei B sind diese das dann in der Seitenleiste zu sehen. Im Idealfall ist eine bestimmte Präferenz erkennbar.

6.1.4 Ablauf

1. Vorstellungsphase (10 Minuten)

Am Anfang geht es darum, den Teilnehmern einen Einstieg zu geben. Indem sie sich mir vorstellen können, wird Achtung für ihren Hintergrund vermittelt. Dabei ist es nicht sehr wichtig, die Informationen über ihren Hintergrund neu zu erfragen, da diese Informationen über die Interviews bekannt sein sollten. Inwiefern das System bisher genutzt wurde ist eine wichtige Information, um den erforderlichen Informationsbedarf zu ermitteln.

Fragen

- Allgemeine Vorstellung
- Wie stark haben Sie smobil bisher bereits genutzt?
- Was benutzen Sie ansonsten für Technik?

2. Gegenüberstellung des Event-Menüs (unbeeinflusst vom Recommender) mit der Recommender-Leiste (10 Minuten)

Das Mentale Modell der Teilnehmer über das Systemverständnis kann helfen, das System so zu erklären, dass es mental erfasst wird. Gleichzeitig baut es einen Anknüpfungspunkt an den nächsten Schritt, in dem das eigentliche System erklärt wird. Auch könnten hier schon Äußerungen fallen, die auf die Akzeptanz des Systems schließen lassen ("Das liest ja wohl hoffentlich nicht meine Profildaten!")

Fragen

- Welche der beiden Veranstaltungslisten gefällt Ihnen besser?
- Was vermuten Sie, warum die unterschiedlich sind?

3. Erklärung der Vorgehensweise des Recommenders mit den Gruppen (10 Minuten)

Es wäre für die Teilnehmer wahrscheinlich unbefriedigend, ihnen keine Erklärung des Recommenders zu geben. Wird die Erklärung verstanden, lädt das auch zum (im Zweifel späteren) Ausprobieren des Systems ein - insbesondere im Hinblick auf folgende Frage: Sieht man die Events des Nachbarn? Außerdem schafft es die Grundlage, um über die Möglichkeiten zu diskutieren.

Fragen: -

4. Datenauswahl (20 Minuten)

Welche Daten für die Empfehlungen genutzt werden können, ohne dass die Nutzer sich daran stören, ist die andere Frage neben den verschiedenen möglichen Empfehlungsarten. Daher wird den Teilnehmern eine grafische Repräsentation der Daten in die Hand gegeben, in Form der Kategoriekästen im Profil. Diese sollen dann einsortiert werden in die Kategorien "nutzen" und "nicht nutzen"

5. Abschließend die Frage, wie sie die Empfehlungen finden (10 Minuten)

Die Akzeptanz des Systems zu erfragen ist der Ziel des Ganzen. Äußerungen zur Nützlichkeit und Einfachheit können Hinweise darauf geben (siehe TAM), aber auch klare direkte Bewertungen könnten hier erfolgen. Die Frage nach dem Empfehlen von Teilnehmern soll wahrgenommene Grenzen ausloten (es ist derzeit nicht geplant) - wird selbst das akzeptiert, sollte das System generell unkritisch sein. Wird diese Art Empfehlung abgelehnt, liegt zumindest die derzeitige Grenze unter dieser Schwelle.

Fragen

- Sind diese Empfehlungen einfach zu benutzen?
- Könnten Sie sich vorstellen, da auch Teilnehmer empfohlen zu sehen?

Am Ende wird den Teilnehmern ein TAM-Fragebogen, das Gegenstück zur Evaluation der Gesamtplattform bei Projektbeginn, gegeben.

6.2 Fragebogen

Am Ende des Workshops wurde ein an TAM angelegter Fragebogen ausgegeben, übernommen an den aus [Beil, 2013]. Damit sollte ein Vergleich mit der vorherigen Bewertung der Gesamtplattform durch die Teilnehmer ermöglicht werden, bei gleichzeitiger Quantifizierung der Akzeptanz des Empfehlungssystems. Diese Bewertungsmöglichkeiten wurden mit einer Likert-Skala von 1 bis 5 gestellt:

1. Das Nutzen der Empfehlungen könnte meinen Alltag erleichtern
2. Ich fände die Empfehlungen nützlich
3. Ich denke, die Benutzung der Empfehlungen wird leicht sein
4. Ich glaube, ich könnte die Empfehlungen auch nutzen, wenn keiner da wäre, der mir sagt, was ich tun muss

5. Ich denke, beim Umgang mit den Empfehlungen würde ich mich unwohl fühlen
6. Ich denke, ich habe die notwendigen Fähigkeiten, um die Empfehlungen zu nutzen
7. Ich denke, die Empfehlungen zu nutzen würde mir Spaß machen
8. Wenn ich Zugang zu den Empfehlungen hätte, würde ich sie nutzen

6.3 Ablauf

Dies wurde mit zwei Gruppen durchgeführt. Die erste Gruppe hatte 9 Teilnehmer, wovon 8 den Fragebogen ausfüllten. Die zweite hatte 5 Teilnehmer. Alle Teilnehmer waren bereits geschult im Umgang mit Smartphones, damit auch generell im Umgang mit modernen IT-Systemen und Webseiten. Im Vorfeld des Workshops wurde eine einstündige Einführung in die Funktionsweise der Webseite selbst gegeben und dann überleitet in den Workshop.

7 Auswertung

Die Auswertung der Evaluation erfolgt in zwei Schritten: Zuerst wird eine Beschreibung der qualitativen Daten gegeben. Es wird also beschrieben, wie die Teilnehmer reagierten, welche Fragen sie stellten und was ihre formulierte Einschätzung des Empfehlungssystem war, insbesondere wo sie Probleme sahen. Im zweiten Abschnitt wird der von den Teilnehmern ausgefüllte Fragebogen ausgewertet, und sowohl mit den Aussagen im Workshop als auch mit den Technikazpetanzdaten der Studie aus Heidelberg verglichen. In beiden Abschnitten wird versucht werden, Diskrepanzen und Parallelen mit früheren Aussagen und Daten der Teilnehmer zu finden: wenn ein Teilnehmer beispielsweise sich im Vorfeld der Studie selbst als besonders technikfreundlich einschätzt und das Empfehlungssystem besonders unkritisch bewertete, soll das erwähnt werden.

7.1 Workshop

Der Rahmen des Workshops war ein Treffen, das die Teilnehmer des S-Mobil-Projekts regelmäßig durchführen, um dort Unterstützung von den Projektmitarbeitern zu bekommen und über das weitere Vorgehen informiert zu werden. Bei beiden Treffen waren neben mir Mitarbeiter des Projekts anwesend, die den Teilnehmern bekannt waren.

7.1.1 Gruppe 1

Das Ergebnis der ersten Gruppe war nicht eindeutig. Auf der einen Seite wurde die Idee eines Empfehlungssystems überraschend deutlich begrüßt. Auch wurde die Schlichtheit der Implementation insofern kritisiert, als dass die simple Auflistung der Events zu wenig sei, also für transparentere sowie tiefergehende Empfehlungen plädiert (dies bestätigt tendentiell die Ergebnisse von [Hu and Pu, 2009a]). In der Frage der verwendeten Daten wurden größtenteils die meisten im Profil repräsentierten Daten zur Verwendung

freigegeben, mit zwei Einschränkungen: Das Profilbild sowie die Heimatadresse wurden mehrheitlich als kritisch eingestuft. Hier dachten die Teilnehmer aber einen Schritt weiter: Es ging ihnen nicht primär um die Verwendung durch das Empfehlungssystem bei der Datenverarbeitung, Fragen in die Richtung wurde eher achselzuckend begegnet. Es ging ihnen um die Darstellung der Empfehlungen, also welche Informationen ein Empfehlungssystem anzeigen sollte. Sie erkannten schließlich den Konflikt zwischen ihrem Wunsch nach Privatsphäre und dem Wunsch nach einer möglichst vollumfänglichen Information über andere Teilnehmer der Mobilitätsplattform und waren dadurch im Kern der Designabwägung angelangt, ohne diesen Konflikt letztendlich auflösen zu können.

Die Teilnehmer bezeichneten sich selbst überwiegend als Technikern. Eine Teilnehmerin (T11) beschrieb sogar Technikangst, in der Form einer inneren Distanz zur Technik und Angst vor ihren Konsequenzen, im Hinblick auf Sicherheit. Nur zwei der Teilnehmer (T14, T7) zeichneten ein Selbstbild von ausgesprochener Techniknähe. Gleichzeitig und im Gegensatz zur Selbstbeschreibung beschrieben jedoch noch drei weitere Teilnehmer (T9, T6, T13), dass sie durchaus über technisches Hintergrundwissen und Erfahrung im Umgang mit Computersystemen verfügen. Beachtet werden muss, dass in [Beil, 2013] ermittelt wurde, dass die Teilnehmer des Projektes generell über eine überdurchschnittliche Technikerfahrung verfügen, und sie im Laufe des Projektes sowohl im Umgang mit Computern, Smartphones und Apps geschult wurden. Eine Verallgemeinerung der Erkenntnisse ist demnach nur ganz schwer möglich.

Bei der Vorstellung des Empfehlungssystems wurden direkt Verständnisschwierigkeiten im Hinblick auf die Transparenz der Empfehlung geäußert. Das System wurde demonstriert, indem in einem Video ein Nutzer Events auswählte, ein zweiter diese dann empfohlen bekam, da beide die einzigen Mitglieder der gleichen Gruppe waren. Bei der Erklärung wurde dann vermutet, dass hier nicht allgemein etwas empfohlen wird, sondern die Events eines spezifischen Nutzers betrachtet werden, eben die von dem, der zuvor im Video gezeigt wurde. Die anderen Teilnehmer korrigierten diesen Eindruck. Dies ist jedoch ein Hinweis darauf, wie das mentale Modell des Nutzers für Empfehlungen war: Dass sie angezeigt werden, indem die Events von einem bestimmten Nutzer angezeigt werden, und diese auch dem Nutzer zugeordnet werden können.

Auf die Erklärung, dass hier ein Empfehlungssystem am Werke ist, fiel diese Mischung aus Aussage und Nachfrage:

T8: Was am häufigsten ausgewählt wird, das steht unten?

Dies ist eine valide Beschreibung eines simplen Empfehlungssystems, wieder ein Hinweis auf die Erwartungshaltung gegeben der intransparenten Darstellung.

Konfrontiert mit der Frage, welche Systeme sie besser finden würden - eine Liste erstellt aus einem Empfehlungssystem, maßgeschneidert für sie, oder einer Standardliste - fiel mehrfach die Antwort, dass man beides bräuchte. Nur die empfohlenen zu sehen, weil man da einmal angegeben habe, sich für Sprachen zu interessieren, wurde als zu einseitig kritisiert - man wollte auch breiter gefächerte Informationen sehen.

Allerdings wurden später selbständig alternative Szenarien beschrieben, in denen ein Empfehlungssystem nützlich sein könnte, insbesondere bei der Fahrersuche. Wenn für eine konkrete Fahrt eine Liste an Fahrern vorgeschlagen werden würde, dann könnte eine

Empfehlungssystem dieses ja so sortieren, dass der vom System am besten eingestufte Fahrer oben wäre.

Aber nur maßgeschneiderte Listen anzuzeigen sei auch deswegen problematisch, da die Einordnung in Gruppen nicht vollständig sein könne. Einer Teilnehmerin missfiel diese Idee auch generell, sich über diese Gruppenzuordnung zu kategorisieren. Dies war auch die Teilnehmerin, die anfangs in ihrer Vorstellung Sicherheitsbedenken gegen Technik formulierte. Die Vermutung liegt nahe, dass zumindest in diesem Fall der Widerstand gegen die Kategorisierung mit dem Misstrauen gegenüber Computersystemen zusammenhängt.

Hinzu kam die Einschränkung, dass Gruppenzugehörigkeit nur im lokalen Umfeld interessant sei - jemand in Frankfurt, der sich in die gleiche Gruppe eintrüge, sei für die Fahrten im Umkreis ja nicht von Belang. Dies passt zu meiner aus den Interviews gewonnenen Erkenntnis, dass für ein Empfehlungssystem von Fahrgemeinschaften die lokale Nähe der Teilnehmer für diese Zielgruppe wichtig ist, als Alternative oder zusätzlich zur Zusammengehörigkeit zu Gruppen, die sich regelmäßig treffen. Dementsprechend wünschte sich diese Teilnehmerin eine Beschränkung auf Menschen im Umkreis einer bestimmten Postleitzahl.

Die Empfehlungsliste dagegen habe einen anderen Einsatzzweck als die allgemeine Liste:

T7: Ich würde mich gerne mit Leuten vernetzen und dann zusammen gehen, oder ich weiß ja, die gehen auch da hin die seh ich da, die treff ich dann.

T8: Oder vielleicht eine andere Veranstaltung wird interessant dadurch, weil ich weiß, Frau ... geht da auch hin. Interessanter, als wenn ich alleine gehen würde, so rum.

Um solche Szenarien zu unterstützen müsste der Prototyp wesentlich mehr Informationen anzeigen, als er es bisher tat. Entsprechend wurde das gefordert und diskutiert. Denn hier wurden auch Probleme gesehen:

T8: Es kann ja auch sein, ich will mal ganz alleine irgendwo hin und will überhaupt nicht, dass Frau ... das weiß. Was mache ich dann?

Gelächter

T14: Dann muss ich das ja nicht eintragen, wenn ich das nicht will.

T11: Ja, das kommt ja automatisch, wenn ich das jetzt richtig verstehe.

T8: Also, genau, kann ich das noch filtern, wer das sieht, das wäre noch die Frage?

Der Prototyp konnte das nicht. Der Wunsch nach Privatsphäre und nach Filtereinstellungen muss als Featurewunsch für die nächste Iteration aufgefasst werden. Dass er so früh in der Diskussion aufkam, ohne dass dies vorher von Seiten der Gruppenleiter thematisiert worden war, zeigt, dass dies ein in den Augen der Teilnehmer wichtiger Aspekt ist. Vorschlag der Teilnehmer war, dass man dies doch pro Veranstaltung freischalten können sollte, ob die Teilnahme anderen angezeigt wird.

Die Frage nach Filterung und Datenschutz lieferte den passenden Moment für eine Provokation: Der Frage, ob man denn dort nicht auch Nutzer als gute Fahrer empfehlen können sollte. Dies wurde mit etwas Zögern verneint. Die Teilnehmer sahen hier eine Reihe von praktischen Problemen:

- Wie solle ein guter Fahrer bestimmt werden? Ein Vorschlag war, dass man dann Bewertungen einführen müsste. Das wurde jedoch als kritisch bewertet.

T9: Ähnlich wie bei Ebay, da kann ich das ja bewerten, dass ich dann nach einer abgeschlossenen Fahrt über drei Symbole, Gut, Mittel, die Bewertung zu setzen, sodass dann nachher das in dem Profil eben erscheint, also so viele positive Bewertungen.

T14: Oder einfach die Fahrten aufsummieren, mengenmäßig, wenn einer schon hundertmal jemanden mitgenommen hat, oder dreißigmal, ist er vermutlich besser, als wenn er das zum ersten oder zweiten mal macht.

...

T11: Also menschliche Bewertungen, das ist immer so ein gefährliches Thema ...

T13: ... Also ich finde das auch zu persönlich.

T12: Ja, möchte ich nicht haben.

Zustimmung in der Runde

Die Mitfahrer zu zählen sei auch deshalb kritisch, wurde später geäußert, weil das auch einfach das Ergebnis einer beliebten Strecke sein könnte, unabhängig vom Fahrer.

- Was gewinnt ein Fahrer dadurch? Eine solche Empfehlung könnte dazu führen, dass er öfter um Mitfahrgelegenheiten gebeten werden würde, als ihm lieb ist. Ein Fahrer selbst habe dadurch keinen Nutzen
- Wie bestimmt man einen guten Fahrer? Selbst wenn die Mitfahrer bewerten würden, gäbe es mehrere Kriterien für gut, wie das persönliche Verhalten dem Mitfahrer gegenüber oder einen sicheren Fahrstil und es gäbe noch einige weitere Kriterien. Eine Bewertung durch Teilnehmer sei deshalb schwierig und eventuell aussagegelos. Außerdem: Wie unterscheidet man zwischen Mitfahrern und Fahrern bei solchen etwaigen Bewertungen? Auch ein Mitfahrer könnte ja eine negative oder positive Bewertung verdienen, dies sei jedoch eine andere Ebene als die Bewertung eines Fahrers.

Verbunden mit den später klar formulierten Bedenken dagegen, Daten wie das Profilfoto anzuzeigen, wurde jedoch deutlich, dass es nicht nur um praktische Probleme ging. Die Teilnehmer empfanden es auch als unangenehme Vorstellung, in solcher Form präsentiert zu werden. Außerdem:

T9: Das würde ja auch im Umkehrschluss bedeuten, die die nicht empfohlen werden sind schlechte Fahrer ... oder nicht so gute.

T8: Das würde mich dann ja viel mehr interessieren!

Vor schlechten Fahrern gewarnt zu werden schien für die Teilnehmer nämlich ein Thema zu sein. Mehrere erzählten, in ihrem Umfeld Fahrer zu kennen, mit denen sie nicht gerne fahren, weil sie sich mit ihnen nicht sicher fühlten. Teilweise sei das diesen Fahrern auch bewusst. Deswegen sei auch der Wunsch da, das Alter von anderen Teilnehmern zu sehen, denn es wurde vermutet, dass besonders alte Fahrer normalerweise entsprechend unsicher seien.

Generell wurde die Nützlichkeit der Empfehlungen bejaht. An einer Stelle kam aber auch Widerspruch:

T7: Also wir telefonieren miteinander - ja hör mal, wir gehen am Sonntag da und da hin, kommt ihr auch?

Für diese Teilnehmerin wurde der wahrgenommene Nutzen der Empfehlungen also durch einen vorhandenen alternativen Informationsfluss minimiert. Doch generell wollten mehr Teilnehmer gerne sehen, wer zu welchen Veranstaltungen geht, um sich gegebenenfalls anschließen zu können und wahrscheinlich auch als Element sozialer Teilhabe, wobei das nicht explizit formuliert wurde.

Doch welche Daten genau sollten dafür genutzt und angezeigt werden (siehe Abbildung 16). Wie oben beschrieben argumentierten die Teilnehmer eher auf Ebene des Anzeigens, nicht auf der Datenverarbeitungsebene des Systems. Nur zwei der Teilnehmer stuften alle ihre Profilinformatoren als anzeigbar ein. Ansonsten gab es drei Datensätze, die generell kritisch waren: Der volle Name, das Profilfoto und die Heimatadresse. Die Heimatadresse wurde von allen anderen als kritisch eingestuft. Nur zwei weitere sahen keine Probleme im Teilen des Fotos, genauso beim Namen.

Zusätzlich als kritisch eingestuft wurden:

- Zweimal die private Telefonnummer
- Einmal die eigenen Interessen

Interessant im folgenden ist der Hintergrund der Teilnehmer, die bei keiner der Datenarten Probleme sahen, T7 und T8, und der Sonderfall T14.

T7 beschrieb, dass sie seit Jahren selbstständig ist und auf ihrem Auto einige Informationen zeigt, unter anderem ihre Telefonnummer. Dies habe nie zu Problemen geführt, sie habe nur selten darüber überhaupt Anrufe erhalten, die nicht auch durch einen anderen Anlass legitimiert waren. Desweiteren war sie einer der Teilnehmer, die sich als techniknah beschrieben, mit jahrelanger Erfahrung mit Computersystemen und auch einer eigenen Homepage. Neu für sie sei nur der Umgang mit dem Smartphone. Ihre Daten seien auch in ganz vielen Seiten im Internet vorhanden. Unter diesem Hintergrund sah sie mit dem Empfehlungssystem keinerlei Probleme, auch wenn es alle im Profil vorhandene Daten zugänglich machen würde. Sie beschrieb allerdings, dass ihr geraten worden



Abbildung 16: Von einem Teilnehmer sortierte Profilkarteikarten.

war, die Heimatadresse nicht anzugeben, weil ihr Umfeld da Bedenken sah. Sie selbst teile die allerdings eigentlich nicht.

T8 beschrieb sich als jemand mit Technikerfahrung, aber als jemand mit eingeschränkter Souveränität. Grundsätzlich sei sie auf Erklärungen angewiesen, bevor sie etwas am Computer beherrschen würde, den sie auch noch nicht so lange nutze. Wie bei T1 sei das Smartphone für sie neu gewesen. Sie saß neben T1, es kann natürlich nicht ausgeschlossen werden, dass sie von ihr beeinflusst wurde. Sie beschrieb, dass sie die Aufgabe von der anderen Sichtweise angegangen war: Alle Datensätze seien solche, die sie interessieren würden, wenn sie bei jemanden mitfahren wollte. Allerdings formulierte sie Bedenken dahingehend, dass die Daten umfassend genug seien um auch missbraucht werden könnten.

T14 ist ein Sonderfall in der Hinsicht, dass er zwar nicht alle Daten verfügbar machen wollte - Profilfoto und Name sollten privat bleiben - aber eine Steuerung der Freigabe ansprach. Beispielsweise sollte statt dem Namen ein Nutzernamen angezeigt werden. Dieses Datum, den realen Namen, wollte er auch am liebsten gar nicht in das System eingeben. Auch T3 hatte einen technikhnen Hintergrund beschrieben: Viele Jahre hatte er mit Computersystemen zu tun, anfangs sogar noch mit Lochkartenprogrammierung, beschrieb sich aber als jemand, der nur selten direkt mit der Programmierung (und damit der Nutzung) solcher Systeme zu tun hatte, wobei auch dies mal zumindest ansatzweise gelernt wurde. Allerdings sagte er auch, dass er Daten einfach nicht eintragen würde, wenn er sie nicht im System haben wollen würde - oder eine Funktion nutzen würde, sie nicht anzuzeigen, wenn es die gäbe.

Am Ende wurde noch auf einen Umstand hingewiesen, der für diese Plattform speziell ist: Es gibt bei ihr Fahrer und Mitfahrer, wobei einzelne Teilnehmer kein eigenes Auto benutzen. Eine Teilnehmerin formulierte, dass man sich dadurch minderwertig fühle.

T12: Deswegen ist man immer so ein bisschen ... man fühlt sich so minderwertig.

Lauter Protest der anderen

T12: Doch, weil wir haben keinen Führerschein, es müssen immer die anderen fahren. Man fühlt sich immer so als kleines Würstchen, ja, die können ja keinen mitnehmen, die haben kein Auto - und wir sind drauf angewiesen, ne?

...

Das begleitete einen das ganze Leben lang, weil man ist immer auf andere angewiesen.

Es blieb unklar, ob die von ihr wahrgenommene und formulierte soziale Verpflichtung, sich den Fahrern unterzuordnen - man wolle ja nicht zur Last fallen - auch ihr Verhalten im Workshop beeinflusste. Vorstellbar ist, dass sie sich in der Diskussion nicht völlig traute, Dinge vorzuschlagen, die anderslautend als die Meinungen der Fahrer waren.

Doch war sie mit ihrer Wahrnehmung eines Unterschieds nicht alleine. Für Mitfahrer jedoch habe eine Empfehlung von Fahrern einen höheren Wert, äußerte auch eine andere

Teilnehmerin, was auch die Bewertung eines Empfehlungssystems ändern würde, das direkt Nutzer als Fahrer empfiehlt:

T8: Die Position, in der ich mich befinde, ändert meine Haltung.

Insgesamt wurde im ersten Workshop das Empfehlungssystem positiv bewertet. Soweit sie sich äußerten konnten die Teilnehmer sich vorstellen, das System zu nutzen, sie fanden die Idee gut und nützlich und hatten auch keine Bedenken im Hinblick auf die einfache Handhabung des Systems. Im Gegenteil, ihre Vorschläge würden es eher schwieriger zu nutzen machen, was aber scheinbar zumindest teilweise akzeptabel zu sein schien.

7.1.2 Gruppe 2

Die fünf Teilnehmer der zweiten Gruppe taten sich meinem Eindruck nach insgesamt schwerer mit dem Workshop. Zwar konnten auch sie der Grundidee positive Seiten abgewinnen - so wurde als einer der ersten Redebeiträge ein mentales Modell vorgestellt, nach dem das Empfehlungssystem Kontextinformationen bei Veranstaltungen anzeigen würde und dass dies positiv sei. Doch insgesamt dauerte es länger, bis das System verstanden wurde, die Diskussion verlief weniger eigenständig und die Bewertung klang negativer. Dies führte dazu, dass Mitarbeiter des Projekts sich an der Diskussion beteiligten und Fragen stellten, was bei der ersten Gruppe nicht vorkam und auch nicht nötig war, um das Gespräch in Gang zu halten. Interessanterweise gab es in dieser Gruppe deutliche Abweichungen zur ersten Gruppe bei der Bewertung, welche Profilinformatoren vom System problemlos genutzt werden dürften.

Die Teilnehmer beschrieben, dass sie durchaus Erfahrung im Umgang mit IT-Systemen besitzen. Nur eine Teilnehmerin (T1) zeichnete ein Selbstbild mit wenig Erfahrung, die sich auf die Dateneingabe im Beruf beschränke. Bei den anderen vier waren im unterschiedlichen Maße einiges an Erfahrung vorhanden. Eine Teilnehmerin (T2) beschrieb, dass sie verschiedene Dinge wie Videobearbeitung mit dem Computer mache und darüber neue Sachen lerne, wenn sie diese bräuchte. Die beiden männlichen Teilnehmer hatten Erfahrung im Beruf und auch einen Rechner zuhause. Beachtenswerterweise nutzte der erste der beiden Teilnehmer (T3) die Gelegenheit, bestehende Probleme mit der Plattform zu betonen und klang dabei fast ein bisschen frustriert angesichts der seiner Beschreibung nach langsamen Fortschrittes. Der ältere der beiden (T4) hatte die Erfahrung im Beruf als gering bezeichnet und beschrieb, dass er früher Distanz zum Computer seiner Frau gehalten hätte, aus Angst etwas kaputtzumachen, inzwischen aber den eigenen Laptop frei nutze. Seine anwesende Frau (T5) beschrieb, dass sie anfänglich von ihrem Sohn einiges gelernt habe und inzwischen seit Jahren den Computer zum Verfassen von Texten nutze. Das Smartphone allerdings sei sehr umfangreich in seinen Möglichkeiten, eine Einschätzung, die zweifelsfrei stimmt, aber als Beschreibung der Schwierigkeiten dabei, seine Beherrschung zu erlernen, verstanden werden sollte. Wobei seine Möglichkeiten, wie Chatten mit den Enkelkindern und gelegentliches Spielen, durchaus schon jetzt genutzt würden.

Wie in Gruppe 1 waren die Teilnehmer in der Diskussion nicht an dem verhaftet, was ihnen vom Prototypen gezeigt wurde

Ich würde das anklicken und schauen, was angeboten wird, ob jemand fährt oder... wenn nicht, fahre ich selbst.

Das allerdings ist eine Kontextinformation, die im Prototypen nicht angezeigt wird. T3 sagte später zu dieser Idee:

Muss ja nicht jeder wissen, ob ich ins Kino, ins Apollo gehe, am Sonntag oder am Samstag. Ich denke mal, das ist mir zu privat.

Was zeigt, dass diese Kontextinformation durchaus nicht unkritisch ist. T5 war der Nutzen einer solchen Empfehlung nicht sofort ersichtlich, zeigte sich jedoch mit einer Erklärung von T2, dass es ja doch interessant sein könnte, um schließlich gemeinsam irgendwo hinzugehen, zufrieden. Woraufhin T3 einwarf:

Kurzes Auflachen Das ist ja so ähnlich wie hier! *hebt das Smartphone*

Gemeint war, dass die Teilnehmer schon jetzt in einer WhatsApp-Gruppe organisiert waren. Ihren Aussagen zufolge würde schon jetzt darüber gelegentlich Mitfahrten organisiert und bekanntgegeben, wer wann wohin geht.

Diese Information, ob Bekannte an einer bestimmten Veranstaltung teilnehmen, wurde auf Nachfrage hin als interessant deklariert. Auch T3, der in dieser Runde der kritischste Teilnehmer war, fand diese Idee gut.

Der Unterschied in der Präsentation der Veranstaltungslisten wurde thematisiert. Die kleinere Liste mit den Empfehlungen wurde zuerst positiv bewertet, als übersichtlicher, dann wurde von einem anderen Teilnehmer die Praxistauglichkeit kritisiert, weil die Liste mit mehreren Einträgen nicht skalieren würde. Später wurde diese Kritik erweitert: Während die normale Liste über alle Veranstaltungen es ermöglichen, frei auszuwählen, würde die Liste des Empfehlungssystem

...schon wieder etwas vorgeben

Die Frage, ob das Empfehlungssystem als störend empfunden werden würde, wurde allgemein verneint. Allerdings solle das Empfehlungssystem nicht die eigene Auswahl einschränken, sondern eine Option bleiben. Über die Frage, ob auch direkt Nutzer empfohlen werden sollten, herrschte Uneinigkeit. T2 bezeichnete dies direkt als valide Möglichkeit, die sie nicht stören würde. Allerdings brachte auch sie das Argument, dass auf einer Mitfahrplattform dies nur begrenzt sinnvoll sei, weil in ihr Auto ja nur eine begrenzte Menge an Mitfahrern passen würden. T1 empfand das einer kurzen Äußerung zufolge als kritisch. T3 verneinte zuerst entschieden, meinte aber die Information, dass er speziell irgendwohin gehen würde. Die reine Empfehlung von ihm als möglichen Fahrer, die konnte er sich nach etwas Zögern dann doch vorstellen. Später formulierte er, dass die ganze Idee für ihn etwas überraschend sei und er das erstmal verarbeiten müsse.

Kritisch jedoch wurde von ihm das Konzept mit den Gruppen bewertet. Denn in die könnten die Leute ja frei eintreten, was dazu führen könnte, dass zuviele Personen die Empfehlungen beziehungsweise die eigenen Aktivitäten sehen würden.

Plötzlich habe ich hundert Leute, die zum FC Köln fahren und gucken Fußball.

Dies äußerte T3, aber bekam Zustimmung von T5 und keinen Widerspruch. Auf die Frage einer Mitarbeiterin des Projektes, ob eine geschlossene kleine Gruppe helfen würde:

T3: Ich meinte, wenn das da eine kleine Gruppe ist, die ein besonderes Hobby haben und sich da austauschen, dann ist das ja in Ordnung.

M: Oder ist dann das Konzept, dass man da keine Kontrolle haben würde

T3: Ne, hat man ja nicht.

M: Dass man da gar nicht sieht, wer ist überhaupt Teil der Gruppe

T3: Ich meine, der Gedanke ist ja vielleicht nicht verkehrt, aber ich kann mich im Moment nicht damit anfreunden. Vielleicht ist die Idee ja nicht schlecht, aber im Moment bin ich da ein bisschen... skeptisch.

Damit war ersichtlich das ganze Konzept eines Empfehlungssystems gemeint.

Daraufhin kam T2 ins Zögern:

Ich bin da auch positiv, aber mir war nicht bewusst, dass ich darüber einen so großen Kreis angesprochen hätte. Ich hatte da eigentlich mehr so einen kleineren Kreis *lacht* wo vielleicht zwei Personen sind die sagen das würde mich interessieren, wir würden gerne mitfahren.

T1 hielt dagegen, dass ja auch mit dem vorgestellten Konzept eine kleine Veranstaltung in Siegen nur wenig Reaktionen hervorrufen würde, wenn die empfohlen würde.

Der Gedanke einer geschlossenen Gruppe oder die in Gruppe 1 von den Teilnehmern geforderten Filteroptionen wurden von der ganzen Gruppe gutgeheißen, nachdem sie von einem Mitarbeiter angesprochen wurden. T1 meinte:

Man hat das Gefühl, man ist da ein bisschen geschützt, wenn das so ist.

Überraschend bei der Bewertung der Informationen, die bei den Empfehlungen angezeigt werden dürften, war der Pragmatismus der Antworten. Insgesamt waren wieder die Informationen des Kernbereiches des privaten Lebens die, welche als kritisch bewertet wurden. Aber es wurde mehrfach erkannt, dass genau diese für einen Praxisbetrieb des Empfehlungssystems und der Plattform selbst nützlich sein könnten.

Es begann T1. Name, Adresse und Profilfoto, das waren die Informationen, die sie nicht gerne in den Empfehlungen hätte. Aber sie fügte hinzu, dass in einem zweiten Schritt, beim konkreten Aushandeln der Mitfahrgelegenheit, die Adresse ja sowieso geteilt werden müsste, und dass das Profilfoto dann auch nützlich sein würde.

Diese Zweiteilung des Prozesses übernahm T2. Sie wollte gerne so wenige Informationen wie möglich herausgeben, dahingehend bezeichnete sie sich selbst als datensparsam. Diese Informationen beinhalteten das Profilfoto, den Namen und die häufigen Ziele. Aber die anderen Informationen könnten eventuell später von ihr freigegeben werden, wenn sich die Begegnung entsprechend entwickelt. Sie beschrieb, dass sie ein solches Verfahren gerne erstmal austesten und ausprobieren wollen würde.

Das Profilfoto war für T3 etwas, was er (ähnlich wie A in Gruppe 1) gerne überhaupt nicht erst angeben wollen würde. Entsprechend könnte ein Empfehlungssystem es auch

nicht anzeigen. Dazu kommen die häufigen Ziele, die er nicht bekanntgeben wollen würde, ebenso wie die eigenen Interessen. Name und Adresse seien dagegen unproblematisch.

Eine pragmatische Betrachtungsweise wählte T4. Name, Bild und Adresse würde er anzeigen lassen, diese betrachtete er als unkritische Informationen, die einen Mitfahrer auch interessieren würden. Ebenso wie das eigene Fahrzeug. Aber eine Reihe weiterer Informationen, nämlich insbesondere die eigenen Interessen, wollte er nicht gerne freigeben, weil das genau die Themen seien, über die er sich mit einem Gast sowieso während der Fahrt unterhalten würde:

Ich fahre ja nicht mit einem Stummen. Und dann kann er alles erfahren, auch Interessen, na sicher, und so weiter und so weiter.

...

Der Mitfahrer will ja nur wissen: Hat der nur ein Fahrrad oder hat der auch ein Auto. *Lachen*

T5 unterschied in dem Szenario zwischen den Telefonnummern. Sie würde statt der Handynummer, die auf dem Beispielprofil angegeben war, lieber ihre Festnetznummer angeben, aus Bedenken, dass sie zu viele ungewollte Anrufe erhielt. Zuhause könne sie besser damit umgehen. Die häufigen Ziele seien ebenfalls kritisch. Das Profildfoto, der Name und die Adresse dagegen waren für sie nicht kritisch. Zu dem Foto meinte sie:

Ich denke immer, das ist wenn man jemanden sieht auch ein bisschen Vertrauenssache, wenn man mit jemand fremden mitfahren will, dass das schon so ein bisschen persönlich sein sollte.

Also bewertete auch sie pragmatisch das Nutzungsszenario.

7.1.3 Fazit der Workshops

Die Bedenken gegen den Kategorisierungsansatz nahm ich als ernstzunehmend wahr. Dies entspricht auch der Empfehlung von [Jones and Pu, 2007], den Aufwand für den Nutzer möglichst gering zu halten. Dementsprechend sollte das Empfehlungssystem um Möglichkeiten erweitert werden, auch ohne die Kategorisierung zu funktionieren, beispielsweise über die Nähe des Wohnortes der Teilnehmer.

Generell wurde eine größere Transparenz der Empfehlungen gewünscht, vor allem, um einen größeren Nutzen aus ihnen ziehen zu können. Dementsprechend sollte bei einer Empfehlung angezeigt werden, wer zu welcher Veranstaltung geht. Das kollidiert jedoch mit dem Wunsch nach Privatsphäre der Nutzer, da sie nicht wollten, dass alle Informationen ihres Profils für andere zugänglich sind. Dementsprechend müsste eine solche Funktion für die Nutzer steuerbar sein und die Einführung nochmals gesondert begleitet werden. Der Wunsch nach Filtern in der ersten Gruppe und die Bedenken über den Kreis der Empfehlungen in der zweiten zeigen, dass der Kontrollwunsch der Teilnehmer ausgeprägt ist. Dementsprechend ist das Konzept mit den offenen Gruppen, die eher selbst vergebaren Tags entsprechen, wieder im Schlaglicht, weil ein Konzept mit weniger ausgeprägter Öffentlichkeit damit nur schwer umzusetzen ist.

Die Ablehnung des Vorschlags, Nutzer direkt zu empfehlen, zeigt, dass die im Vorfeld von den Projektleitern vermuteten Bedenken gegen ein solches Modell eines Sozialen Netzwerks berechtigt waren. Durch den gleichzeitig sogar von den Nutzern formulierten Gewöhnungseffekt (siehe auch Iachello and Hong [2007] für die Erwartung, dass dieser der technologischen Entwicklung folgt) steht jedoch zu vermuten, dass diese Beschränkung im Laufe der Zeit immer wieder neu evaluiert werden sollte. Es ist möglich, dass nach Gewöhnung an und dadurch gesteigertem Vertrauen in das System eine solche Funktion überwiegend akzeptiert werden würde. Außerdem wurden Szenarien beschrieben, in denen sowieso Fahrer angezeigt würden, in denen eine Nutzung des Empfehlungssystems sinnvoll sein würde. Dies könnte ja auch neutral ohne Begründung gemacht werden.

Die Bedenken gegen ein Bewertungssystem und die ausgiebige Diskussion über Gütekriterien von Fahrern und Mitfahrern zeigt, dass auch ein solches Bewertungssystem kritisch wäre, wollte man es einführen, um das Empfehlungssystem zu erweitern. Dies könnte der Beobachtung in [Dearden and Watts, 2004] entsprechen, dass Nutzer ihre Selbstdarstellung kontrollieren wollen.

Erfahrung im Umgang mit Techniksystemen scheint die Akzeptanz solcher Empfehlungssysteme beziehungsweise den Umfang dieser Akzeptanz zu beeinflussen. Es ist nun keineswegs so, dass jeder mit Technikerfahrung automatisch und vollumfänglich alle seine privaten Daten für das Empfehlungssystem freigegeben hätte und das Ganze unkritisch sah. Das beste Gegenbeispiel dafür war T3 in Gruppe 2, der trotz laut Eigenbeschreibung erheblicher Technikerfahrung das Konzept wohl am kritischsten sah. Aber im ersten Workshop waren die Leute, die das System am unkritischsten sahen auch die, welche am meisten von vorheriger und ausgeprägter Technikerfahrung berichteten.

Allerdings unterschieden sich die Bewertungen der Informationen in den Gruppen doch beträchtlich. Während Gruppe 1 generell vermuten ließe, dass die Profildaten der eigenen Identität nicht geteilt werden sollten (im Gegensatz zur gängigen Praxis Sozialer Netzwerke wie Facebook), sah dies in Gruppe 2 schon ganz anders aus. Dort waren die häufigen Ziele, die ja auch durchaus in der engen Privatsphäre enthalten sind, das häufigstgenannte kritische Datum. Es steht daher zu vermuten, dass die Bewertung privater Informationen höchst beeinflussbar vom konkret ausgemalten Bedrohungsszenario, der pragmatischen Abwägung des gewünschten Nutzungsszenarios und den wortführenden Diskussionsteilnehmern ist. Was vermuten ließe, dass die Bewertung konkreter Umsetzungen zumindest hinsichtlich des Themenfeldes Datenschutz und Privacy erheblich von der Bewertung des Umfelds oder allgemeiner von meinungsproduzierenden Stellen abhinge, demnach ein Technikakzeptanzmodell wie TAM2 (Venkatesh and Davis [2000]), welches diese soziale Umgebung beinhaltet, stark zu bevorzugen wäre. Eine negative Kritik des Systems in den Medien, wie Facebook das erlebte, wäre hochproblematisch.

7.2 Fragebögen

Am Ende des Workshops wurden die Teilnehmer gebeten, einen an TAM angelehnten und von [Beil, 2013] übernommenen Fragebogen auszufüllen. Mit dem Fragebogen sollten verschieden Ziele erreicht werden:

1. Der subjektive Eindruck, wie die Workshops abliefen, geprüft werden.
2. Die Akzeptanz des Empfehlungssystem quantifiziert werden, soweit möglich.
3. Eine zusätzliche Möglichkeit geschaffen werden, Verbindungen zwischen der Akzeptanz einzelner Teilnehmer und Aussagen aus den Interviews/Ergebnissen der vorherigen Studie zu prüfen (auch wenn bei der Teilnehmeranzahl keine statistisch signifikanten Ergebnisse vermutet wurden).

7.2.1 Verhältnis zwischen Workshop und Fragebogen

Frage	T3*	T5*	T4*	T2*	T1*	T6	T7	T8	T9	T10	T11	T12	T13
PU1	4	3	3	2	4	3	5	5	5	4	4	3	5
PU2	4	2	3	2	4	4	5	5	5	4	4	5	3
PEOU1	3	3	2	3	5	4	5	5	3	5	4	3	4
PEOU2	3	5	4	1	3	4	4	5	3	3	4	3	3
Gefühl	5	5	3	4	5	5	5	5	3	5	2	2	3
PEOU3	4	5	2	3	2	4	5	5	5	5	4	3	3
Spaß	5	3	4	4	1	4	5	5	5	4	4	5	3
BI	5	3	4	4	4	4	5	5	5	4	3	5	5
Summe	33	29	25	23	28	32	39	40	34	34	29	29	29

Tabelle 4: Ergebnis des Fragebogens beider Gruppen. Die mit einem * markierten Teilnehmer nahmen am zweiten Workshop teil. Die Likert-Skala reichte von 1-5. Siehe Abschnitt 6.2.

Gesamt	Gruppe1	Gruppe 2
31,09	33,25	27,6

Tabelle 5: Durchschnittliche Gesamtbewertung.

Aus den Ergebnissen (siehe Tabelle 4) lässt sich ablesen, dass der generell positive Eindruck der Workshops bestätigt werden kann. Die durchschnittliche Gesamtbewertung von 31,09 (siehe Tabelle 5), bei einer bestmöglichen von 40, lässt auf eine gute Akzeptanz schließen.

Falls der Fragebogen tatsächlich die Einstellung der Teilnehmer gegenüber dem Empfehlungssystem widerspiegelt, wäre zu erwarten, dass die Bewertung mit Aussagen der Teilnehmer in den Workshops zu verbinden ist. Tatsächlich lässt sich das an einzelnen Aussagen und Eindrücken ableiten. Ein Beispiel dafür ist T8 aus dem ersten Workshop. Ihre Bewertung mit der Maximalpunktzahl würde nur dann passen, wenn sie sich im Workshop als sehr positiv eingestellt gezeigt hat. Dies war der Fall. Sie war noch dazu eine der Wortführerinnen in dieser Gruppe. Sie war die Teilnehmerin, die sagte, dass je nach Position (Fahrer oder Mitfahrer) ihre Akzeptanz des Empfehlungssystems steige,

und meinte, dass sie dem System nun wesentlich positiver gegenüberstehe als noch am Anfang des Workshops.

Genau entgegengesetzt war die Entwicklung von T2 im zweiten Workshop. Sie äußerte bei der Vorstellung des Prototypen spontan Zustimmung und bewertete die Empfehlungen zuerst als sehr positiv. Aber im Verlauf des Workshops wandelte sich ihr Eindruck. Kurz vor Ende sagte sie, dass ihr nicht bewusst gewesen sei, wieviele Leute eine solche Empfehlung sehen würden. Auch war sie sehr zurückhaltend bei der Beschreibung, welche Informationen sie freigeben würde und betonte, dass sie so wenig Daten wie möglich zeigen wollen würde (aber pragmatisch abwägend, wieviel für die Empfehlungen nötig seien). Und dazu passend gab sie eine geringe Bewertung von 23 ab. Wobei es dann doch überraschend war, dass sie trotz der anfänglichen Zustimmung die niedrigste Bewertung aller Teilnehmer aussprach.

Es findet sich kein Gegenbeispiel, bei dem ein Teilnehmer die Empfehlungen im Workshop erst durchgängig lobte, um dann deutlich negativ zu bewerten, und auch kein umgekehrtes, nämlich positive Bewertung trotz durchgängig negativer Äußerungen. Stattdessen scheinen die Bewertungen durchaus mit den Gesprächsverläufen zusammenzupassen. Trotzdem gab es Fälle, bei denen die Bewertung nicht direkt vorhergesehen wurde. T3 als einer der Wortführer in der zweiten Gruppe beispielsweise bewertete mit 33 Punkten die Empfehlungen positiver als erwartet, wobei er im Verlauf des Workshops sich auch zunehmend positiver äußerte, aber eben bei anfänglicher Ablehnung. Und T4 in der gleichen Gruppe bewertete mit 25 Punkten schlechter als vermutet, wobei er sich im Verlauf des Workshops auch nicht positiv (und generell nur selten) äußerte, wobei seine Informationsbewertung sehr pragmatisch und er daher der Idee gegenüber eher offen erschien.

7.2.2 Vergleich mit der Bewertung der Gesamtplattform

Frage	T3*	T5*	T4*	T2*	T1*	T6	T7	T8	T9	T10	T11	T12	T13
PU1	5	0	5	3	5	4	5	5	5	4	4	.	5
PU2	5	0	5	5	5	4	5	5	5	5	5	3	5
PEOU1	4	0	4	3	5	4	5	4	5	4	4	3	5
PEOU2	5	0	4	3	5	3	4	5	5	5	4	3	5
Gefühl	5	0	5	5	4	4	1	5	5	5	5	4	5
PEOU3	5	0	5	3	5	5	5	5	5	5	5	4	5
Spaß	5	0	5	5	5	4	5	5	5	5	5	5	5
BI	5	0	5	5	5	4	5	5	5	5	5	5	5
Summe	39	0	38	32	39	32	35	39	40	38	37	27	40

Tabelle 6: Ergebnis des Fragebogens der Teilnehmer in der vorherigen Studie von Beil, siehe Beil [2013]. Die mit einem * markierten Teilnehmer nahmen am zweiten Workshop teil. Die Likert-Skala reichte von 1 bis 5.

Allerdings ist die Bewertung des Empfehlungssystems insgesamt wesentlich negativer

Gesamt	Gruppe1	Gruppe 2
36,73	36,57	37

Tabelle 7: Durchschnittliche Gesamtbewertung in der vorherigen Studie, siehe Beil [2013].

als die Bewertung, die in der vorherigen Studie für die Gesamtplattform erreicht wurde (siehe Tabelle 6 und 7). Das Ergebnis dort von den gleichen Teilnehmern war 36,73, also etwa 5,5 Punkte höher.

Die schlechtere Bewertung des Empfehlungssystems gegenüber der Bewertung für die Gesamtplattform kommt überwiegend aus der schlechteren Bewertung der zweiten Gruppe. Während die Teilnehmer im ersten Workshop noch durchschnittlich eine Bewertung von 33,25 vergaben - immer noch niedriger, aber nur um 3 Punkte - lag das durchschnittliche Ergebnis der zweiten Gruppe bei nur 27,6 (siehe Tabelle 5). Eine solche Abweichung ist bei der vorherigen Studie bei der Bewertung der Gesamtplattform nicht zu beobachten, ein Hinweis darauf, dass die Akzeptanz des Empfehlungssystems nicht grundsätzlich mit der Akzeptanz der Plattform korreliert. Allerdings bestätigt dies den im vorherigen Abschnitt beschriebenen subjektiven Eindruck aus den Workshops, dass die zweite Gruppe dem Empfehlungssystem kritischer gegenüberstand.

Die nicht ersichtliche Korrelation zwischen der Bewertung der Gesamtplattform, die insgesamt sehr positiv bewertet wurde, und der des Empfehlungssystems lässt vermuten, dass sich auch bei näherer Betrachtung der Hintergrunddaten der Teilnehmer über die Fragebögen der vorherigen Studie kein Faktor finden lassen wird, der gleichermaßen mit der Akzeptanz beider Systeme korreliert. Stattdessen dürfte die individuelle Technikerfahrung und das eigene Verhältnis zwischen dem Wunsch nach Datenschutz und Kommunikationsbedürfnis sich spezifisch auf ein so spezielles System wie ein Empfehlungssystem auswirken. Es wurde geprüft, ob sich bei den Fragen zur Technikeinstellung und Technikerfahrung ein Faktor finden lässt, der die Akzeptanz des Empfehlungssystems beeinflusst.

Sicherheitshalber wurde die Spearman-Rangkorrelation zwischen den Antworten der Teilnehmer zu ihrer Technikeinstellung (in Beil [2013]) und der Bewertung des Empfehlungssystems berechnet. Es ergaben sich zwar bis zu mittelstarke Zusammenhänge, beispielsweise mit der Bewertung der Aussage

Ein Beruf, der mit Technik zu tun hat, wäre nichts für mich gewesen.

Das wäre auch nicht abwegig gewesen, denn die Bewertung dieser Aussage spiegelt möglicherweise sowohl die bewahrte Begeisterungsfähigkeit für Technik als auch die gefühlte Kompetenz gegenüber einem technischen System wieder, was mit der Bewertung eines nicht unbedingt uninteressanten Empfehlungssystems im Rahmen der Diskussion seiner Möglichkeiten tatsächlich korrelieren könnte. Allerdings war keiner dieser Zusammenhänge statistisch signifikant.

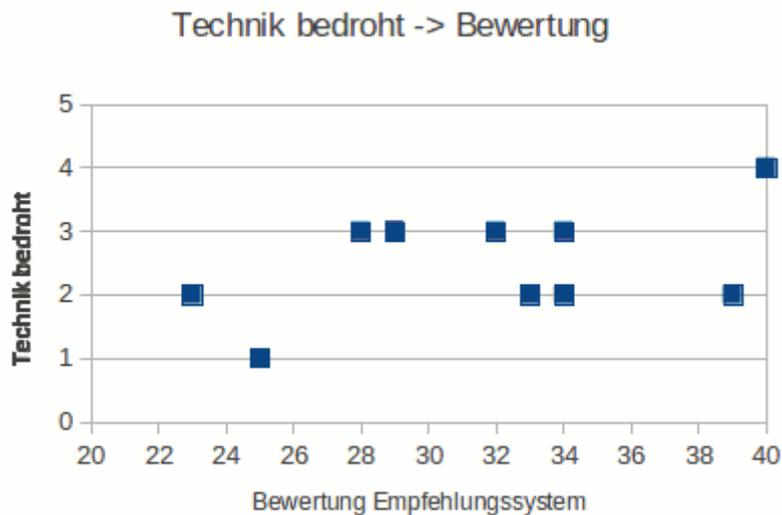


Abbildung 17: Als Beispiel: Graph des Zusammenhangs zwischen Zustimmung zur These, dass Technik bedrohlich ist, und der Bewertung des Empfehlungssystems.

7.2.3 Vergleich mit den Interviews

Solche Faktoren sind auch aus den Interviews nicht ersichtlich. T2, Teilnehmerin im zweiten Workshop, sei hier ein Beispiel. In dem Interview beschrieb sie, dass sie einiges an Technikerfahrung besitze. So nutze sie aktiv das Internet für Recherchezwecke, beispielsweise für die gelegentliche Reise - und nutzt dafür mit Google ein Empfehlungssystem. Außerdem sei der Computer für sie generell ein Werkzeug, um Filme zu schneiden, als Hobby. Außerdem zeigte sie Interesse an der Mobilitätsplattform an sich - wenn auch eher in der Rolle eines Fahrers - und beschrieb, dass sie vorhabe, in naher Zukunft ein Navi auszuprobieren.

Gleichzeitig besitze sie, wie viele der Teilnehmer, eine ausgeprägte Fahrgemeinschaftspraxis. Regelmäßig werden Freunde und Bekannte transportiert, moderiert durch feste und spontane Absprachen. Fahrgemeinschaften, die nicht nur zur Alltagsbewältigung wie dem Einkaufen genutzt werden, sondern auch, um gemeinsam zu Veranstaltungen zu fahren.

Es ist bei ihr also definitiv die Bereitschaft vorhanden, neue Technikartefakte auszuprobieren. Wie viele der Teilnehmer ist sie überdurchschnittlich technikaffin. Auch war sie nicht älter als der Rest der Teilnehmer. Es sind also keine Faktoren erkennbar, die im Interview erfragt wurden und die nun im Nachhinein als Warnzeichen erkennbar sind und die geringe Bewertung des Empfehlungssystems erklären.

Ihre Interviewaussagen passen auch nicht wirklich zu Einzelbewertungen. Beispielsweise bewertete diese Aussage mit einer 1, also dass sie nicht zutreffe:

Ich glaube, ich konnte die Empfehlungen auch nutzen, wenn keiner da wäre,

der mir sagt, was ich tun muss

Zwar beschrieb sie in dem Interview, dass sie ausgiebig auf Hilfe zurückgreife, wenn etwas am Computer nicht funktioniere. Aber ansonsten vermittelte sie durchaus den Eindruck, Nutzungsbarrieren auch ohne Hilfestellung überwinden zu können. Vielleicht ist diese Diskrepanz mit einem Fehler beim Ausfüllen des Fragebogens erklärbar.

Hypothese war, dass Aussagen zur eigenen Technikdistanz mit einer geringeren Bewertung korrelieren könnten. T2 soll beispielhaft zeigen, dass dies bei den niedrigen Bewertungen nicht beobachtet wurde. Erst im Workshop äußerte sie, dass sie bei der Preisgabe privater Daten sehr vorsichtig sein wolle, wobei unklar blieb, ob dies als generelle Aussage gemeint war oder eher auf diese spezielle Situation bezogen wurde.

Teilweise zeigte sich das Gegenteil, dass Teilnehmer trotz entsprechender Aussagen in den Interviews das Empfehlungssystem positiv bewerteten, auch ohne gegenteilige Aussagen in den Workshops. Dies passierte beispielweise bei T9, die das System mit 34 Punkten bewertete.

In ihrem Interview äußerte sie zum Beispiel über das Internet, dass dessen Nutzung nicht ungefährlich sei. Sie beschrieb aber auch, dass diese Bewertung nicht alleine die ihre sei, sondern auf den Einfluss eines Familienmitgliedes zurückgehe.

Andererseits könnte dies natürlich schon Auswirkungen haben, aber teilweise durch andere Faktoren ausgeglichen werden. Denn in den Workshops wurde formuliert, dass die Empfehlungen für Teilnehmer ohne Führerschein oder Auto einen stärkeren Nutzen haben könnten. T9 war eine der Teilnehmerinnen, die kein eigenes Auto besitzt. Demzufolge könnte dies die Bewertung erhöht haben.

8 Fazit

Anhand von vorhandenen Interviews der Teilnehmer eines Mobilitätsprojektes für Menschen über 60 in der Region Siegen-Wittgenstein wurden Anforderungen für ein Empfehlungssystem als Teil des Projektes abgeleitet. Dieses wurde implementiert und daraufhin in zwei Workshops mit 13 Teilnehmern evaluiert, und zusätzlich mit einem an TAM angelehnten Fragebogen die Akzeptanz des Systems erfragt.

Deutlich wurde, dass ein Empfehlungssystem von den Teilnehmern nicht grundlegend abgelehnt wurde. Dies war eines der möglichen Szenarien und wurde nicht bestätigt. Zwar wurde das Empfehlungssystem negativer bewertet als die Gesamtplattform, aber insgesamt war die Bewertung immer noch gut. Zusätzlich belegt wird das durch Aussagen der Teilnehmer in den Workshops. Eine tiefsitzende Angst samt der Vorstellung eines intransparenten System, das irgendwie die privaten Daten durchsucht und verknüpft und daraus Empfehlungen ableitet, schien nicht zu bestehen.

Aussagen der Teilnehmer in den Workshops legen nahe, dass von dieser Zielgruppe eine Empfehlung von Mitbenutzern, verbunden mit der Preisgabe von Profilinformatio- nien, kritisch gesehen wird. Es war aber nicht eindeutig zu entscheiden, welche Profilinfor- mationen als besonders schützenswert eingestuft werden. Dies variierte und wurde teils sehr pragmatisch im Hinblick auf die Funktion der Mobilitätsplattform abgewogen. Aber die

erwartete Tendenz hin zum Schutz der privaten Kerndaten wie Adresse, Profilfoto und Telefonnummer konnte bestätigt werden.

Es konnten keine Faktoren identifiziert werden, mit denen man vorher hätte absehen können, ob, beziehungsweise wie ein solches Empfehlungssystem akzeptiert werden würde. Allerdings wurde mittels der Literaturrecherche die Transparenz und Erklärung als wichtiger Faktor entdeckt, die beide durch Aussagen der Teilnehmer im Workshop gestützt wurde.

Diese Aussagen begründen ein weiteres Fazit: Zumindest einzelne Teilnehmer wünschten sich, weitergehende Informationen über die Empfehlungen zu sehen. Mithilfe der Empfehlungen sollten Kontextinformationen angezeigt werden, um verschiedene Nutzungsszenarien zu unterstützen - wie dem Prüfen, ob an einer Veranstaltung Bekannte teilnehmen. Dies steht allerdings im Konflikt mit der oben erwähnten Brisanz der Preisgabe von als privat empfundenen Informationen. Eine Konfigurierbarkeit über Filtereinstellungen für solche Möglichkeiten scheint erforderlich

Diese Arbeit zeigte auch den Nutzen der Kombination verschiedener Hilfsmittel zum Ermitteln von Nutzerrückmeldung. So diente der TAM-Fragebogen als Bestätigung des subjektiven Eindrucks, dass die Teilnehmer des zweiten Workshops dem Empfehlungssystem weniger positiv gegenüberstanden als die des ersten Workshops, was es ermöglichte, die Aussagen unter diesem Gesichtspunkt nochmal zu prüfen. Gleichzeitig eigneten sich die Workshops als Werkzeug, um Ideen für die weitere Entwicklung des Empfehlungssystems wie einer geeigneteren Darstellung zu gewinnen. Während die Fragebögen wiederum es zusätzlich ermöglichten, die Bewertung des Systems mit den von Beil ermittelten Hintergrundinformationen wie der Technikeinstellung der Teilnehmer zu vergleichen, auch wenn dieser Vergleich keine statistisch signifikanten Korrelationen zeigte.

Literatur

- Gediminas Adomavicius and Jingjing Zhang. Stability of recommendation algorithms. *ACM Trans. Inf. Syst.*, 30(4):23:1–23:31, November 2012. ISSN 1046-8188. doi: 10.1145/2382438.2382442. URL <http://doi.acm.org/10.1145/2382438.2382442>.
- Louise Barkhuus. The mismeasurement of privacy: using contextual integrity to reconsider privacy in hci. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 367–376, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1015-4. doi: 10.1145/2207676.2207727. URL <http://doi.acm.org/10.1145/2207676.2207727>.
- Judith Beil. Beschreibung der Nutzerstichprobe. 2013.
- Izak Benbasat and Henri Barki. Quo vadis tam? *J. AIS*, 8(4), 2007. URL <http://dblp.uni-trier.de/db/journals/jais/jais8.html#BenbasatB07>.
- Matthias Bender. Konzeption und Entwicklung eines Frameworks zur Evaluation semantischer Recommender Systeme. 2011.
- Kevin Beyer, Jonathan Goldstein, Raghu Ramakrishnan, and Uri Shaft. When is "nearest neighbor" meaningful? In *In Int. Conf. on Database Theory*, pages 217–235, 1999.
- Michele Bezzi. Expressing privacy metrics as one-symbol information. In *Proceedings of the 2010 EDBT/ICDT Workshops*, EDBT '10, pages 29:1–29:5, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-990-9. doi: 10.1145/1754239.1754272. URL <http://doi.acm.org/10.1145/1754239.1754272>.
- Philip Bonhard, Clare Harries, John McCarthy, and M Angela Sasse. Accounting for taste: using profile similarity to improve recommender systems. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1057–1066. ACM, 2006.
- Margot Brereton, Paul Roe, Marcus Foth, Jonathan M. Bunker, and Laurie Buys. Designing participation in agile ridesharing with mobile social software. In *Australasian Computer-Human Interaction Conference*, pages 257–260, 2009. doi: 10.1145/1738826.1738868.
- Bundesverfassungsgericht. BVerfGE 65, 1 – Volkszählung. Urteil des Ersten Senats vom 15. Dezember 1983 auf die mündliche Verhandlung vom 18. und 19. Oktober 1983 – 1 BvR 209, 269, 362, 420, 440, 484/83 in den Verfahren über die Verfassungsbeschwerden, 1983. <http://www.datenschutz-berlin.de/gesetze/sonstige/volksz.htm> [Access date 21 November 2006].
- Sylvain Castagnos, Nicolas Jones, and Pearl Pu. Eye-tracking product recommenders' usage. In *Proceedings of the fourth ACM conference on Recommender systems*, RecSys '10, pages 29–36, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-906-0. doi: 10.1145/1864708.1864717. URL <http://doi.acm.org/10.1145/1864708.1864717>.

- Li Chen and Pearl Pu. A cross-cultural user evaluation of product recommender interfaces. In Pearl Pu, Derek G. Bridge, Bamshad Mobasher, and Francesco Ricci, editors, *RecSys*, pages 75–82. ACM, 2008. ISBN 978-1-60558-093-7. URL <http://dblp.uni-trier.de/db/conf/recsys/recsys2008.html#ChenP08>.
- Li Chen, Wen Wu, and Liang He. How personality influences users’ needs for recommendation diversity? In *CHI ’13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’13, pages 829–834, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1952-2. doi: 10.1145/2468356.2468505. URL <http://doi.acm.org/10.1145/2468356.2468505>.
- T. Cover and P. Hart. Nearest neighbor pattern classification. *Information Theory, IEEE Transactions on*, 13(1):21–27, 1967. ISSN 0018-9448. doi: 10.1109/TIT.1967.1053964.
- Henriette S. M. Cramer, Vanessa Evers, Satyan Ramlal, Maarten Van Someren, Lloyd Rutledge, Natalia Stash, Lora Aroyo, and Bob J. Wielinga. The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-adapted Interaction*, 18:455–496, 2008. doi: 10.1007/s11257-008-9051-3.
- Henriette S.M. Cramer, Vanessa Evers, Maarten W. van Someren, and Bob J. Wielinga. Awareness, training and trust in interaction with adaptive spam filters. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’09, pages 909–912, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-246-7. doi: 10.1145/1518701.1518839. URL <http://doi.acm.org/10.1145/1518701.1518839>.
- Paolo Cremonesi, Franca Garzotto, Sara Negro, Alessandro Papadopoulos, and Roberto Turrin. Comparative evaluation of recommender system quality. In *CHI ’11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’11, pages 1927–1932, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0268-5. doi: 10.1145/1979742.1979896. URL <http://doi.acm.org/10.1145/1979742.1979896>.
- Fred Davis. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *Management Information Systems Quarterly*, 1989.
- A. Dearden and L. Watts, editors. *Instant Messaging and Privacy*, Leeds, England, September 2004.
- Paul Dourish and Ken Anderson. Collective information practice: Exploring privacy and security as social and cultural phenomena. *Human-Computer Interaction**Human-Computer Interaction*, 21(3):319–342, 2006.
- Cynthia Dwork. Differential privacy. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, 2006. ISBN 978-3-540-35907-4. doi: 10.1007/11787006_1. URL http://dx.doi.org/10.1007/11787006_1.

- M. Fishbein and I. Belief Ajzen. *Belief, attitude, intention and behavior: an introduction to theory and research*. 1975.
- Jill Freyne, Michal Jacovi, Ido Guy, and Werner Geyer. Increasing engagement through early recommender intervention. In *Proceedings of the third ACM conference on Recommender systems*, RecSys '09, pages 85–92, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-435-5. doi: 10.1145/1639714.1639730. URL <http://doi.acm.org/10.1145/1639714.1639730>.
- Werner Geyer, Casey Dugan, David R. Millen, Michael Muller, and Jill Freyne. Recommending topics for self-descriptions in online user profiles. In *Proceedings of the 2008 ACM conference on Recommender systems*, RecSys '08, pages 59–66, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-093-7. doi: 10.1145/1454008.1454019. URL <http://doi.acm.org/10.1145/1454008.1454019>.
- Ido Guy, Naama Zwerdling, David Carmel, Inbal Ronen, Erel Uziel, Sivan Yogev, and Shila Ofek-Koifman. Personalized recommendation of social software items based on social relations. In *Proceedings of the third ACM conference on Recommender systems*, RecSys '09, pages 53–60, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-435-5. doi: 10.1145/1639714.1639725. URL <http://doi.acm.org/10.1145/1639714.1639725>.
- Jeerold L. Hale, Brian J. Householder, and Kathryin L. Greene. *The Theory of Reasoned Action*, pages 259–287. SAGE Publications, Inc., 0 edition, 2002. URL <http://dx.doi.org/10.4135/9781412976046>.
- Jiawei Han and Micheline Kamber. *Data Mining. Concepts and Techniques*. Morgan Kaufmann, 2nd ed. edition, 2006. ISBN 1558609016. URL <http://www.amazon.de/Mining-Concepts-Techniques-Kaufmann-Management/dp/1558609016>.
- M. Hassenzahl, M. Burmester, and F. Koller. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. pages 187–196, 2003.
- Marc Hassenzahl. The interplay of beauty, goodness, and usability in interactive products. *Hum.-Comput. Interact.*, 19(4):319–349, December 2008. ISSN 0737-0024. doi: 10.1207/s15327051hci1904_2. URL http://dx.doi.org/10.1207/s15327051hci1904_2.
- T. Hennig-Thurau, E. C. Malthouse, C. Frieger, S. Gensler, L. Lobschat, A. Rangaswamy, and B. Skiera. The Impact of New Media on Customer Relationships. *Journal of Service Research*, 13:311–330, 2010. doi: 10.1177/1094670510375460.
- Sascha Herr, Andreas Rösch, Christoph Beckmann, and Tom Gross. Informing the design of group recommender systems. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '12, pages 2507–2512, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1016-1. doi: 10.1145/2212776.2223827. URL <http://doi.acm.org/10.1145/2212776.2223827>.

- Andreas Hotho, Robert Jäschke, Christoph Schmitz, and Gerd Stumme. FolkRank: A ranking algorithm for folksonomies. In *UNIVERSITY OF HILDESHEIM, INSTITUTE OF COMPUTER SCIENCE*, pages 111–114, 2006.
- Rong Hu and Pearl Pu. Acceptance issues of personality-based recommender systems. In *Proceedings of the third ACM conference on Recommender systems, RecSys '09*, pages 221–224, New York, NY, USA, 2009a. ACM. ISBN 978-1-60558-435-5. doi: 10.1145/1639714.1639753. URL <http://doi.acm.org/10.1145/1639714.1639753>.
- Rong Hu and Pearl Pu. A comparative user study on rating vs. personality quiz based preference elicitation methods. In *Proceedings of the 14th international conference on Intelligent user interfaces, IUI '09*, pages 367–372, New York, NY, USA, 2009b. ACM. ISBN 978-1-60558-168-2. doi: 10.1145/1502650.1502702. URL <http://doi.acm.org/10.1145/1502650.1502702>.
- Rong Hu and Pearl Pu. A Study on User Perception of Personality-Based Recommender Systems. In *User Modeling, Adaptation, And Personalization, Proceedings*, volume 6075 of *Lecture Notes in Computer Science*, pages 291–302. Springer-Verlag New York, Ms Ingrid Cunningham, 175 Fifth Ave, New York, Ny 10010 Usa, 2010.
- Giovanni Iachello and Gregory D. Abowd. Privacy and proportionality: Adapting legal evaluation techniques to inform design in ubiquitous computing. pages 91–95, Portland, Oregon, USA, 2005. ACM Press, New York. Conf. on Human Factors in Computing Systems.
- Giovanni Iachello and Jason Hong. End-user privacy in human-computer interaction. *Found. Trends Hum.-Comput. Interact.*, 1(1):1–137, January 2007. ISSN 1551-3955. doi: 10.1561/1100000004. URL <http://dx.doi.org/10.1561/1100000004>.
- Nicolas Jones and Pearl Pu. User Technology Adoption Issues in Recommender Systems. In *Proceedings of the 2007 Networking and Electronic Commerce Research Conference*, pages 379–394, Riva del Garda, Italy, 2007. URL http://atsma.org/NAEC_2007/.
- Young Ae Kim and Jaideep Srivastava. Impact of social influence in e-commerce decision making. In *Proceedings of the ninth international conference on Electronic commerce, ICEC '07*, pages 293–302, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-700-1. doi: 10.1145/1282100.1282157. URL <http://doi.acm.org/10.1145/1282100.1282157>.
- Bart P. Knijnenburg, Svetlin Bostandjiev, John O'Donovan, and Alfred Kobsa. Inspectability and control in social recommenders. In *RecSys*, pages 43–50, 2012.
- Joseph A. Konstan and John Riedl. Recommender systems: from algorithms to user experience. *User Modeling and User-Adapted Interaction*, 22(1-2):101–123, April 2012. ISSN 0924-1868. doi: 10.1007/s11257-011-9112-x. URL <http://dx.doi.org/10.1007/s11257-011-9112-x>.

- Younghwa Lee, Kenneth A. Kozar, and Kai R. T. Larsen. The technology acceptance model: Past, present, and future. *Communications of the Association for Information Systems*, 12:752–780.
- Arto Lehtiniemi. Evaluating supermusic: streaming context-aware mobile music service. In *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, ACE '08*, pages 314–321, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-393-8. doi: 10.1145/1501750.1501826. URL <http://doi.acm.org/10.1145/1501750.1501826>.
- Ninghui Li, Tiancheng Li, and S. Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and l-diversity. In *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*, pages 106–115, 2007. doi: 10.1109/ICDE.2007.367856.
- Brian Y. Lim, Anind K. Dey, and Daniel Avrahami. Why and why not explanations improve the intelligibility of context-aware intelligent systems. In *Computer Human Interaction*, pages 2119–2128, 2009. doi: 10.1145/1518701.1519023.
- G. Linden, B. Smith, and J. York. Amazon.com recommendations: item-to-item collaborative filtering. *Internet Computing, IEEE*, 7(1):76–80, 2003. ISSN 1089-7801. doi: 10.1109/MIC.2003.1167344.
- Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkatasubramanian. L-diversity: Privacy beyond k-anonymity. *ACM Trans. Knowl. Discov. Data*, 1(1), March 2007. ISSN 1556-4681. doi: 10.1145/1217299.1217302. URL <http://doi.acm.org/10.1145/1217299.1217302>.
- Carman Neustaedter and Saul Greenberg. The Design of a Context-Aware Home Media Space for Balancing Privacy and Awareness. In *UbiComp 2003: Ubiquitous Computing*, pages 297–314, 2003. URL <http://www.springerlink.com/content/RE70Q1D4E7CDX46Y>.
- Helen Nissenbaum. Privacy as Contextual Integrity. *Washington Law Review*, 79(1), 2004.
- Donald A. Norman. *The Design of Everyday Things*. Basic Books, 2002. ISBN 9780465067107.
- Fatih Kursat Ozenc, Lorrie F. Cranor, and James H. Morris. Adapt-a-ride: understanding the dynamics of commuting preferences through an experience design framework. In *Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces, DPPI '11*, pages 61:1–61:8, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-1280-6. doi: 10.1145/2347504.2347571. URL <http://doi.acm.org/10.1145/2347504.2347571>.
- A. Patrick, S. Marsh, and P. Briggs. In *Security and Usability: Designing Secure Systems that People Can Use*. 2005. ISBN 9780596553852.

- Jennifer Peerce and Ben Shneiderman. The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction*, 1, 2009. URL <http://aisel.aisnet.org/thci/vol1/iss1/5/>.
- Andreas Pfitzmann and Marit Hansen. A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management. *Version 0.34 Aug*, 10, 2010.
- Constance Elise Porter and Naveen Donthu. Using the technology acceptance model to explain how attitudes determine internet usage: The role of perceived access barriers and demographics. *Journal of Business Research*, 59(9):999–1007, September 2006. URL <http://ideas.repec.org/a/eee/jbrese/v59y2006i9p999-1007.html>.
- Pearl Pu, Maoan Zhou, and Sylvain Castagnos. Critiquing recommenders for public taste products. In *Proceedings of the third ACM conference on Recommender systems*, RecSys '09, pages 249–252, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-435-5. doi: 10.1145/1639714.1639760. URL <http://doi.acm.org/10.1145/1639714.1639760>.
- Pearl Pu, Li Chen, and Rong Hu. A user-centric evaluation framework for recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems*, RecSys '11, pages 157–164, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0683-6. doi: 10.1145/2043932.2043962. URL <http://doi.acm.org/10.1145/2043932.2043962>.
- Pearl Pu, Li Chen, and Rong Hu. Evaluating recommender systems from the user's perspective: survey of the state of the art. *User Modeling and User-Adapted Interaction*, 22(4-5):317–355, 2012. ISSN 0924-1868. doi: 10.1007/s11257-011-9115-7. URL <http://dx.doi.org/10.1007/s11257-011-9115-7>.
- Tim Reichling and Volker Wulf. Expert recommender systems in practice: evaluating semi-automatic profile generation. In *CHI*, pages 59–68, 2009.
- Tim Reichling, Michael Veith, and Volker Wulf. Expert recommender: Designing for a network organization. *Comput. Supported Coop. Work*, 16(4-5):431–465, October 2007. ISSN 0925-9724. doi: 10.1007/s10606-007-9055-2. URL <http://dx.doi.org/10.1007/s10606-007-9055-2>.
- Guy Shani and Asela Gunawardana. Evaluating recommendation systems. In Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor, editors, *Recommender Systems Handbook*, pages 257–297. Springer US, 2011. ISBN 978-0-387-85819-7. doi: 10.1007/978-0-387-85820-3_8. URL http://dx.doi.org/10.1007/978-0-387-85820-3_8.
- Martin Svensson, Kristina Höök, Jarmo Laaksolahti, and Annika Waern. Social navigation of food recipes. In *Computer Human Interaction*, pages 341–348, 2001. doi: 10.1145/365024.365130.

- Kirsten Swearingen and Rashmi Sinha. Beyond Algorithms: An HCI Perspective on Recommender Systems. 2001.
- LATANYA SWEENEY. k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY
1. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, 2002.
- Panagiotis Symeonidis, Alexandros Nanopoulos, and Yannis Manolopoulos. Moviexplain: a recommender system with explanations. In *RecSys*, pages 317–320, 2009.
- Nava Tintarev and Judith Masthoff. Effective explanations of recommendations: user-centered design. In *Proceedings of the 2007 ACM conference on Recommender systems*, pages 153–156, Minneapolis, MN, USA, 2007. ACM. ISBN 978-1-59593-730-8. doi: <http://doi.acm.org/10.1145/1297231.1297259>.
- Viswanath Venkatesh and Hillol Bala. Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, 39(2):273–315, 2008. ISSN 1540-5915. doi: [10.1111/j.1540-5915.2008.00192.x](http://dx.doi.org/10.1111/j.1540-5915.2008.00192.x). URL <http://dx.doi.org/10.1111/j.1540-5915.2008.00192.x>.
- Viswanath Venkatesh and Fred D. Davis. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. 2000.
- Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3): pp. 425–478, 2003. ISSN 02767783. URL <http://www.jstor.org/stable/30036540>.
- Thomas Vander Wal. Folksonomy coinage and definition, 2007. URL <http://vanderwal.net/folksonomy.html>.
- Yiwen Wang, Natalia Stash, Lora Aroyo, Peter Gorgels, Lloyd Rutledge, and Guus Schreiber. Recommendations based on semantically enriched museum collections. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(4):283–290, 2008. ISSN 1570-8268. doi: <http://dx.doi.org/10.1016/j.websem.2008.09.002>. URL <http://portal.acm.org/citation.cfm?id=1464505.1464602>.
- A. Westin. Opinion Surveys: What Consumers Have to Say About Information Privacy. 2001.
- Volker Wulf, Markus Rohde, Volkmar Pipek, and Gunnar Stevens. Engaging with practices: design case studies as a research framework in cscw. In Pamela J. Hinds, John C. Tang, Jian Wang, Jakob E. Bardram, and Nicolas Ducheneaut, editors, *CSCW*, pages 505–512. ACM, 2011. ISBN 978-1-4503-0556-3. URL <http://dblp.uni-trier.de/db/conf/csw/csw2011.html#WulfRPS11>.
- Markus Zanker. The influence of knowledgeable explanations on users’ perception of a recommender system. In *RecSys*, pages 269–272, 2012.

Eidesstattliche Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmitteln benutzt habe, insbesondere keine anderen als die angegebenen Informationen aus dem Internet. Diejenigen Paragraphen der für mich gültigen Prüfungsordnung, welche etwaige Betrugsversuche betreffen, habe ich zur Kenntnis genommen. Der Speicherung meiner Masterarbeit zum Zweck der Plagiatsprüfung stimme ich zu. Ich versichere, dass die elektronische Version mit der gedruckten Version inhaltlich übereinstimmt.

Siegen, den 24.10.2013

(Malte Paskuda)