

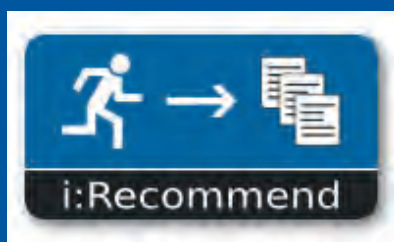
BAND 38

BIBLIOTHEK **B.I.T.** INFORMATION TECHNOLOGIE *online*

INNOVATIV

Analyse von Recommendersystemen in Deutschland

Literaturstudie



INNOVATIONSPREIS 2012

Katrin Gärtner

B.I.T.online – Innovativ

DINGES & FRICK

Band 38

Band 38

Innovationspreis 2012

**Analyse von Recommendersystemen
in Deutschland
Literaturstudie**

2012

Verlag: Dinges & Frick GmbH, Wiesbaden

Innovationspreis 2012

Analyse von Recommendersystemen
in Deutschland
Literaturstudie

von

Katrin Gärtner

2012

Verlag: Dinges & Frick GmbH, Wiesbaden

B.I.T.online – Innovativ

Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

ISBN 978-3-934997-43-1

ISBN 978-3-934997-43-1

ISSN 1615-1577

© Dinges & Frick GmbH, 2012 Wiesbaden

Alle Rechte vorbehalten, insbesondere die des Nachdrucks und der Übersetzung.

Ohne Genehmigung des Verlages ist es nicht gestattet, dieses Werk oder Texte in einem photomechanischen oder sonstigen Reproduktionsverfahren oder unter Verwendung elektronischer Systeme zu verarbeiten, zu vervielfältigen und zu verbreiten.

Satz und Druck: Dinges & Frick GmbH, Wiesbaden

Printed in Germany

Vorwort

B.I.T.online Innovationspreis 2012

Die diesjährigen Preisträgerinnen des B.I.T.online-Innovationspreises, ausgewählt von der Kommission für Ausbildung und Berufsbilder des Berufsverbandes Information Bibliothek (BIB), schrieben ihre Diplom-bzw. Bachelorarbeiten an den Hochschulen in Hamburg, Potsdam und Leipzig.

Die drei mit zukunftsweisenden Aspekten der Bibliotheksarbeit befassten Arbeiten wurden auf dem Hamburger Bibliothekartag 2012 von den ausgezeichneten Absolventinnen vorgestellt; dies geschah in diesem Jahr sehr nah bei den Young Professionals, da das Innovationsforum erstmalig am Stand der Konferenz der informations- und bibliothekswissenschaftlichen Ausbildungs- und Studiengänge (KIBA) ausgerichtet wurde; anschließend erfolgte die Ehrung der drei Preisträgerinnen mit dem B.I.T.-online-Innovationspreis des Jahres 2012.

Das Themenspektrum der prämierten Arbeiten reicht von einer

- Analyse von **Recommendersystemen** in Deutschland (Diplomarbeit – Fachhochschule Potsdam – von *Katrin Gärtner*) über
- Wie können Bibliotheken das PR-Instrument **Storytelling** für ihre PR-Arbeit nutzen: Analyse erfolgreicher Fallbespiele und Ableitung von Handlungsstrategien auf Bibliotheken. (Bachelorarbeit – Hochschule für angewandte Wissenschaften Hamburg – von *Josefine Bäßler*) bis zu
- **Emotion Selling** – ein Impuls für die Kommunikationsgestaltung in Öffentlichen Bibliotheken (Bachelorarbeit – HTWK Leipzig – von *Kerstin Laufs*).

Auf ein vor allem für wissenschaftliche Bibliotheken zukunftsweisendes Thema geht *Katrin Gärtner* in ihrer Diplomarbeit ein, in dem sie Recommendersysteme und deren Anwendung in Deutschland darstellt und die Relevanz der automatisch generierten Empfehlungen dieser Systeme beispielhaft überprüft. Darunter werden heute Softwareanwendungen und -techniken verstanden, die Empfehlungen von Objekten/Gegenständen bzw. mittlerweile vielen Dingen des täglichen Lebens für die Benutzer bereitstellen.

Die Relevanz solcher Recommendersysteme auch im Bibliothekssektor und deren möglicher Mehrwert für die Kunden verdeutlicht – basierend auf umfassendem Literaturstudium – die mit über 100 Seiten Hauptteil umfängliche Abschlussarbeit.

Vor dem Hintergrund des sich ändernden Informationsverhaltens und zeitgleich knapper öffentlicher Mittel greifen die beiden Bachelorarbeiten von **Josefine Bäßler** und **Kerstin Laufs** neue Kommunikationsmittel zur besseren „Vermarktung“ von (öffentlichen) Bibliotheken auf und zeigen mit Storytelling und Emotion Selling mögliche neuartige Wege, das Standing von Bibliotheken in der Öffentlichkeit zu verbessern und damit zugleich ihre Finanzierung im härter werdenden Verteilungskampf auch zukünftig in ausreichendem Maße zu sichern.

Storytelling, das Erzählen von Geschichten, um Kunden emotional zu überzeugen und für die eigenen Produkte und Werte zu begeistern, wird als Kommunikationsmittel von Unternehmen im privatwirtschaftlichen Sektor seit den 90er Jahren erfolgreich eingesetzt. Im öffentlichen Bereich und damit auch in Bibliotheken gibt es hingegen bislang nur wenige Fallbeispiele. **Josefine Bäßler** stellt dar, wie Storytelling, unabhängig von Größe oder Art der Einrichtung, gerade für oftmals mit Imageproblemen kämpfende Bibliotheken Möglichkeiten bietet, mehr Aufmerksamkeit bei Nutzern zu erlangen und neue Kunden zu gewinnen.

Leser- oder Kundengewinnung ist auch das Ziel von Emotion Selling durch Bibliotheken – bezogen auf diese untersucht in der Abschlussarbeit von **Kerstin Laufs**. Kunden nutzen alle Arten von Einrichtungen dann oft und gerne, wenn sie sich in ihnen wohlfühlen. Basierend auf einem dementsprechenden Kommunikationskonzept der Bibliothekseinrichtung kann positive Emotionalisierung zur (verstärkten) Identifikation der Nutzer mit der Bibliothek führen – mit dem beabsichtigten Ziel dass die Inanspruchnahme der Bibliothek als anzustrebender Zustand – extrem formuliert – als „must-have“ erscheint. Dies mit der weiteren Absicht, dass auch potentielle Kunden den Wunsch entwickeln, sich in Bibliotheken aufzuhalten und diese zu nutzen. Auch hier ist ein Blick auf die freie Wirtschaft und deren Werbestrategien geeignet, Anregungen zu geben, wie Bibliotheken sich vorteilhaft im Bewusstsein der Bevölkerung und Behörden verankern können.

Die vollständigen Texte der ausgezeichneten Arbeiten werden 2012 als drei separate Titel in der Buchreihe B.I.T.online – INNOVATIV veröffentlicht.

Karin Holste-Flinspach

(Kommission für Ausbildung und Berufsbilder, Vorsitzende)

Analyse von Recommendersystemen in Deutschland Literaturstudie

Diplomarbeit
zur Erlangung des akademischen Grades
Diplom-Bibliothekarin (FH)
am Fachbereich Informationswissenschaften
der Fachhochschule Potsdam

vorgelegt von:
Katrin Gärtner

Erstgutachter: Prof. Dr. Stephan Büttner
Zweitgutachter: Dr. Jürgen Plieninger
Eingereicht am: 27.06.2011

Abstract

In dieser Arbeit werden in Deutschland vorkommende Empfehlungssysteme analysiert.

Im ersten Teil der Arbeit werden die Grundlagen zu den Empfehlungssystemen, ihre grundlegende Funktionsweise und die zugrundeliegenden Techniken von Recommendersystemen erläutert.

Darauf aufbauend werden Empfehlungssysteme in Deutschland anhand von ausgewählten Internetseiten vorgestellt, analysiert und in die aktuell gültige Klassifikation eingeteilt.

Um beispielhaft die Relevanz von automatisch generierten Empfehlungen zu überprüfen, wird eine stichprobenartige Umfrage unter Wissenschaftlern mehrerer Fachgebiete durchgeführt.

Inhaltsverzeichnis

Abstract	8
Abbildungsverzeichnis	12
Tabellenverzeichnis	15
Abkürzungsverzeichnis	16
1 Prozess der Entscheidungsfindung auf Basis von Empfehlungen	17
1.1 Problemstellung	17
1.2 Zielsetzung und Lösungsansatz	18
1.3 Aufbau der Diplomarbeit	18
2 Klärung von Fachbegriffen	20
2.1 Web 2.0	20
2.2 Katalog 2.0 in Bibliotheken	21
2.3 Mashup	23
2.4 Social Bookmarking Dienste	24
2.5 Social Tagging System	25
2.6 Folksonomie	26
3 Recommendersysteme	28
3.1 Definition	28
3.2 Datengewinnung – implizit/explicit	30
3.3 Vorteile	30
3.4 Nachteile	31
4. Klassifikationsmöglichkeiten von Empfehlungssystemen	33
4.1 Grad der Personalisierung	33
4.2 Berechnungsmethoden	33
4.3 Produktähnlichkeit	34
4.4 aktuelle Forschungsansätze zur Einteilung von Recommender- systemen nach Burke 2011	34
5 Techniken von Recommendersystemen nach Burke (2007)	37
5.1 Content-based Empfehlungssysteme (CB)	37
5.1.1 Definition	37
5.1.2 Vor- und Nachteile von content-based Empfehlungstechniken	40

5.2 Kollaborative Filterung (CF)	41
5.2.1 Definition	41
5.2.2 Vor- und Nachteile von kollaborativer Filterung	44
5.3 Demographische Empfehlungssysteme	46
5.3.1 Definition	46
5.3.2 Vor- und Nachteile	47
5.4 Wissensbasierte Empfehlungssysteme	47
5.4.1 Definition	47
5.4.2 Vor- und Nachteile	48
5.5 Community-basierte Empfehlungssysteme	49
5.5.1 Definition	49
5.5.2 Vor- und Nachteile	52
5.5.3 Exkurs: Einsatz von Social Tagging Systems im Bereich der Recommendersysteme	53
5.5 Hybride Systeme	56
6 Anwendungen von Empfehlungssystemen in Deutschland	60
6.1 klassische Empfehlungssysteme	60
- mündliche Empfehlung	60
- Bestsellerlisten	61
- Bibliographien, Literaturlisten, Literaturverzeichnisse, Publikationslisten	63
6.2 automatisierte Empfehlungssysteme	64
6.2.1 Content-Based Techniken	64
- Spiegel Online	64
- WorldCat.org	69
- Empfehlungen über Podcast am Beispiel der Seite podcast.de	71
- Empfehlung mit Web 2.0 Technik - Empfehlung über RSS	75
6.2.2 Empfehlungssysteme mit kollaborativer Technik	76
- Last.fm	76
- Moviepilot	78
6.2.3 Demographisch/Statistische Empfehlungssysteme in Deutschland	80
- Antolin	80
- Onleihe – virtuelle Bibliotheken	85

6.2.4 Wissensbasierte Empfehlungssysteme in Deutschland	88
- Onlinemodegeschäfte	88
- CHIP Online	91
6.2.5 Community-basierte Empfehlungssysteme	96
- Qype	96
- Reliwa	99
- Lovely Books	101
- Empfehlungen mit Location Based Services	103
- HeyStaks	104
6.2.6 Empfehlungen mit Tags	107
- Delicious	107
- Mister Wong	109
- CiteULike	111
- BibSonomy	114
- LibraryThing	116
6.2.7 Hybride Ansätze bei Empfehlungssystemen	119
- BibTip	119
- Empfehlungselemente bei Suchmaschinen, Beispiel Google	123
- Amazon	126
7 Stichprobe – Bewertung der Relevanz der Empfehlungen bei Amazon und BibTip	134
7.1 Aufbau der Stichprobe	134
7.2 Durchführung der Stichprobe	135
7.3 Auswertung der Stichprobe	137
8. Fazit	141
Literatur- und Quellenverzeichnis	144

Abbildungsverzeichnis

Abbildung 1 - WorldCat als Beispiel für einen Katalog 2.0	22
Abbildung 2 - Taggingmöglichkeit in WorldCat	22
Abbildung 3 - Mashup: eingebundene Karte aus Google Maps auf der Webpage der Hochschule für Musik Würzburg	23
Abbildung 4 - Buchcoverabbildung im Onlinekatalog der Universitätsbibliothek Würzburg	24
Abbildung 5 - Bookmarklets auf Heise Online	25
Abbildung 6 - TagCloud aus LibraryThing	26
Abbildung 7 - Empfehlungsmodell als Blackbox nach Jannach	29
Abbildung 8 - Taxonomy of knowledge sources in recommendation	35
Abbildung 9 - Einfluss von Bereichsfaktoren auf die unterschiedlichen Empfehlungstechniken	36
Abbildung 10 - Genereller Ablauf der content-based Empfehlung nach Färber et al., 2003	38
Abbildung 11 - Ablauf des content-based Filterns nach Schaffert et al., 2010	39
Abbildung 12 - Kollaboratives Filtern bei Emailüberflutung	42
Abbildung 13 - Ablauf der auf CF beruhenden Empfehlungen nach Färber et al. , 2003	43
Abbildung 14 - Weltkarte, die die Verbindungen von Personen bei Facebook sichtbar macht	50
Abbildung 15 - Überblick über die bedeutendsten Forschungsarbeiten zum Thema Hybride Systeme nach Burke	57
Abbildung 16 - Hybrides Design "monolithic" nach Jannach et al.	58
Abbildung 17 - Hybrid Design "parallelized" nach Jannach et al.	58
Abbildung 18 - Hybrides Design "pipelined" nach Jannach et al.	59
Abbildung 19 - Logo der Rubrik Buchempfehlung (vom 01.06.2011) auf der Seite Gründerszene	61
Abbildung 20 - Bestseller aus „Der Spiegel“ 9/2011	62
Abbildung 21 - Startseite Spiegel Online	65

Abbildung 22 - Empfehlung ähnlicher Artikel auf Spiegel Online	66
Abbildung 23 - Hinweis auf verwandte Themen zum Artikel	66
Abbildung 24 - kommerzielle Empfehlungen auf Spiegel Online	67
Abbildung 25 – Möglichkeiten auf Spiegel Online einen Artikel weiterzuempfehlen	68
Abbildung 26 - Startseite WorldCat	69
Abbildung 27 - Trefferanzeige Worldcat	70
Abbildung 28 - TagCloud zu Rebecca Gable aus WorldCat Identities	71
Abbildung 29 - Ansicht eines Titels auf Podcast.de	72
Abbildung 30 - Startseite Podcast.de	74
Abbildung 31 - Datei:Feed-icon.svg (Originally distributed by the Mozilla Foundation under a MPL/GPL/LGPL tri-license)	75
Abbildung 32 - Onlinekatalog der Universität Karlsruhe – RSS	76
Abbildung 33 - Bewertungsbuttons Herz und Verbotsschild bei Last.fm	77
Abbildung 34 - Last.fm Profilseite mit Empfehlungen	78
Abbildung 35 - Filmempfehlung bei moviepilot.de	79
Abbildung 36 - Startseite Antolin	81
Abbildung 37 - Buchbeschreibung mit Bewertung und Empfehlung von weiteren Titeln	82
Abbildung 38 - Empfehlungen für Jungen/Mädchen	83
Abbildung 39 - Explizite Eingabe der Buchbewertung bei Antolin	84
Abbildung 40 - Verdeckte Werbung für kostenpflichtige Lernseiten	84
Abbildung 41 - Startseite der Virtuellen Stadtbücherei Würzburg	86
Abbildung 42 - explizite Empfehlungsfunktion bei der Onleihe	87
Abbildung 43 - Online-Shop von Esprit	88
Abbildung 44 - Suche nach einer Espritfiliale in der Nähe	90
Abbildung 45 - Mashup- Storefinder von Esprit	90
Abbildung 46 - Internetseite Chip-Online	92
Abbildung 47 - Shop-Empfehlung auf Basis redaktioneller Artikel zu Produkten	93
Abbildung 48 - explizite Bewertungsmöglichkeit für Downloads	95

Abbildung 49 - Empfehlungen für Frankfurt	98
Abbildung 50 - Medienbeschreibung bei Reliwa	100
Abbildung 51 - Buchfrage Widget von LovelyBooks	102
Abbildung 52 - Buchfrage	102
Abbildung 53 - Googlesuche mit HeyStaks-Toolbar	105
Abbildung 54 - Die Architektur des HeyStaks-Systems und ein Grundriss des Empfehlungsmodells	106
Abbildung 55 - Google-Suchergebnis mit integrierten HeyStaks Empfehlungen	106
Abbildung 56 - Tags bei Delicious	108
Abbildung 57 - Tag-Suche bei Delicious	109
Abbildung 58 - Startseite Mister Wong	110
Abbildung 59 - Suche auf Mister Wong	110
Abbildung 60 - CiteULike.org - Ergebnis der Suche nach Recommender System	112
Abbildung 61 - Neun mögliche Empfehlungsarten bei CiteULike nach Bogers	113
Abbildung 62 - Empfehlungen bei CiteULike	114
Abbildung 63 - Suchergebnis bei BibSonomy	115
Abbildung 64 - LibraryThing Lokale Suche	117
Abbildung 65 - gefundener Titel auf LibraryThing	118
Abbildung 66 - KIT-Katalog mit BibTip-Empfehlung	120
Abbildung 67 - Architektur von BibTip	122
Abbildung 68 - Google - Vorschläge für evtl. gemeinte Suchwörter	123
Abbildung 69 - Google - Verwandte Suchbegriffe	123
Abbildung 70 - content-based Elemente bei Google	124
Abbildung 71 - Google Maps – Routenvorschläge	125
Abbildung 72 - Empfehlung des kooperativen Recommendersystems von Amazon	127
Abbildung 73 - Empfehlungen des nutzerspezifischen Recommendersystems von Amazon	128

Abbildung 74 - Versuchte Transparenz der Empfehlung - Ansicht der beobachteten Objekte	128
Abbildung 75 - Abgabe von expliziter Bewertung bei Amazon	129
Abbildung 76 - Explizite Bewertung bei Amazon über den „Gefällt mir“-Button	130
Abbildung 77 - Rezension bei Amazon zu Klahold: Empfehlungssysteme	130
Abbildung 78 - Tags zum gleichen Artikel	131
Abbildung 79 - Funktionsweise des Amazon-Algorithmus nach Linden et. al.	132
Abbildung 80 - Screenshot aus der Umfrage für ein Medium mit Empfehlungen von BibTip	135
Abbildung 81 - Screenshot aus der Umfrage für den gleichen Titel wie Abb. 80 mit Empfehlungen aus Amazon	136
Abbildung 82 - Ausschnitt aus der Befragung - Ratingskala und offene Ergänzungsaufgabe	136

Tabellenverzeichnis

Tabelle 1 - Schlüsselforschungsansätze Empfehlungen mit sozialen Netzen nach Arrazy	52
Tabelle 2 - Häufigkeit der vergebenen Relevanzstufen bei Amazon und BibTip	138
Tabelle 3 - Durchschnittliche Bewertung der Relevanz bei Amazon und BibTip	138

Abkürzungsverzeichnis

CB/CN	content-based Filterung
CF	kollaborative Filterung
DM	demographische Filterung
KB	Knowledge/wissensbasierte Filterung
RS	Recommendersystem
RSWK	Regeln für den Schlagwortkatalog
STS	Social Tagging System

1 Prozess der Entscheidungsfindung auf Basis von Empfehlungen

1.1 Problemstellung

Die erste nachgewiesene Bibliothek, die Bibliothek Assurbanipals (668-627 v.Chr.), die in dessen Residenzstadt Ninive stand, wird auf eine Größe von 5.000-10.000 Tontafeln, die mit Keilschrift beschrieben waren, geschätzt.¹

Am 08. Juni 2011 berichtet die Frankfurter Allgemeine Zeitung vom erfolgreichen Test-Tag für das Internetprotokoll IPv6, da die 4,3 Milliarden Internetadressen, die seit 1973 vergeben wurden, nicht ausreichen.²

Die Anzahl der Informationen, die einem Menschen zur Verfügung stehen, ist rasant angewachsen. Im Jahr 2009 ist die Anzahl der Informationen im „Digitalen Universum“ über 62 Prozent auf annähernd 800.000 Petabyte (ein Petabyte sind eine Millionen Gigabyte) angewachsen. Bei dieser steigenden Wachstumsrate ist das Internet bis 2020 44mal größer als im Jahr 2009.³

Um bei der Vielzahl an vorhandenen Informationen nicht überfordert zu sein, werden Selektionskriterien benötigt, die den „Information Overload“ filtern, sonst lassen sich die folgende Fragen: Welches der vielen Bücher zum Thema xy ist das relevanteste / beinhaltet die von mir gesuchte Information, welchen Film kann / sollte man anschauen bzw. welche Produkte soll man kaufen bzw. welche würden zu den bereits gekauften Produkten passen, nur schwer beantworten.

Wenn die informationssuchende Person⁴ selbst nicht über ausreichendes Wissen verfügt, um eine begründete Entscheidung zu treffen, kann sie als Selektionskriterium Empfehlungen von anderen Personen heranziehen.

Oftmals müssen die Entscheidungen allerdings in kürzester Zeit getroffen werden. Hierbei ist nicht immer ausreichend Zeit, Empfehlungen von Personen des Vertrauens einzuholen und diese zu befragen. Aufgrund der Schnelllebigkeit, die auch durch die steigende Digitalisierung der Welt hervorgeru-

¹ Jochum 2007, S. 14–15

² Schmidt 08.06.2011

³ Gantz, Reintzel Mai 2010

⁴ Aus Gründen der Vereinfachung und der besseren Lesbarkeit wird in dieser Arbeit durchgehend die männliche Form verwendet.

fen wird, reichen die Empfehlungen durch persönliche Kommunikation von Mensch zu Mensch nicht mehr aus.

Diese Funktion bieten Recommendersysteme, die manchmal auch als Vorschlagssysteme oder Empfehlungssysteme bekannt sind.

1.2 Zielsetzung und Lösungsansatz

Die Aufgabe der vorliegenden Arbeit besteht darin, die Grundlagen und Funktionsweisen von Recommendersystemen, anhand einer Literaturstudie, vorzustellen. In der aktuellen Literatur sind bereits erste Beiträge zu Recommendersystemen zu finden. Dabei werden meist nur Teilbereiche behandelt, wobei jüngst erstmals ein umfassendes Handbuch Ricci (2011)⁵ erschienen ist. Darüber hinaus sollen die aktuelle Situation und die Anwendungen von Empfehlungssystemen in Deutschland erforscht werden, da hierüber noch keine Darstellungen vorliegen. Dazu werden unterschiedliche Beispiel-Internetseiten mit integrierten Empfehlungstechniken betrachtet und analysiert.

In diesem Zusammenhang lautet die zentrale Fragestellung dieser Arbeit:

Welche Arten von Recommendersystemen gibt es in Deutschland und wie ist ihre aktuelle Anwendung?

Mithilfe einer stichprobenhaften Umfrage unter Wissenschaftlern soll die Relevanz von Empfehlungen bei zwei ausgewählten Empfehlungssystemen, Amazon und BibTip, überprüft werden.

1.3 Aufbau der Diplomarbeit

Die Arbeit ist in acht Kapitel gegliedert, die aufeinander aufbauend mit jedem Kapitel eine Spezialisierung bzw. Weiterführung der aus den vorangegangenen Kapiteln gewonnenen Erkenntnisse darlegen.

In Kapitel 2 werden einige Begriffe aus dem Bereich des Web 2.0 und der Sozialen Webanwendungen erklärt, für die es noch keine allgemeingültigen Begriffsdefinitionen gibt; daher liefert dieses Kapitel die Grundlage für die folgende Nutzung in der Arbeit.

⁵ Ricci 2011

Kapitel 3 gibt eine Erklärung des Begriffs „Recommendersysteme“ und zeigt deren allgemeine Vor- und Nachteile auf.

Nach der Vorstellung von verschiedenen Möglichkeiten Recommendersysteme einzuteilen (Kapitel 4), wird anhand der Klassifikation nach Burke (2007)⁶ (Kapitel 5) auf die einzelnen Formen von Vorschlagssystemen eingegangen.

In einem Exkurs (Kapitel 5.5.3) werden die Social Tagging Systeme vorgestellt, da diese in letzter Zeit an Bedeutung zugenommen haben.

Im darauf folgenden Kapitel 6 werden beispielhaft einzelne Empfehlungssysteme, die in Deutschland angeboten werden, analysiert und versucht, diese in die gängige Klassifikation einzuteilen.

Anschließend wird mithilfe einer Stichprobe (Kapitel 7) geklärt, wie Wissenschaftler die Relevanz von Empfehlungen durch den kommerziellen E-Shop-Anbieter Amazon im Vergleich zu den Empfehlungen des Bibliotheksempfehlungssystems BibTip bewerten.

Die Arbeit schließt in Kapitel 8 mit einer Zusammenfassung sowie einem kurzen Ausblick auf weiterführende Forschungsarbeiten.

⁶ Burke 2007

2 Klärung von Fachbegriffen

Die wichtigsten Fachtermini aus dem Umfeld von Empfehlungssystemen werden in diesem Kapitel vorgestellt. Für diese Begriffe gibt es meist noch keine allgemeingültige Definition. Deswegen werden sie in diesem Kapitel definiert und diese Definitionen sollen als Grundlage für die weiteren Kapitel dienen.

2.1 Web 2.0

Der Begriff „Web 2.0“ ist ein sehr häufig verwendeter Begriff. „Die Bezeichnung "2.0" ist aus der Software-Branche entlehnt, in der die Versionen von Computerprogrammen im Zuge ihrer Weiterentwicklung durchnummeriert werden und ein Sprung auf eine neue Versionsstufe in der Regel mit sehr grundlegenden Veränderungen in den Funktionen und im Design einhergeht.“⁷ Im Bereich des Internets bedeutet der Begriff also eine tiefgreifende Veränderung gegenüber dem früheren Verständnis und der früheren Benutzung. So werden die Veränderungen in einer Reihe von technischen, ökonomischen und gesellschaftlichen Entwicklungen sichtbar, die das heutige Internet prägen.

O'Reilly Media, die den Begriff für eine Konferenzreihe kreierten, definiert Web 2.0 so:

„Web 2.0 is the network as platform, spanning all connected devices; Web 2.0 applications are those that make the most of the intrinsic advantages of that platform: delivering software as a continually-updated service that gets better the more people use it, consuming and remixing data from multiple sources, including individual users, while providing their own data and services in a form that allows remixing by others, creating network effects through an architecture of participation, and going beyond the page metaphor of Web 1.0 to deliver rich user experience.“⁸

Durch Web 2.0 wurden und werden die Möglichkeiten geschaffen, Inhalte für andere zugänglich zu machen, zu bearbeiten und weiter zu verbreiten. Beispiele hierfür sind Wikis, Weblogs, YouTube usw.

⁷ Hochschulrektorenkonferenz (Hg.) 2010, S. 11

⁸ O'Reilly 2005

Somit entsteht Web 2.0 dann, wenn der Benutzer nicht mehr nur Anwender ist, sondern interaktiv und kollaborativ an der Gestaltung des „Netzes“ mitarbeitet.

2.2 Katalog 2.0 in Bibliotheken

In Bibliotheken stehen die Kerndienstleistungen Bestandsaufbau und Erschließung des Bestandes im Zentrum. „ [Dabei gibt es] zwei Kerndienstleistungen einer Bibliothek. Die eine ist der Bestand und sein Profil, gemeinhin Bestandsaufbau genannt. Die andere ist die Erschließung des Bestandes“⁹ in einem Katalog. Früher erfolgte diese Erschließung in einem Zettelkatalog, heute ist dieser als „Online Public Access Catalog“ im Internet suchbar.

„Katalog 2.0 bezeichnet sogenannte "next generation catalogs", in denen Elemente des Web 2.0 implementiert und die auf die Prinzipien der Bibliothek 2.0 ausgerichtet sind.“¹⁰

Solche Funktionen können sein, dass man sich Titel in Merklisten speichern und nach Anmeldung gruppieren kann oder das Veröffentlichen dieser Titelliste für andere, indem man sie mit ihnen teilt und/oder sogar gemeinsam in Projektgruppen diese bearbeitet. Eine weitere Möglichkeit, die der Katalog 2.0 bietet, ist das Abonnieren von RSS-Feeds z. B. zu speziellen Autoren, Systematikstellen oder Schlagworten.

⁹ Drauz, Plieninger 2010, S. 40

¹⁰ Kneifel 2010, S. 38

inhalt (Dokumente, Bilder und Videos) im eigenen Browser gespeichert werden. Das Wiederfinden der gespeicherten Links wird mit inhaltlicher Erschließung z.B. durch Kategorien und Tags ermöglicht.



Abbildung 5 - Bookmarklets auf Heise Online¹⁸

Durch das individuelle Abspeichern und Ordnen der eigenen Quellensammlung entstehen soziale Metainformationen, wie z.B. die Frage, in wie vielen persönlichen Sammlungen ein Aufsatz gespeichert ist? Welcher Aufsatz verbreitet sich bei welchen Benutzern? Welche Art von Verweisen wird benutzt? Diese Metainformationen sind keine direkten Nutzerbewertungen wie in Recommendersystemen, sondern entstehen aufgrund von sozialen Beziehungen.

2.5 Social Tagging System

Unter dem Englischen Begriff „to tag sth. to sth.“¹⁹ ist laut Langenscheidt die „Auszeichnung oder Etikettierung eines Gegenstandes“ zu verstehen.

Im Bereich der Informationswissenschaften wird unter dem Begriff „Taggen“ die Vergabe von freien, regellosen Schlagworten verstanden. „Getaggt“ werden kann alles: Bücher, Filme, Videos, Musik, Fotos, Bookmarks usw.

Beim „Social Tagging“ wird dieser Prozess gemeinschaftlich von vielen Anwendern gemeinsam verfolgt. Diese beschreiben ihre eigenen oder fremden Ressourcen (z.B. Kaufrezensionen) mit ihren persönlichen Worten und machen sie anderen Nutzer zugänglich; man spricht auch vom „Kollaborativen Indexieren“. Die Verschlagwortung durch viele bietet eine gute inhaltliche Erschließung, die eine Verbesserung der Suchergebnisse ermöglicht. Sie ist allerdings im Vergleich zu einer „bibliothekarischen“ Erschließung ungeordneter und breiter gefächert, dafür näher an der Alltagssprache.

Die Probleme, die beim „Social tagging“ auftreten, werden von Balby (2011) beschrieben: „[...] social tagging can present problems, such as tag sparsity

¹⁸ Heise Online 2011: <http://heise.de/-1218777>

¹⁹ Sinclair-Knight 2008

3 Recommendersysteme

In diesem Kapitel sollen die Grundbegriffe und -funktionen von Recommendersystemen erläutert werden. So wird auf die implizite und explizite Datengewinnung eingegangen, bevor der Nutzen von Empfehlungssystemen durch Abwägen der Vor- und Nachteile vorgestellt wird. Im darauffolgenden Kapitel 4 wird dann auf die verschiedenen Klassifizierungsansätze von Empfehlungssystemen eingegangen.

3.1 Definition

Resnick/Varian²⁵ haben im Jahr 1997 den Begriff „Recommendersystem“ (im Deutschen spricht man auch von „Empfehlungssystemen“ oder „Vorschlagsystemen“) geprägt.

Bereits davor gab es Systeme (z.B. Tapestry), die automatische Empfehlungen ausgaben, allerdings wurde der Vorgang hier noch als „collaborative filtering“ bezeichnet. Hierzu beschreiben Resnick und Varian:

„We prefer the more general term ‚recommender system‘ for two reasons. First, recommenders may not explicitly collaborate with recipients, who may be unknown to each other. Second, recommendations may suggest particularly interesting items, in addition to indicating those that should be filtered out.“²⁶

Unter Recommendersystemen werden heute Softwareanwendungen und -techniken verstanden, die Empfehlungen von Objekten/Gegenständen bzw. mittlerweile allen Dingen des täglichen Lebens für den Benutzer bereitstellen.

„Empfehlungssysteme (Recommender Systems) dienen dazu, ihren Nutzern bestimmte Produkte zu empfehlen. Aufgrund eines zuvor erstellten Nutzerprofils sind dessen Vorlieben für bestimmte Produkte bekannt, außerdem können weitere Produkte ermittelt werden, die für ihn von Interesse sein könnten.“²⁷

²⁵ Vgl. Resnick, Varian 1997

²⁶ Resnick, Varian 1997, S. 56

²⁷ Meier, Stormer 2007, S. 1455

Anwendungsgebiete sind nach Klahold (2009)²⁸ die Empfehlungen von Inhalten wie z.B. Produkte, Texte, Bilder, Audioempfehlungen. Es werden aber auch Prozesse z.B. Fahrtrouten oder Email-Spamfilter und auch Personen wie z. B. Experten von Recommendersystemen empfohlen.

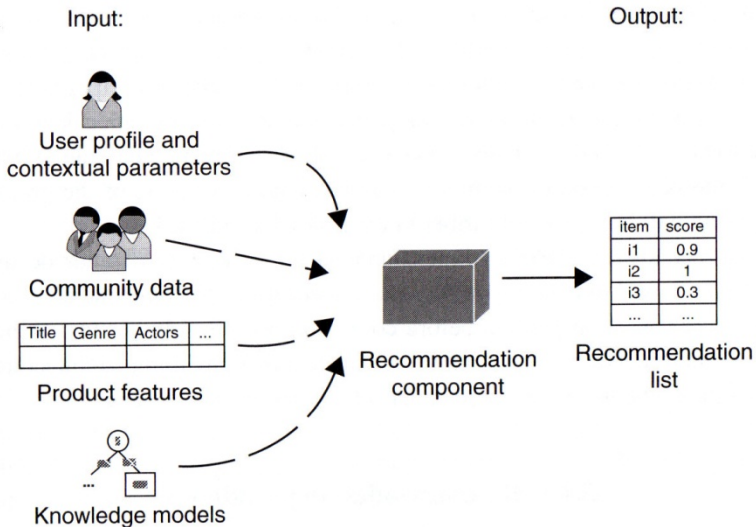


Abbildung 7 - Empfehlungsmodell als Blackbox nach Jannach²⁹

In Abbildung 7 wird das Recommendersystem als eine „blackbox“, die Eingabedaten in eine gewertete Objektausgabeliste umwandeln und ausgeben kann, dargestellt.

Die potentiellen Eingabedaten sind hierbei: Benutzermodelle, textabhängige Informationen, Gemeinschaftsdaten, Produktdaten und Wissensmodelle, diese werden mithilfe der Empfehlungsverfahren umgewandelt.

²⁸ Vgl. Klahold 2009, S. 4

²⁹ Jannach 2011, S. 125

4. Klassifikationsmöglichkeiten von Empfehlungssystemen

Empfehlungssysteme lassen sich auf verschiedene Weise klassifizieren. Hier werden die verschiedenen Ansätze zur Klassifikation vorgestellt, bevor der aktuelle Forschungsansatz nach Burke 2011³⁸ für die Einteilung von Recommendersystemen erläutert wird. Im nachfolgenden fünften Kapitel werden die einzelnen Formen von Empfehlungssystemen anhand der (im Moment) allgemein anerkannten Klassifikation nach Burke 2007³⁹ veranschaulicht.

4.1 Grad der Personalisierung

So können Vorschlagssysteme danach gegliedert werden, ob die Empfehlungen personalisiert oder nichtpersonalisiert erfolgen:

Bei der personalisierten Empfehlung wird für jeden Nutzer eine individuelle Empfehlung ermittelt und abgegeben.

Eine nichtpersonalisierte (anonyme) Empfehlung ist für alle Nutzer identisch. Ein Beispiel sind die Buchempfehlungen von Amazon, die mit dem Satz „Kunden, die dieses Buch gekauft haben, haben auch folgendes Buch gekauft“ eingeleitet werden. Jeder Nutzer erhält dieselben Empfehlungen.

4.2 Berechnungsmethoden

Es gibt auch Gliederungsansätze von Recommendersystemen nach der Berechnungsmethode selbst.

So erfolgt bei memorybasierten Ansätzen die gesamte Ähnlichkeitsberechnung unmittelbar bei der Anfrage. Die Bewertungsmatrix muss im Arbeitsspeicher vorhanden sein (daher „memorybasiert“).

Bei den modellbasierten Ansätzen wird dagegen in einem ersten Schritt eine Vorberechnung vorgenommen. Bei einer Anfrage kann auf diese Ergebnisse

³⁸ Burke, Ramezani 2011

³⁹ Burke 2007

5.2 Kollaborative Filterung (CF)

5.2.1 Definition

Die ursprünglichste und einfachste Art, Empfehlungen zu generieren, ist der Ansatz des kollaborativen Filterns (CF)⁵². Hierbei werden Verhaltensmuster von Benutzergruppen ausgewertet, um auf die Interessen des Einzelnen zu schließen.

Die Ähnlichkeit bei den Vorlieben wird aufgrund einer Ähnlichkeit in der Bewertungsgeschichte der Nutzer, der „Nachbarschaft“ berechnet. Aus diesem Grund wird das CF auch als „Mensch-zu-Mensch-Korrelation“ bezeichnet.

Das erste bekannte System, das mit kollaborativer Filterung arbeitete, war Tapestry von Goldberg, Nichols, Oki und Terry (1992). Sie erklären kollaborative Filterung wie folgt:

„Collaborative filtering simply means that people collaborate to help one another perform filtering by recording their reactions to documents they read. Such reactions may be that a document was particularly interesting (or particularly uninteresting). These reactions, more generally called annotations, can be accessed by others' filters.“⁵³

Sie veranschaulichten das kollaborative Filtern am Beispiel von Emailüberflutung, da solche Annotationen dazu genutzt werden können, um Spam zu identifizieren. Wenn viele Empfänger diese Email als Spam markiert haben, kann sie bei allen weiteren ausgefiltert werden (Abbildung 12).

„A pure collaborative recommendation system is one which does no analysis of the items at all - in fact, all that is known about an item is a unique identifier. Recommendations for a user are made solely on the basis of similarities to other users.“⁵⁴

Ein rein kooperatives Empfehlungssystem, wie von Balabanović und Shoham (1997)⁵⁵ beschrieben, abstrahiert völlig vom Inhalt der Dokumente und arbeitet ausschließlich mit den Ähnlichkeiten des aktuellen Nutzers mit anderen Benutzern.

⁵² Im Deutschen wird es manchmal auch als kooperatives System bezeichnet.

⁵³ Goldberg u. a. 1992, S. 61

⁵⁴ Balabanović, Shoham 1997, S. 67

⁵⁵ Balabanović, Shoham 1997

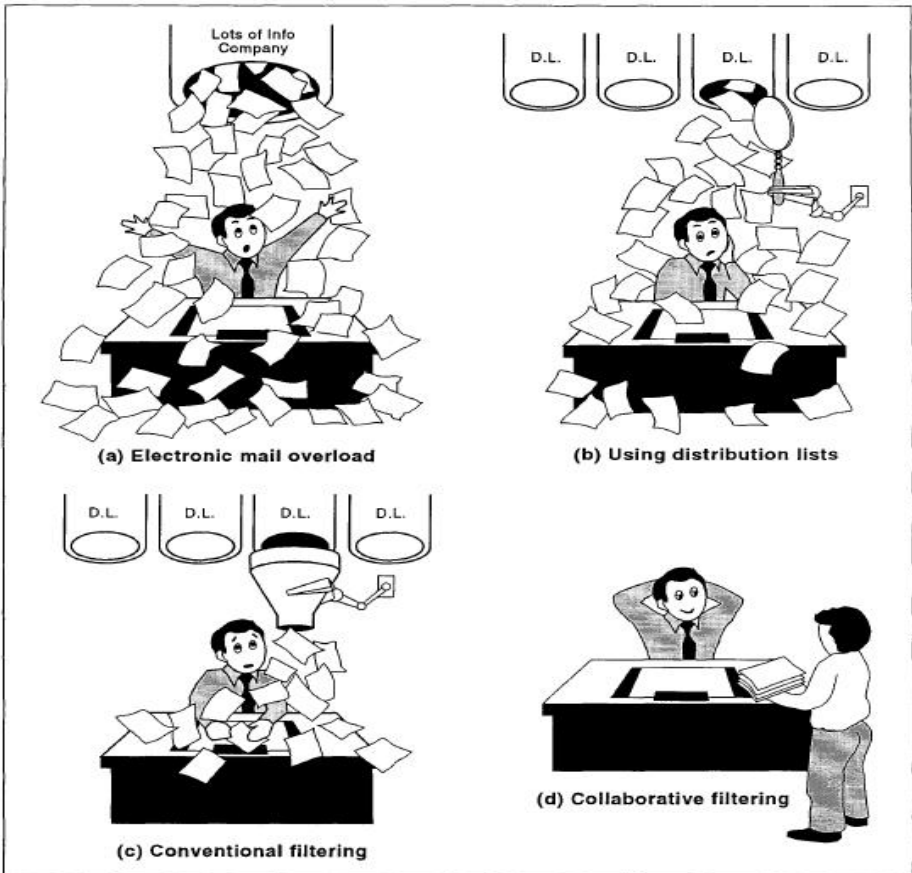


Abbildung 12 - Kollaboratives Filtern bei Emailüberflutung⁵⁶

Dieser Prozess der Nachbarschaftsbildung, Ähnlichkeitsberechnung und die dazugehörigen Algorithmen sind der Mittelpunkt der Empfehlungen, die auf CF beruhen (Abbildung 13).

⁵⁶ Goldberg u. a. 1992, S. 62

5.3.2 Vor- und Nachteile

Die Vorteile von demographischen Empfehlungssystemen sind gleich denen von kollaborativen Filtern. Es können Empfehlungen über Genre Grenzen hinweg gegeben werden, und es ist ebenfalls kein Bereichswissen nötig, da „Mensch-zu-Mensch-Empfehlungsdaten“ verglichen werden. Außerdem verbessert sich die Qualität des Systems bei längerer Laufzeit, da diese Systeme lernfähig sind. Einzig das beim CF genutzte implizite Feedback fällt bei demographischen Systemen weg, da der Benutzer selbst seine persönlichen Daten eingibt.

Die Nachteile sind laut Burke (2002)⁶⁴ das „neue Benutzer“-Problem, das „Grey-Sheep“-Problem, die Abhängigkeit der Qualität von den großen Altdatenbeständen, sowie die schwere Veränderbarkeit des Benutzerprofils. Der gewichtigste Nachteil ist allerdings, dass der Benutzer persönliche Daten über sich preisgeben muss.

5.4 Wissensbasierte Empfehlungssysteme

5.4.1 Definition

Höfeld/Kwiatkowsk (2007)⁶⁵ benutzen die Erklärung von Burke (2000)⁶⁶, um die Funktionsweise von „Knowledge-based“, zu Deutsch „wissensbasierte Empfehlungssysteme“, zu beschreiben:

„Wissensbasierte Empfehlungssysteme ziehen Schlussfolgerungen, um aus den Präferenzen der Nutzer funktionales Wissen über die Korrelation der Bedürfnisse des Nutzers und bestimmter Objekte abzuleiten. Das funktionale Wissen solcher Systeme ermittelt bspw. den Nutzen eines bestimmten Produkts für eine Person.“⁶⁷

Die Empfehlungen werden unabhängig von den Bewertungen eines Benutzers berechnet, entweder in Form von Ähnlichkeiten zwischen den Bedürfnis-

⁶⁴ Burke 2002

⁶⁵ Höfeld, Kwiatkowski 2007

⁶⁶ Burke 2000

⁶⁷ Höfeld, Kwiatkowski 2007

6 Anwendungen von Empfehlungssystemen in Deutschland

In diesem Kapitel werden beispielhaft Anwendungen von Recommendersystemen in Deutschland und ihre Techniken vorgestellt. Die Einteilung weicht von Burke ab, da auch die „klassischen“ Empfehlungen, wie z.B. die mündliche Mensch-zu-Mensch-Empfehlung, in diesem Kapitel ihren Platz finden sollen, sie sind im „Offline-Bereich“ immer noch gültig.

6.1 klassische Empfehlungssysteme

Unter „klassische“ Empfehlungssysteme fallen die Empfehlungen, die es schon immer gab und die auch ohne technische Hilfe und ohne Internet gegeben werden. So haben wir schon immer den Empfehlungen von Menschen, die wir kennen, oder den gedruckten Rezensionen vertraut. Zu diesen „klassischen“ Empfehlungen gehören die mündliche Empfehlung, Bestsellerlisten und Bibliographien.

- mündliche Empfehlung

Die mündliche Empfehlung von Mensch zu Mensch findet auch heute noch immer Anwendung. So fragt man z.B. beim Einchecken im Hotel, welches Restaurant zu empfehlen ist oder im Freundeskreis, welcher Club „in“ ist oder die freundliche Bibliothekarin in der Bibliothek, welchen Krimi (nicht zu blutrünstig, intelligenter Ermittler) sie uns vorschlagen kann. Bei dieser Empfehlungsart kann generell alles empfohlen werden, Inhalte, Prozesse oder auch Personen.

Aber auch diese Empfehlungsart hat sich bereits ins Internet verlagert, so finden sich bei der Suche in Google¹¹⁴ nach „Buchempfehlungen“ mehr als 1.360.000 Treffer. Dabei sind nicht nur Treffer von Buchhändlern und Autoren,¹¹⁵ sondern auch Webseiten wie Gründerszene¹¹⁶, bei der wöchentlich

¹¹⁴ Google: <http://www.google.de>

¹¹⁵ Leseberater: <http://www.leseberater.de>

¹¹⁶ Gründerszene: <http://www.gruenderszene.de>

wechselnd Werbeunternehmer Buchempfehlungen für Geschäftsneugründer geben, zu finden.



Abbildung 19 - Logo der Rubrik Buchempfehlung (vom 01.06.2011) auf der Seite Gründerszene¹¹⁷

- Bestsellerlisten

Bei den „klassischen“ Empfehlungsarten spielen die Bestsellerlisten oder „Top Ten“-Listen eine große Rolle.

Bestsellerlisten finden sich in jedem Lebensbereich, z.B. bei Autos ist es die Pannenstatistik¹¹⁸, im Bereich der Musik sind es die Charts¹¹⁹, im Filmbereich die Kinocharts¹²⁰, im Bereich Ausgehen & Essen gibt es für jede größere deutsche Stadt die Restaurantbewertungen, z.B. für Frankfurt¹²¹ usw. Auch hier kann und wird generell keine Unterscheidung nach dem Anwendungsgebiet der Empfehlung gemacht, es kann alles empfohlen werden.

Will man wissen, welche Bücher im Moment am gefragtesten sind, so schaut man in der Spiegel-Bestsellerliste, einer Liste, die jene Bücher aufzeigt, die über einen bestimmten Zeitraum am häufigsten verkauft wurden. Diese Liste wird von „Buchreport“, dem Netzwerk der deutschen Buchhändler, bereits seit 1971 wöchentlich erstellt und erscheint in der Zeitschrift „Der Spiegel“.

¹¹⁷ Gründerszene: <http://www.gruenderszene.de>

¹¹⁸ ADAC: <http://www.adac.de/infotestrat/unfall-schaeden-und-panne/pannenstatistik/default.aspx?ComponentId=73430&SourcePageld=49424>

¹¹⁹ ChartKing: <http://www.charts-weltweit.de>

¹²⁰ Zelluloid: <http://www.zelluloid.de/charts/>

¹²¹ Genussmagazin: <http://www.genussmagazin-frankfurt.de/>

Verlag Dinges & Frick GmbH, Wiesbaden

ISBN 978-3-934997-43-1

ISSN 1615-1577

€ 24,50