



INSTITUT FÜR INFORMATIK

Dissertation

**Maschinelle Generierung von Empfehlungen zur
Lehr-/Lernunterstützung im Hochschulkontext**

vorgelegt durch

Benedikt Engelbert

Dezember 2016

Dissertation zur Erlangung des Doktorgrades (Dr. rer. nat.)

des Fachbereichs Mathematik und Informatik

der Universität Osnabrück

Betreuer: Prof. Dr. Oliver Vornberger

Zusammenfassung

Das Internet, in dem Informationen schnell und komfortabel abgerufen werden können, hat das Leben im Informationszeitalter nachhaltig verändert. Die Vorteile eines einfachen Zugriffs auf Daten, Dienste und Informationen sind dabei genauso greifbar wie die stetig anwachsende Menge an Angeboten, die das Internet in gleicher Weise vielfältig und unübersichtlich gestalten. Die Auswahl geeigneter Angebote ist für Benutzer ein häufiges Problem, für welches sich unterstützende Systeme unter dem Namen der Empfehlungsdienste etabliert haben. Empfehlungsdienste helfen bei der Selektierung von Informationen auf der Grundlage von Benutzerpräferenzen oder Bedürfnissen in einer Vielzahl von Anwendungskontexten. Online Versandhandel, in denen Empfehlungsdienste Vorschläge für interessante oder nützliche Artikel aussprechen, ist eines der populärsten und gängigsten Anwendungsszenarien. Aber auch das Lernen wurde durch das Internet in den letzten Jahren stark geprägt. Ein Lernmanagementsystem (LMS) ist im universitären Kontext mittlerweile ein gängiger Standard, wodurch der Prozess der Distribution und des Zugriffs auf digitale Lernmaterialien stark vereinfacht wurde. Über das Internet finden Lernende die Möglichkeit ihr Repertoire an Lernmaterialien zu erweitern, was sich angesichts der Menge an verfügbaren Materialien, sowie der Komplexität an Lerninhalten im universitären Kontext nicht notwendigerweise als Vorteil herausstellt. Ein Lernender sieht sich einem schwer zu sichtenden Überangebot ausgesetzt, welches sich im ungünstigsten Fall auf die Motivation oder Leistung auswirken könnte. In jedem Fall bedarf es bei einer Suche Zeit, um eine genügende Sichtung durchführen zu können, obgleich im Regelfall bereits Lernmaterialien durch die/den Lehrende/n zur Verfügung stehen. Es wird an dieser Stelle der Bedarf an unterstützenden Systemen im Kontext des digitalen Lernens bzw. E-Learning gesehen, die eine assistierende Rolle bei der Auswahl an Inhalten und Materialien einnehmen. Die vorliegende Arbeit untersucht die Entwicklung eines maschinellen Ansatzes zur automatisierten und inhaltsbezogenen Herleitung von Empfehlungen, die kontextrelevante Inhalte in Lernmaterialien aufdecken und mit alternativen Materialien verknüpfen. Hierfür präsentiert die Arbeit die Entwicklung eines kollaborativen Tagging Ansatzes, um den Zielvorgaben zu genügen. Das entwickelte System verfolgt die Intention die Nutzung von bestehenden Materialien Lehrender zu erleichtern und des Weiteren den Prozess des Auffindens relevanter, alternativer Lernmaterialien zu vereinfachen. Die einfache Integration in die Lehre steht bei der Entwicklung ebenfalls im Fokus, so dass ein Mehraufwand für Lehrende vermieden wird und sich möglichst Mehrwerte durch hergeleitete Informationen des Systems ergeben. Die Arbeit beschreibt zudem

die Evaluation des Systems, die in zwei Evaluationsszenarien durchgeführt wurde und diskutiert die Ergebnisse im Zusammenhang mit vergleichbaren Ansätzen.

Abstract

The Internet where information can be retrieved fast and comfortably has changed peoples lives within the information age. The advantages which came along with the easy access to data, services and information are as present as the continuous growth of offerings. Therefore the internet can be equally diverse and overwhelming for a user, which can be difficult for the right selection regarding the user's needs. Recommender systems help users to identify the most appropriate content on the basis of user preferences and needs. These systems exist for a variety of applications. One of the most popular application fields are online mail order services, where the system generates suggestions for products the user can purchase. Also the way of learning has changed evidently with the increasing usage of the internet. Learning Management Systems (LMS) at universities are widespread whereby distributing and consuming of Learning Materials (LM) became easy. From a student's point of view the access to LM provided by a course teacher is straightforward. On the internet students have access to further LM which can be used in addition. But using the internet to find new learning material is not necessarily successful. With the variety of available LM and its complexity, a student needs time to examine the material in order to make a proper selection. At worst the retrieval leads to demotivation or has negative effects on students learning process. There is a demand for assistive systems in the area of digital learning, to support students to choose appropriate content or learning material. The following thesis examines the development of a machinable approach to deduce automatic and content-based recommendations. These recommendations reveal relevant content in available LM and connect them to further appropriate material from the internet. For this, the thesis presents a collaborative tagging approach to fulfill the desired purposes. The intention of the developed system is to encourage and facilitate the use of available course material and to simplify the retrieval of additional material from the internet. A main focus of the system is the easy implementation in learning scenarios to avoid extra time for teachers to setup a certain system. At best a teacher shall benefit equally as the students from the information collected by the system. Furthermore, the thesis presents two types of user studies to evaluate the system. The results are discussed with current research approaches from the literature.

Danksagung

Sowohl die Durchführung meines Promotionsvorhabens als auch die Fertigstellung der vorliegenden Dissertation wäre ohne die Unterstützung verschiedener Personen nicht möglich gewesen, denen ich nachfolgend meinen Dank aussprechen möchte:

Allen voran bedanke ich mich bei Herrn Prof. Dr. Oliver Vornberger für die Möglichkeit, mein Promotionsvorhaben durchzuführen sowie für jegliche damit verbundene Unterstützung.

Ebenso bedanke ich mich bei Herrn Prof. Dr. Karsten Morisse, dessen Rat und Anmerkungen maßgeblich zum Gelingen der Dissertation beigetragen haben. Ohnehin bedarf es einem großen Dank für das uneigennützig Engagement in jeglichem Belang vor und während der Promotionszeit.

Auch einige Kollegen möchte ich an dieser Stelle erwähnen, die mir während meines Promotionsvorhabens helfend zur Seite standen. Eduard Wolf danke ich für den regen fachlichen Austausch bei der Entwicklung des Systems. Beate Siegel, Julius Hoyer und Svenja Wichelhaus möchte ich für die Anregungen und die organisatorische Hilfe im Kontext der Evaluation danken. Ohnehin danke ich allen genannten Kollegen für viele motivierende Gespräche und für viel Verständnis, wenn mir die Promotion Zeit und Nerven abverlangte.

Auch Natalie Pfeifer danke ich für viele motivierende Worte und für das sorgsame Aufspüren von fehlerhafter Rechtschreibung und Grammatik. Ebenso danke ich meiner Mutter Maria Engelbert für viele gesetzte Kommata. Gewiss gebührt meinen Eltern Maria und Herbert Engelbert ein besonderer Dank. Nicht nur die Ermöglichung des Studiums, sondern auch das in mich gesetzte Vertrauen hat entscheidend dazu beigetragen, dass ich die vorliegende Dissertation überhaupt anfertigen konnte.

Abschließend sage ich Kim Verena Grabbe danke, die mir in einer intensiven Zeit mit viel Verständnis entgegengetreten ist und die oftmals meine Belange selbstlos und selbstverständlich in den Vordergrund gestellt hat.

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Motivation und Ausgangslage.....	1
1.2	Problemstellung.....	3
1.3	Forschungsziele und Einordnung der Arbeit.....	5
1.4	Aufbau der Arbeit.....	8
2	Voruntersuchung	10
2.1	Ziele	12
2.2	Methode.....	12
2.3	Aufbau.....	15
2.4	Präsentation der Ergebnisse	16
2.4.1	Lerntagebuch.....	16
2.4.2	Fragebogen.....	19
2.5	Diskussion der Ergebnisse.....	21
2.6	Schlussfolgerung.....	24
3	Grundlagen und Stand der Forschung	26
3.1	Empfehlungsdienste Grundlagen: Terminologie und Ziele	27
3.2	Empfehlungsdienste Grundlagen: Verfahrensarten.....	34
3.2.1	Inhaltsbasierte Verfahren	38
3.2.2	Kollaborative Verfahren	42
3.2.3	Hybride Verfahrensarten.....	49
3.2.4	Erweiterte Analyseverfahren: Clustering.....	51
3.2.5	Evaluationsmethoden	54
3.3	Social Tagging Grundlagen.....	63
3.3.1	Grundlagen Social Tagging	64
3.3.2	Der Begriff der Folksonomy	66
3.3.3	Social Tagging und Empfehlungsdienste: Verfahrensarten.....	67
3.3.4	Nachteile Social Tagging Systems im Kontext von Empfehlungsdiensten	73
3.4	Empfehlungsdienste im Kontext E-Learning: Zusammenhänge und Abgrenzung.....	74
3.5	Empfehlungsdienste im Kontext E-Learning: Überblick relevanter Forschungsansätze....	81
3.5.1	Lernressource	82

3.5.2	Learning Objects	89
3.5.3	Kontextinformationen	92
3.5.4	Empfehlungen für Lehrende	95
3.6	Zusammenfassung.....	98
4	Der Empfehlungsdienst LAOs	103
4.1	Ziele und Anforderungen	104
4.2	Konzeption	107
4.3	Systemimplementation.....	109
4.4	Empfehlungsdienst Algorithmik.....	119
4.4.1	Systemdefinition	120
4.4.2	Datenaufbereitung.....	122
4.4.3	Datenanalyse	126
4.4.4	Empfehlung.....	127
5	Evaluation	138
5.1	Problemstellung und Zielvorgabe.....	138
5.2	Hypothesen	141
5.3	Aufbau der Studie	143
5.3.1	Methode	144
5.3.2	Benutzerstudie unter Laborbedingungen	145
5.3.3	Online Simulation in realem Kursumfeld	148
6	Präsentation der Ergebnisse	157
6.1	Benutzerstudie unter Laborbedingungen.....	157
6.2	Online Simulation in realem Kursumfeld.....	162
7	Interpretation der Evaluationsergebnisse	177
8	Diskussion	188
9	Fazit & Ausblick	205
10	Literaturverzeichnis	209
11	Anhang.....	221

Abbildungsverzeichnis

Abbildung 1 Beispielhafter Eintrag im Lerntagebuch	13
Abbildung 2 Eintrag erstellen im Lerntagebuch	14
Abbildung 3 Nutzungshäufigkeit Lernmaterialien.....	17
Abbildung 4 Bewertung der Lernmaterialien.....	18
Abbildung 5 Grund des Lernens in Abhängigkeit der Lerntagebucheinträge.....	19
Abbildung 6 Auswertung Frage 1 und 2	19
Abbildung 7 Dauer der Nutzung neuer Lernmaterialien (in Prozent).....	20
Abbildung 8 Anlaufstelle zur Suche neuer Lernmaterialien	21
Abbildung 9 Suche im Internet nach Lernmaterialtypen	21
Abbildung 10 Verweis auf YouTube Video im Skript.....	22
Abbildung 11 Aufbau Empfehlungsdienst.....	33
Abbildung 12 User-Item-Matrix	36
Abbildung 13 Schritte eines Empfehlungsdienst Verfahrens.....	37
Abbildung 14 k-Means Clustering Algorithmus.....	53
Abbildung 15 Agglomeratives (hierarchisches) Clustering.....	54
Abbildung 16 Technology Acceptance Model (TAM) nach Davis	62
Abbildung 17 TAM Merkmale für Empfehlungsdienste	63
Abbildung 18 Darstellung eines Social Tagging Systems	65
Abbildung 19 Folksonomy als 3-dimensionale Matrix.....	67
Abbildung 20 Tag-Cloud am Beispiel der häufigsten Tags im System BibSonomy.....	69
Abbildung 21 Projektion der ternären Relation Y auf zwei Teilmatrizen (entnommen aus [129] S.627)	71
Abbildung 22 Variablen in einem E-Learning Szenario (entnommen aus [101] S. 29)	80
Abbildung 23 System Implementation.....	111
Abbildung 24 Taggingkomponente Textdokument	112
Abbildung 25 Taggingkomponente Videodokument.....	113
Abbildung 26 Übersicht Bewertungskriterien des Evaluationsschemas LORI (entnommen aus [89]).....	115
Abbildung 27 Quicktag im System LAOs	116

Abbildung 28 Abhängigkeit Relation zwischen User, Tag und Resource.....	118
Abbildung 29 Multimediakoordinaten für Text und Video	119
Abbildung 30 Systemdefinition	121
Abbildung 31 Ablauf Clustering Verfahren.....	123
Abbildung 32 Empfehlung in LAOs	128
Abbildung 33 Struktureller Aufbau einer Empfehlung.....	129
Abbildung 34 Herleitung textuelle Aussagen	132
Abbildung 35 Herleitung Handlungsempfehlung	133
Abbildung 36 Übersicht Lehrender Analysetools.....	135
Abbildung 37 Lehrender Detaillierte Materialanalyse.....	136
Abbildung 38 Lehrender Detaillierte Metarialeinsicht	137
Abbildung 39 Lehrender Detaillierte Nutzeranalyse.....	137
Abbildung 40 Empfehlungsdienstprozess LAOs	140
Abbildung 41 Zusammenhang Datenbasis und Qualität.....	141
Abbildung 42 Evaluationsablauf online Simulation	151
Abbildung 43 Hinweis auf Empfehlungen im System LAOs	152
Abbildung 44 Bewertungsschema für Empfehlungen.....	153
Abbildung 45 Klassifikationsabgleich	155
Abbildung 46 Klassifikation (TP, FP) in Abhängigkeit Anzahl Tags Benutzerstudie	161
Abbildung 47 Klassifikation (TN, FN) in Abhängigkeit Anzahl Tags Benutzerstudie	161
Abbildung 48 Klassifikation (TP, FP) in Abhängigkeit Anzahl Tags online Simulation.....	165
Abbildung 49 Klassifikation (TN, FN) in Abhängigkeit Anzahl Tags online Simulation.....	166
Abbildung 50 Auswertung Bewertungen Aggregiert.....	167
Abbildung 51 Aggregierte Auswertung der Bewertungen in Abhängigkeit der Klassifikation (online Simulation).....	171
Abbildung 52 Bewertungen in Abhängigkeit Anzahl an Tags online Simulation.....	172
Abbildung 53 Bewertungen je Resource in Abhängigkeit Anzahl an Tags online Simulation ..	172
Abbildung 54 Bewertung in Abhängigkeit Anzahl alternativer Lernmaterialien online Simulation	173
Abbildung 55 Auswertung Akzeptanzbefragung Aggregiert.....	174
Abbildung 56 Auswertung Akzeptanzbefragung (Kategorie Qualität).....	175
Abbildung 57 Auswertung Akzeptanzbefragung (Kategorie Aufwand).....	175
Abbildung 58 Auswertung Akzeptanzbefragung (Kategorie Akzeptanz allgemein).....	176

Abbildung 59 Beispiel False Positive Klassifikation.....	194
Abbildung 60 Klassifikation auf Grundlage der Verweildauer (Skript, online Simulation).....	196
Abbildung 61 Klassifikation auf Grundlage Verweildauer, Tagaufkommen (Skript, online Simulation).....	197
Abbildung 62 Auswertung Testaufgaben Benutzerstudie.....	199

Tabellenverzeichnis

Tabelle 1 Beispiel Bewertungstabelle für Filme	35
Tabelle 2 Klassifikationstabelle	59
Tabelle 3 Vor-/Nachteile aktueller Verfahren.....	99
Tabelle 4 Übersicht Tags im System.....	114
Tabelle 5 Lernmaterialien 1. Evaluationsschritt.....	147
Tabelle 6 Lernmaterialien 2. Evaluationsschritt.....	150
Tabelle 7 Fragebogen Akzeptanz gemäß TAM	154
Tabelle 8 Auswertung Gesamtübersicht Benutzerstudie.....	158
Tabelle 9 Klassifikationstabelle Benutzerstudie	159
Tabelle 10 Klassifikationstabelle nach Material Benutzerstudie	160
Tabelle 11 Auswertung Gesamtübersicht Online Simulation	163
Tabelle 12 Klassifikationstabelle Online Simulation.....	163
Tabelle 13 Klassifikationstabelle nach Material Online Studie	164
Tabelle 14 Bewertung Gesamtübersicht Online Simulation	166
Tabelle 15 Auswertung Bewertungen in Abhängigkeit True-Positives online Simulation	168
Tabelle 16 Auswertung Bewertungen in Abhängigkeit True-Negatives online Simulation.....	168
Tabelle 17 Auswertung Bewertungen in Abhängigkeit False-Positives online Simulation	169
Tabelle 18 Auswertung Bewertungen in Abhängigkeit False-Negatives online Simulation.....	169
Tabelle 19 Auswertung Bewertungen in Abhängigkeit keiner Klassifikation online Simulation	170
Tabelle 20 Klassifikation auf Grundlage Verweildauer, Tagaufkommen (Auszug).....	198

1 Einleitung

1.1 Motivation und Ausgangslage

Das Lehren und Lernen an deutschen Hochschulen und Universitäten hat sich über die Jahre hinweg stark verändert. Der Einsatz digitaler Präsentationsformen, die von Powerpoint über multimediale Inhalte bis hin zu Smartboards reichen ist in vielen Hörsälen heutzutage gängige Praxis. Auch die Form und Vielfalt von Lernmaterialien hat sich im Zuge der Digitalisierung stark verändert. Wo früher kopierte Skripte und Hörsaalmitschriften die Grundlage zum Lernen darstellten und die Bibliothek gegebenenfalls die Möglichkeit zur Erweiterung des Repertoire an Lernmaterialien bot, scheint die Vielfalt und Menge an Lernmaterialien heutzutage kein Ende zu nehmen. Digitale Vorlesungsskripte und Präsentationsfolien avancieren zum quasi Standard in vielen Bereichen der Hochschullandschaft und werden immer häufiger durch multimediales Material¹ wie z.B. Podcasts, Vorlesungsaufzeichnungen oder YouTube Videos ergänzt. Die Digitalisierung wird zusätzlich durch die fortwährende Zunahme von online Kurse z.B. in Form von *Massive Open Online Courses (MOOCs)* oder frei Verfügbaren Bildungsressourcen z.B. in Form von *Open Educational Resources (OER)* vorangetrieben. Diese Trendwende kann prinzipiell positiv gesehen werden, da vor allem Studierende oder generell Lernende von etwaigen Angeboten profitieren. So konnte z.B. der Mehrwert von Vorlesungsaufzeichnungen für Studierende bereits mehrfach bestätigt werden [132, 159]. Die Einbettung inhaltlich aufbereiteter Lernmaterialien in einem didaktischen Lernszenario verspricht ebenso einen weiteren Mehrwert für Studierende [168]. Dem Internet ist es schlussendlich anzulasten, dass sich die Anzahl an zugänglichen Lernmaterialien vervielfacht, da sich die Verbreitung von Daten

¹ Die Begriffe Lernmaterial und Material werden im Fortlaufenden synonym verwendet

und Informationen als komfortabel erweist und viele Lehrende auf die Verbreitung über das Internet setzen. Das daraus entstehende Überangebot erzielt dabei nicht ausschließlich positive Effekte. Zwar besteht durch die Nutzung des Internets die Möglichkeit benötigte Informationen schnell und unkompliziert abzufragen, allerdings kann angenommen, dass sich beim Zugriff auf unselektierte und nicht bewertete komplexe Informationen schwer zu überwindende Hürden ergeben [28]. Für Informationen aus dem universitären Lehrkontext bedarf es so z.B. ein gewisses Maß an Zeit, um Sachverhalte verstehen, bewerten und mit bereits bestehenden Informationen verbinden zu können. Ob Studierende diese Zeit für die Erarbeitung tatsächlich aufbringen bleibt an dieser Stelle zunächst offen.

Dass eine Vielzahl an zusätzlichen Lernmaterialien im Netz zur Verfügung stehen kann mit einem einfachen Experiment bestätigt werden. Nutzt man die Suchmaschine Google mit dem Begriff „Algorithmen und Datenstrukturen“, welches als grundständiges Themenfeld für jeden Informatik Studiengang gesehen werden kann, findet man bereits unter den ersten zehn Treffern² sieben unterschiedliche Internetseiten mit Skripten, Folien und Übungen von deutschen Hochschulen die frei zugänglich sind. Das Ergebnis spiegelt sich in gleicher Form auch auf den übrigen Ergebnisseiten der Suchmaschine wider. Es sei an dieser Stelle deutlich vermerkt, dass dies keine Aussage über die Qualität der gefundenen Lernmaterialien darstellt. Vielmehr verbildlicht das Experiment das bestehende Überangebot von Lernmaterialien im Internet, welches Studierende bei der Suche nach alternativen Informationen überwinden müssen. Als mögliches Resultat aus der Bereitstellung eines etwaigen Überangebots können in diesem Zusammenhang Desorganisation und mentale Überforderung genannt werden [3].

Dass Studierende zusätzliches Material aus dem Internet nutzen, sei an dieser Stelle vorweggenommen. Die im Kontext dieser Arbeit durchgeführte Voruntersuchung (vgl. [45] sowie Kapitel 2) konnte diese These belegen. Auch die Studien in [140] und [67] bestätigen die Hinzunahme des Internets um weitere Informationen zu erhalten. Dies geschieht auch obgleich Lernende das vom Lehrenden zur Verfügung gestellte Lernmaterial als primär zu nutzendes Material sehen. Die Qualität der bestehenden Lernmaterialien spielt dabei eine untergeordnete

² Zum Zeitpunkt der Verfassung der vorliegenden Arbeit

Rolle und Studierende suchen ungeachtet der Güte nach ergänzenden Übungen, Erklärungen oder Beispielen. Fraglich ist an dieser Stelle ob die Suche einen Mehrwert für die Studierenden bringt, da sich gezeigt hat, dass den zusätzlichen Materialien nur eine geringe Bedeutung zugemessen wird.

An dieser Stelle wird der Bedarf gesehen, bestehende Lernmaterialien mit sachdienlichen und kontextbezogenen Zusatzinformationen anzureichern, um ein verbessertes Verständnis zu erwirken und die Nutzung von bestehenden Materialien anzuregen. Des Weiteren bedarf es einer inhaltlichen Verknüpfung zu weiteren Materialien, um Lernenden den Zugang zu neuen Lernmaterialien zu erleichtern. Sowohl die Anreicherung von Lernmaterialien durch Zusatzinformation z.B. in Form von Hilfestellungen, als auch die Verknüpfungen zu weiteren Materialien erspart den Lernenden Zeit bei der Erarbeitung und Suche von Inhalten und kann zudem helfen, Lernende auf relevante und sachdienliche Lernmaterialien oder -Inhalte aufmerksam zu machen. Empfehlungsdienste können an dieser Stelle anknüpfen, um bei der Suche und Aufbereitung geeigneter Informationen zu helfen. Ganz allgemein stellt ein Empfehlungsdienst ein Softwaresystem dar, welches in Abhängigkeit von Benutzerinformationen neue hilfreiche, nützliche oder interessante Informationen – sogenannte Empfehlungen, herleitet. Empfehlungsdienste finden häufig in z.B. Webshops wie Amazon [94] Anwendung, wo auf Grundlage von bewerteten oder gekauften Objekten (Bücher, DVDs, etc.) Empfehlungen in Form von weiteren Objekten abgeleitet werden.

1.2 Problemstellung

Zwei primäre Fragestellungen ergeben sich bei der Entwicklung eines Empfehlungsdienstes, bevor über eine technische/algorithmische Umsetzung nachgedacht werden kann. Zum einen bedarf es der Klärung, welche Datenbasis dem System zu Grunde liegt bzw. auf welcher Datenbasis ein Empfehlungsprozess angestoßen werden kann. In dem in der Motivation angeführten Beispiel eines Webshops werden so z.B. bestehende Artikel und deren Attribute oder Metainformationen – wie z.B. Autor, Genre, etc. genutzt, um ähnliche Artikel ausfindig zu machen. Des Weiteren muss betrachtet werden, welche Zielsetzung ein Empfehlungsdienst verfolgen soll. Empfehlungen welcher Art oder mit welchem Inhalt sollen dem Endanwender präsentiert werden? Im Beispiel des Webshops sind dies alternative Artikel, die zum Kauf

angeboten werden, um z.B. das Kaufverhalten des Käufers weiter anzuregen. Hierzu bedarf es der Information über den Benutzer, welche Käufe bereits durchgeführt wurden.

Betrachtet man das Anwendungsszenario der vorliegenden Arbeit fällt die Beantwortung beider Fragestellungen deutlich schwieriger aus. Die Ziele eines etwaigen Systems wurden in der Motivation bereits kurz umrissen. Es gilt die Vorgabe inhaltlich gebundene Informationen in Lernmaterialien zur Verfügung zu stellen, um das Lernen mit einem Material zu unterstützen und Lernende für bestehende Materialien zu interessieren. Des Weiteren sollen Verknüpfungen sowohl zwischen bestehenden als auch zu neuen Lernmaterialien existieren, um die Suche nach alternativen Übungen, Erklärungen oder Beispielen zu vereinfachen. Die Bereitstellung inhaltlicher Verknüpfungen zwischen Lernmaterialien stellt eine recht klare Zielvorgabe dar. Noch unbestimmt ist allerdings, welche Informationen Lernenden helfen könnten, ein besseres Verständnis oder größeres Interesse für bestehende Materialien zu entwickeln. Verfahren, die einer inhaltlichen Aufbereitung von Lernmaterialien gleichkommt, existieren in der Literatur bislang nur bedingt. Es besteht zwar eine Vielzahl an Ansätzen, die sich generell mit der Empfehlung von Lernressourcen beschäftigen, die sich allerdings häufig auf gesamte Objekte, wie z.B. Skripte, beziehen und inhaltliche Bereiche unberücksichtigt lassen (vgl. in diesem Zusammenhang Abschnitt 3.5.1)³. Auch aus diesem Grund gibt die Literatur bislang noch keinen Aufschluss darüber, welche Informationen in einem Lernmaterial bereitgestellt werden müssen, um die Lernende oder den Lernenden für die Arbeit mit dem Material zu motivieren. Es besteht an dieser Stelle der Bedarf ein entsprechendes Modell für den Anwendungskontext zu entwickeln.

Bei der Betrachtung der zur Verfügung stehenden Datenbasis ergeben sich weitere zu untersuchende Probleme. Eine Datenbasis wie im eingangs erwähnten Beispiel des Webshops in Form von Attributen oder Benutzerbewertungen liegt im hier vorgestellten Anwendungskontext nicht vor. Lernmaterialien verfügen zwar zumeist über einige Angaben wie z.B. Autor, Titel oder Thema, die innerhalb des Materials zu finden sind. Allerdings liegen diese als Metainformationen in der Regel nicht vor. Es kann an dieser Stelle schon vorweg genommen

³ Kapitel 3 wird noch eine umfassende Darstellung gängiger Verfahren mit ihren Vor- und Nachteilen präsentieren.

werden, dass z.B. Autor- oder Fachinformationen für die inhaltliche Bewertung von Lernmaterialien ohnehin unzureichend sind. Das Problem liegt in erster Linie im Umfang und der Komplexität von Lernmaterialien. Vorlesungsskripte im universitären Lernkontext umfassen zumeist einige hundert Seiten und können nicht anhand einiger weniger übergeordneter Informationen miteinander verglichen werden. Bei Lernmaterialien mit geringem Umfang oder z.B. Webseiten kann eine übergeordnete Menge an Informationen reichen (vgl. z.B. [106]). Deswegen ist eine gängige Verfahrensart Lernmaterialien in möglichst kleine und abgeschlossene Inhaltsobjekte aufzuteilen – sogenannte *Learning Objects*. So ist eine feingranulare Aufteilung von Lernmaterialien möglich, in der jedes Learning Object mit individuellen Metainformationen angereichert werden kann. Der Aufwand geht zu Lasten des Lehrenden, der diese Zeit in der Regel nicht investieren möchte (vgl. z.B. [112]). Weitere Verfahrensarten wie z.B. *Lernpfade* sind in der Literatur präsent und werden in Kapitel 3 ausführlicher diskutiert, bergen allerdings ähnliche Nachteile wie die bereits genannten Learning Objects (vgl. [101] S. 8, [117]). Generell ist ein Verfahren wünschenswert, indem ohne zusätzlichen Aufwand die Nutzung jeglichem bestehenden Lernmaterials möglich ist. Des Weiteren sollte jeder Typ von Lernmaterial – also unabhängig ob es ein Text- oder Videodokument ist, Verwendung finden können. Beide Anforderungen grenzen die Möglichkeiten einer einheitlich verwendbaren Datenbasis weiter ein. Gängige Ansätze aus dem Bereich der Empfehlungsdienste – auch im Kontext E-Learning, bergen verschiedene Nachteile die eine Nutzung im vorgestellten Anwendungskontext nicht erlauben.

1.3 Forschungsziele und Einordnung der Arbeit

Durch die zuvor skizzierte Problemstellung ergeben sich recht eindeutige Forschungsziele die im Folgenden benannt werden. Die vorliegende Arbeit untersucht die Entwicklung eines algorithmischen Ansatzes sowohl zur Generierung von inhaltsbezogenen⁴ Hilfestellungen in Lernmaterialien als auch der inhaltsbezogenen Verknüpfung zwischen Lernmaterialien, die

⁴ Inhaltlich oder inhaltsbezogen meint in der vorliegenden Arbeit zunächst den Verweis auf einen Inhalt bzw. Teilbereich innerhalb eines Lernmaterials. Kapitel 4 legt den Sachverhalt dann gesondert dar.

Lernenden als Empfehlung⁵ zur Verfügung gestellt werden kann. Der Arbeit liegt somit nachfolgende Forschungsfrage zugrunde:

Kann ein Empfehlungsdienst entwickelt werden, der maschinell inhaltliche Hilfestellungen für bestehende Lernmaterialien, als auch inhaltliche Verknüpfungen zwischen bestehenden und neuen, ungenutzten Lernmaterialien zur Unterstützung von Lernenden in Form von Empfehlungen generiert?

Vier weitere Teilfragen ergeben sich für die Forschungsfrage, die es insbesondere zu beantworten gilt:

- Was umfasst eine unterstützende inhaltliche Hilfestellung innerhalb des Empfehlungsdienstes?
- Was umfasst eine unterstützende inhaltliche Verbindung zwischen Lernmaterialien innerhalb des Empfehlungsdienstes?
- Welche existierenden Verfahrensarten aus dem Forschungsfeld der Empfehlungsdienste können genutzt werden, um inhaltliche Hilfestellungen und Verbindungen zwischen Lernmaterialien in Form von Empfehlungen herleiten zu können? Ist als Alternative die Neuentwicklung eines Verfahrens notwendig?
- Welche Güte⁶ haben die durch das System zur Verfügung gestellten Empfehlungen?
- Wie hängt die Güte für Empfehlungen mit der Menge an Daten für eine entsprechende Verfahrensart zusammen?

Des Weiteren ergeben sich aus der Argumentation in Abschnitt 1.2 verschiedene Aspekte die es als eine Art Zielvorgabe⁷ mit zu berücksichtigen gilt:

⁵ Sowohl inhaltliche Hilfestellungen, als auch inhaltsbezogene Verknüpfungen werden nachfolgend unter dem Begriff Empfehlung benannt.

⁶ Den Begriff der Güte gilt es im Verlauf der Arbeit noch genauer zu definieren, was im Kontext der Evaluation in Kapitel 5 geschehen wird.

⁷ Sowohl die Forschungsfrage als auch die zusätzlichen Zielvorgaben werden im weiteren Verlauf der Arbeit auch allgemein als Ziele bezeichnet.

- Bestehende Lernmaterialien von Lehrenden müssen im Empfehlungsdienst nutzbar sein
- Eine gesonderte Aufbereitung von Lernmaterialien durch z.B. den Lehrenden ist nicht erforderlich
- Jegliche Form von Lernmaterialien sollen Berücksichtigung finden (z.B. Text, Audio, Video)

Wie sich noch zeigen wird umfasst die Entwicklung eines Empfehlungsdienstes sowohl die Entwicklung eines Verfahrens zur Datenanalyse/-bewertung als auch ein Verfahren zur geeigneten Datenerhebung/-repräsentation. Beide Teilbereiche werden bei der Entwicklung berücksichtigt, so dass ein ganzheitlicher Empfehlungsdienstansatz entsteht. Bedeutsam wird zunächst allerdings die Frage sein, was eine Empfehlung innerhalb des Ansatzes umfasst und welche Informationen präsentiert werden müssen, um einen Mehrwert für Lernende zu erwirken. Auch worauf sich die Terminologie inhaltlich bezieht muss zunächst eindeutig definiert werden (vgl. Kapitel 4). Zunächst sei der Begriff inhaltlich als ein Verweis auf einen Teilbereich innerhalb eines Lernmaterials interpretiert.

Thematisch reiht sich die Arbeit im Forschungsfeld der *Empfehlungsdienste* oder auch *Recommender Systems*⁸ ein. Durch den engen Bezug zum Thema E-Learning spricht man auch vom Forschungsfeld *Empfehlungsdienste im Kontext E-Learning*, wobei die englische Bezeichnung *Recommender Systems for Learning* auch in der deutschsprachigen Forschungscommunity eine gängigere Bezeichnung darstellt. Des Weiteren können die Forschungsfelder (*Educational*)- *Data Mining* und *Learning Analytics* der Arbeit zugerechnet werden. Auch einige Lerntheoretische Probleme werden im Kontext der Arbeit diskutiert, allerdings liegt der Fokus klar auf den technischen Fragestellungen. Wie der vorherige Abschnitt 1.2 angedeutet hat, weisen die genannten Forschungsfelder bereits eine Vielzahl an angrenzenden Verfahren auf, die allerdings im Kontext der vorliegenden Forschungsfrage zumeist entscheidende Nachteile aufweisen. Dies ist zwar häufig einer unterschiedlichen Ausgangslage oder Zielvorgabe geschuldet, allerdings sind Verfahren die eine inhaltliche

⁸ **Anmerkung:** Im Folgenden der Arbeit wird eine englische Begrifflichkeit ohne weitere Erläuterung verwendet, sofern es sich um einen eigenständigen Fachbegriff/-term handelt. Diese sind bei erstmaliger Verwendung kursiv gekennzeichnet. Auf eine explizite Übersetzung in die deutsche Sprache wird bewusst verzichtet.

Analyse, Vernetzung und Empfehlung von Lernmaterialien vorsehen, nur wenig etabliert. Eine Kernherausforderung wird es deshalb sein, einen geeigneten Ansatzpunkt zu finden, um die genannten Zielvorgaben zu erfüllen. Es wird sich im weiteren Verlauf zeigen, dass die Auswertung von inhaltlichen Bewertungen durch Lernende und somit die Analyse von kollektivem Wissen die vielversprechendste Lösung darstellt. In diesem Zusammenhang wird das Forschungsgebiet des *Social Tagging* bzw. der *Social Tagging Systems* untersucht und die Nutzung im Bereich der Empfehlungsdienste geprüft. An dieser Stelle wird die Herausforderung gesehen unterschiedliche Forschungsansätze zu kombinieren, um der genannten Zielvorgabe genügen zu können.

Die Bewertung des zu entwickelnden Verfahrens ist ein weiteres primäres Ziel der Arbeit. Hierzu bedarf es der Entwicklung eines geeigneten Evaluationsdesigns sowie der entsprechenden Durchführung. Die Bewertung der Empfehlungen ist dabei ebenso Bestandteil wie die Prüfung der Akzeptanz. Um dies zu ermöglichen, ist eine prototypische Umsetzung des Systems vorgesehen.

1.4 Aufbau der Arbeit

Voruntersuchung

Das Kapitel präsentiert eine im Rahmen der Arbeit durchgeführte Vorstudie um das Nutzungsverhalten von Studierenden im Kontext einer Lehrveranstaltung zu untersuchen und um nützliche Informationen für die Entwicklung eines Empfehlungsdienstes ableiten zu können. Es wird das Studiendesign, die erhobenen Ergebnisse, sowie eine kurze Diskussion präsentiert.

Grundlagen und Stand der Technik

Innerhalb des Kapitels 3 werden theoretische Grundlagen für die genannten Forschungsbereiche gebildet und bestehende Verfahren die im Zusammenhang mit der vorliegenden Arbeit stehen, präsentiert. Es wird eine klare Abgrenzung zu bestehenden Verfahren vorgenommen und dargelegt, wo die vorliegende Arbeit ansetzt, sowie welche Forschungslücke geschlossen wird.

Der Empfehlungsdienst LAOs

Kapitel 4 stellt die Entwicklung des Empfehlungsdienstes LAOs vor. Das Konzept des Empfehlungsdienstes wird zunächst verbal erläutert, bevor eine formale Definition erfolgt. Final wird die Implementierung in ein prototypisches System beschrieben.

Evaluation

Kapitel 5 beschreibt eine Evaluation, die für den entwickelten Ansatz durchgeführt wird. Neben der Güte der generierten Empfehlungen wird auch die Akzeptanz und Nutzerfreundlichkeit des Systems geprüft. Es werden zu prüfende Hypothesen gestellt und ein entsprechendes Evaluationsdesign zur Validierung der Hypothesen präsentiert.

Präsentation und Deutung der Evaluationsergebnisse

Die Ergebnisse der Evaluation werden in Kapitel 6 beschrieben. Das Kapitel beschränkt sich dabei vornehmlich auf die Präsentation und Deutung der Ergebnisse

Diskussion

Die in Kapitel 6 präsentierten Evaluationsergebnisse werden in Kapitel 7 diskutiert und bewertet. Das Kapitel trifft final eine Aussage, mit welcher Güte der Empfehlungsdienst arbeitet und welcher Mehrwert für Lernende geschaffen wird. Des Weiteren bespricht das Kapitel mögliche Nachteile, die innerhalb der Evaluation ausgemacht wurden.

Fazit und Ausblick

Die Arbeit schließt mit einer kurzen Zusammenfassung über die Ergebnisse der Arbeit und versucht einen Ausblick. Zum einen soll versucht werden mögliche Schritte für eine Weiterentwicklung zu nennen. Des Weiteren soll beantwortet werden, ob und in welcher Form ein Einsatz in der produktiven Lehre möglich erscheint.

2 Voruntersuchung

Über die Nutzungsgewohnheiten von Lernenden⁹, wie und in welchem Umfang sie bestehende oder neue Lernmaterialien nutzen ist in der Literatur nur wenig bekannt. Dass sich Lernende innerhalb des Lernprozesses dem Internet bedienen um zusätzliche Informationen zu sammeln scheint hingegen sicher vergleicht man hierzu die bereits in der Motivation zitierten Studien von Schulmeister [140] mit 2096 Befragten sowie Höver [67] mit 107. Beide Studien geben Auskunft darüber, welche Dienste im Internet von Studierenden genutzt werden. Während die Studie von Schulmeister eine generelle Nutzung von Diensten betrachtet, beziehen Höver et. al. die Nutzung gezielt auf den Lernprozess, was für die vorliegende Arbeit trotz geringerer Anzahl an Befragten einen höheren Informationsgehalt darstellt. Beide Studien sind sich allerdings einig, dass vor allem Online-Enzyklopädien und Suchmaschinen einen hohen Stellenwert im Alltag Studierender darstellen. Höver et. al. beschreiben so z.B. dass 95% der Befragten Dienste wie Wikipedia¹⁰ nutzen, sowie 93% der Befragten Suchmaschinen – wie Yahoo¹¹ oder Google¹². Zwar sagten lediglich ~25% der Befragten aus Video-Diensten wie YouTube oder iTunesU zum Lernen zu nutzen. Allerdings hat sich seit 2010 – Stand der Befragung – die allgemeine Nutzung von Videoportalen in Deutschland bis 2013 verdoppelt und die Anzahl an hochgeladenen Videos bei YouTube zwischen 2010 und 2015 um das 16-fache erhöht, so dass im Jahr 2016 eine deutlich höhere Nutzung erwartet werden kann¹³. Ein Grund für die erhöhte Nutzung von

⁹ Im Nachfolgenden ist ein Lernender im universitären Kontext gemeint und wird mit dem Begriff Studierender synonym verwendet

¹⁰ Wikipedia: <http://www.wikipedia.org>

¹¹ Yahoo: <http://www.yahoo.com>

¹² Google: <http://www.google.com>

¹³ Abgerufen am 17. August 2016 von <http://www.tubefilter.com/2015/07/26/youtube-400-hours-content-every-minute/> und <https://www.youtube.com/yt/press/statistics.html>

Videoportalen durch Lernende ist, dass viele Lehrende Plattformen wie z.B. YouTube für Lehrzwecke für sich entdeckt haben [151]. Ohnehin ist das Angebot an weiteren Videoplattformen für die explizite Bereitstellung von Lernmaterialien in den letzten Jahren stark angestiegen. Dies ist z.B. an der stetig ansteigenden Anzahl an Plattformen für freie Bildungsressourcen oder auch *Open Education Resources (OER)* und für die Bereitstellung von *Massive Open Online Courses (MOOC)* zu sehen [116, 139]. Es kann somit angenommen werden, dass Studierende heutzutage mehr Anlaufstellen im Internet zum Lernen heranziehen.

Dennoch bleibt unklar wie Lernende oder Studierenden tatsächlich mit den gefundenen Informationen oder Lernmaterialien umgehen. Nutzen sie eher die Lernmaterialien des Lehrenden oder sind die Materialien aus dem Internet von größerer Bedeutung? Des Weiteren bleibt die Frage offen wie viel Zeit mit der Suche und schlussendlich mit der Nutzung von alternativen Lernmaterialien verbracht wird. Auch welcher Nutzen oder Mehrwert aus den Informationen aus dem Internet gezogen wird ist als Frage bislang offen. Eben diese Informationen können allerdings Aufschluss darüber geben, ob und welche alternativen Informationen und Materialien von Interesse für Studierende sind.

Das nachfolgende Kapitel stellt eine Untersuchung in Form eines Online-Lerntagebuchs vor, um die bislang offenen Fragen im Kontext der vorliegenden Arbeit zu beantworten. Hieraus ergibt sich auch der Name des Kapitels *Voruntersuchung*, da die Ergebnisse dieser Studie vor allem der Entwicklung eines geeigneten Modells für Empfehlungen dienen. Aus der Literatur ist bislang nämlich noch nicht ersichtlich, ob und aus welchen Informationen ein konkreter Mehrwert für Studierende entsteht. Des Weiteren soll die Annahme geprüft werden, ob Studierende sich weiterer Lernmaterialien – vor allem – von weiteren Hochschulen bedienen oder ob die Suche in einer Online-Enzyklopädie endet. Das Kapitel beschreibt zunächst die Ziele der Untersuchung und mit welchem Aufbau diese hinterfragt bzw. erwirkt werden können. In diesem Zusammenhang wird auch das Online-Lerntagebuch beschrieben, was für den Einsatz innerhalb der Untersuchung entwickelt wurde. Abschließend werden die erhobenen Ergebnisse präsentiert und im Kontext der vorliegenden Arbeit diskutiert. Die Ergebnisse der Voruntersuchung wurden in [45] veröffentlicht.

2.1 Ziele

Durch die Untersuchung soll sowohl das Nutzungsverhalten im Umgang mit digitalen Lernmaterialien einer Studierendengruppe als auch das Suchverhalten nach neuen, alternativen Materialien untersucht werden. Mit Nutzungsverhalten ist grundsätzlich die Häufigkeit, die ein Material zum Lernen herangezogen wurde, gemeint. Dabei geht man davon aus, dass die Nutzung eines Materials in der Regel einen Grund bzw. ein Ziel verfolgt. Auch dieser Zusammenhang soll genauer untersucht werden und findet in Kapitel 2.2 noch genauere Erläuterung. Nachfolgend seien die zu untersuchenden Ziele in Form von Leitfragen formuliert:

- Nutzen Studierende die vom Lehrenden zur Verfügung gestellten Lernmaterialien?
- Sehen Studierende einen Mehrwert in den bestehenden Lernmaterialien?
- Suchen Studierende im Internet nach inhaltlich zum Veranstaltungskontext passenden Lernmaterialien und nutzen sie sie?
- Sehen Studierende einen Mehrwert in den im Internet gefundenen Lernmaterialien?
- Warum suchen Studierende nach neuen Lernmaterialien?
- Welche Art von Lernmaterialien wird im Internet gesucht?

Die Untersuchung verlangt die Überprüfung des Nutzungsverhaltens von Lernmaterialien Studierender im alltäglichen Nutzungskontext und sieht deshalb eine explorative quantitative Datenanalyse vor (vgl. [15] S. 371 ff.). An dieser Stelle werden somit keine Hypothesen formuliert, die in quantitativer Form bestätigt oder abgelehnt werden können. Vielmehr dient die Vorstudie zur geeigneten Bildung von zu erforschenden Hypothesen. Es bedarf einer Methode, die das Lernverhalten von Studierenden über einen bestimmten Zeitraum hinweg dokumentiert, um Zusammenhänge zwischen Intention der Nutzung eines Lernmaterials, Zeitraum der Nutzung, sowie Mehrwert der Nutzung herstellen zu können.

2.2 Methode

Die Dokumentation des Lernverhaltens und des impliziten Nutzungsverhalten von Lernmaterialien wird durch das Führen eines Lerntagebuchs realisiert. Studierende erhalten die Aufgabe die Nutzung eines Lernmaterials innerhalb des Lerntagebuchs zu dokumentieren. Die

Nutzung eines Materials ist dabei mit verschiedenen Informationen verknüpft, die ebenfalls durch die Studierenden hinterlegt werden: Grund, Ziel und Bewertung des genutzten Materials. Der Grund der Nutzung eines Materials ist vordefiniert und kann entweder in einer *Vor-/Nachbereitung* oder in einer *Vorbereitung auf eine Übung* liegen. Eine Vorauswahl für Übungen kann vom Lehrenden hinterlegt werden. Das Ziel kann durch Studierende frei formuliert werden und beschreibt grundsätzlich den gewünschten Lernfortschritt. Die Bewertung eines Lernmaterials erfolgt in den vier Kategorien *Schwierigkeit*, *Unterstützung*, *Verständnis* und *Wichtigkeit* und wird jeweils auf einer Skala von 1 (wenig) bis 5 (sehr) durch die Studierenden bewertet. Ein beispielhafter Eintrag im Lerntagebuch ist zur Verdeutlichung in Abbildung 1 gezeigt. Das Lerntagebuch ist als onlinebasierte Applikation zur Verfügung gestellt und kann im Browser oder auch als mobile Web-App genutzt werden, um den Zugang für Studierende möglichst flexibel zu ermöglichen. Die Einträge des Lerntagebuchs werden als Lernpfad in chronologisch aufsteigender Reihenfolge dargestellt (vgl. ebenfalls Abbildung 1). Eine Liste mit bestehenden Materialien des Lehrenden ist im jeweiligen Veranstaltungskontext vordefiniert und kann durch die Studierenden mit eigenen Materialien – z.B. aus dem Internet – ergänzt werden (vergleiche hierzu auch Abbildung 2).

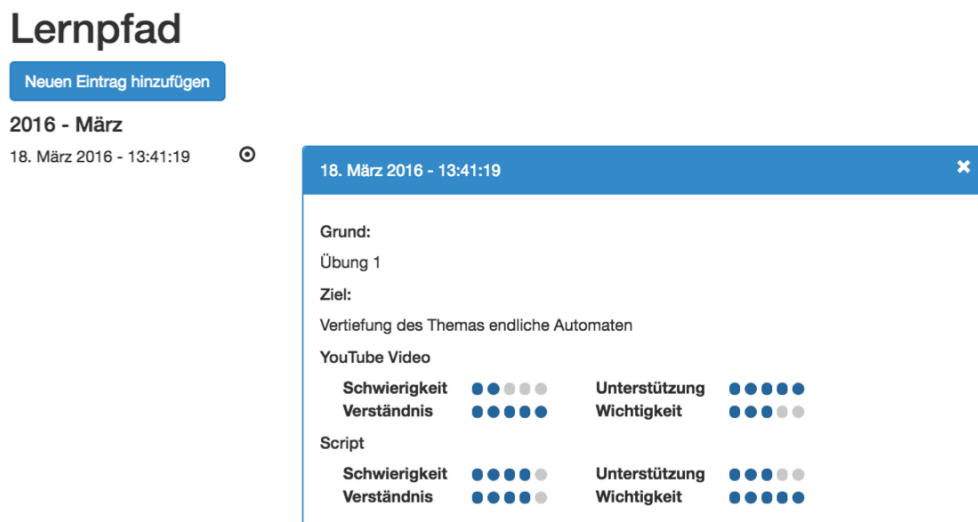


Abbildung 1 Beispielhafter Eintrag im Lerntagebuch

Aus welchem Grund hast Du gelernt?

- Vor- und Nachbereitung
- Übung 1
- Übung 2
- Übung 3
- Übung 4
- Übung 5

Wähle die Materialien, die Du genutzt hast und bewerte diese:

Script	Schwierigkeit	Verständnis	Untersützung	Wichtigkeit	
	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	-
Wegener Theoretische Informatik					+
Sipser Introduction of Computation					+
Erk Theoretische Informatik					+
Hoffmann Theoretische Informatik					+
Hopcroft Einführung in die Automatentheorie					+
Vorlesungsaufzeichnung					+
YouTube Video					+
TI Toolbox					+
Eigenes Material hinzufügen					➔

Abbildung 2 Eintrag erstellen im Lerntagebuch

Der Funktionsumfang der Lerntagebuchanwendung ist bewusst auf ein Minimum reduziert, um den Aufwand für Studierende möglichst gering zu halten. Es gibt somit lediglich die Möglichkeit dem Lernpfad einen neuen Eintrag hinzuzufügen (vgl. Abbildung 1 oben links), der über das Formular in Abbildung 2 final angelegt wird. Einträge können des Weiteren wieder gelöscht werden. Eine Login Funktionalität ermöglicht einen personalisierten Zugriff auf die Lerntagebücher. Die Auswertung der hinterlegten Daten erfolgt anonym.

Ein abschließender Fragebogen hinterfragt sowohl das Suchverhalten nach neuen Materialien als auch die Zufriedenheit über die zur Verfügung gestellten Materialien. Der Fragebogen untersucht weiterhin die durchschnittliche Nutzungsdauer, die Studierende mit dem jeweiligen Lernmaterial aufgewendet haben. Nachfolgende Fragen sind im Fragebogen enthalten:

1. Hast du dir alternative Materialien zum Lernen gesucht?
2. Wie lange hast du dich pro Woche mit den alternativen Materialien beschäftigt?
3. Hast du dich mit dem zur Verfügung gestellten Material in der Vorlesung theoretische Informatik genügend unterstützt gefühlt?
4. Was hat dir bei den zur Verfügung gestellten Lernmaterialien gefehlt?
5. Unabhängig von der Vorlesung theoretische Informatik: Wo suchst du nach alternativen Materialien?
6. Unabhängig von der Vorlesung theoretische Informatik: Nach welcher Art Material suchst du?

2.3 Aufbau

Die Durchführung der Untersuchung erfolgte im Rahmen der Lehrveranstaltung *Theoretische Informatik* im Wintersemester 2014 an der Hochschule Osnabrück. Die Veranstaltung stellt eine Pflichtveranstaltung für die Studiengänge *Technische Informatik* und *Medieninformatik* dar und erfolgt laut Studienverlaufsplan im 4. Semester. Das Curriculum sieht einen Anteil von 2/5 dozentengebundenem Workload sowie einen Anteil von 3/5 als individuelle Arbeitseinteilung vor. Studierende erhalten für die Veranstaltung 5 ECTS-Punkte. Nachfolgende Lernmaterialien wurden vorlesungsbegleitend im LMS der Veranstaltung zur Verfügung gestellt:

- Vorlesungsskript mit 300 Seiten
- Begleitende Vorlesungsaufzeichnungen mit Opencast Mattern mit einer regulären Länge von 90 Minuten
- Themenspezifische YouTube Video mit einer Länge zwischen 1:30 und 10 Minuten
- Fachbezogene Literaturhinweise

Die YouTube Videos sind thematische in Form eines *QR-Codes* mit dem Vorlesungsskript verknüpft und stellen eine digitale Hörsaalmitschrift¹⁴ in Form eines *Screencasts* dar, die durch einen nachträglich gesprochenen Audiokommentar ergänzt wurden. Die

¹⁴ Durchgeführt mit einem Tablett-Computer

Vorlesungsaufzeichnungen beinhalten einen vollständigen Vorlesungsmittschnitt und wurden spätestens einen Tag nach dem jeweiligen Veranstaltungstermin bereitgestellt.

Die Studierenden erhielten in der ersten Veranstaltung des Semesters eine Einführung in die durchzuführende Aufgabe und in das zu nutzende System. Hierbei wurden die Studierenden angehalten das Lerntagebuch über das gesamte Semester hinweg zu führen und neben der Nutzung bestehender Materialien, zusätzlich genutzte Lernmaterialien z.B. aus dem Internet mit in das System aufzunehmen. Der abschließende Fragebogen wurde in der letzten Vorlesungseinheit des Semesters durch die Studierenden ausgefüllt.

2.4 Präsentation der Ergebnisse

Die Ergebnisse der durchgeführten Untersuchung sind im nachfolgenden Abschnitt präsentiert. Zunächst erfolgt eine Auswertung der geführten Lerntagebücher in Abschnitt 2.4.1, bevor die Ergebnisse des Fragebogens in Abschnitt 2.4.2 präsentiert werden. Das Kapitel dient zunächst ausschließlich der Präsentation der Ergebnisse. Die Herleitung von Zusammenhängen innerhalb der Ergebnisse sowie eine entsprechende Diskussion erfolgt in Kapitel 2.5.

2.4.1 Lerntagebuch

Insgesamt 25 Studierende haben im Kontext der genannten Lehrveranstaltung ein Lerntagebuch geführt. Hierdurch konnten 244 Lerntagebucheinträge gesammelt werden, die im Nachfolgenden Aufschluss über das Nutzungsverhalten und über die Bewertung von Lernmaterialien gibt.

In Abbildung 3 ist zunächst die Nutzungshäufigkeit der Lernmaterialien in Abhängigkeit der Lerntagebucheinträge abgebildet.

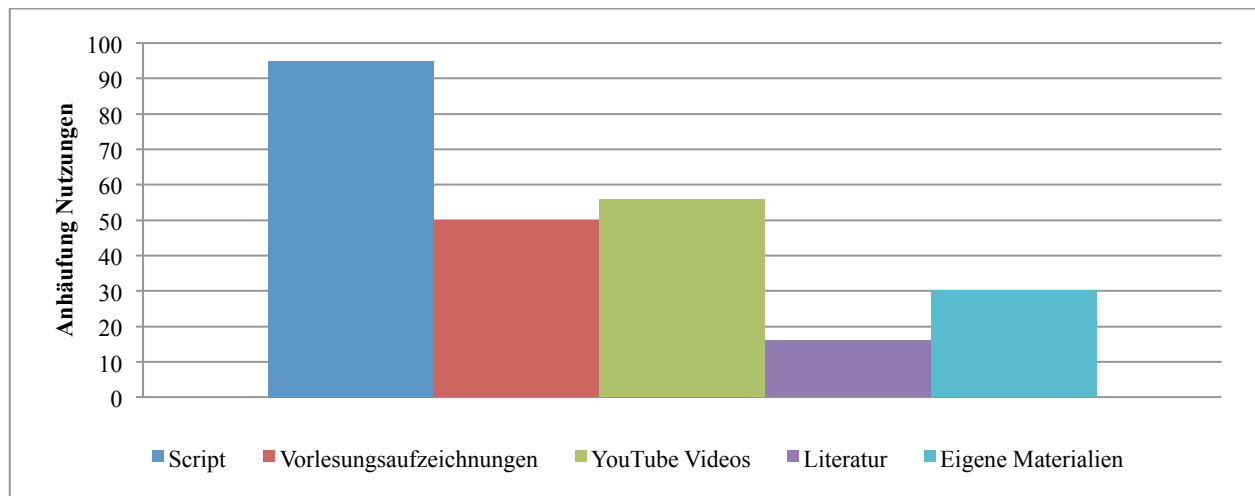


Abbildung 3 Nutzungshäufigkeit Lernmaterialien

Deutlich zu sehen ist die primäre Nutzung des Vorlesungsskriptes. Es kann eine fast doppelt so hohe Anzahl an Nutzungen festgestellt werden als bei YouTube Videos. Die multimedialen Elemente in Form von YouTube Videos und Vorlesungsaufzeichnungen wurden nahezu gleich genutzt. Die Nutzung von eigens gesuchten Materialien wird von den Studierenden vor der zur Verfügung gestellten Literatur bevorzugt.

In Abbildung 4 ist die Bewertung der Lernmaterialien in Abhängigkeit der Lerntagebucheinträge gezeigt. Die Abbildung zeigt die aggregierte Ausprägung in den Kategorien Verständnis, Wichtigkeit, Unterstützung und Schwierigkeit individuell für jedes Material (von 1 (wenig) bis 5 (sehr)):

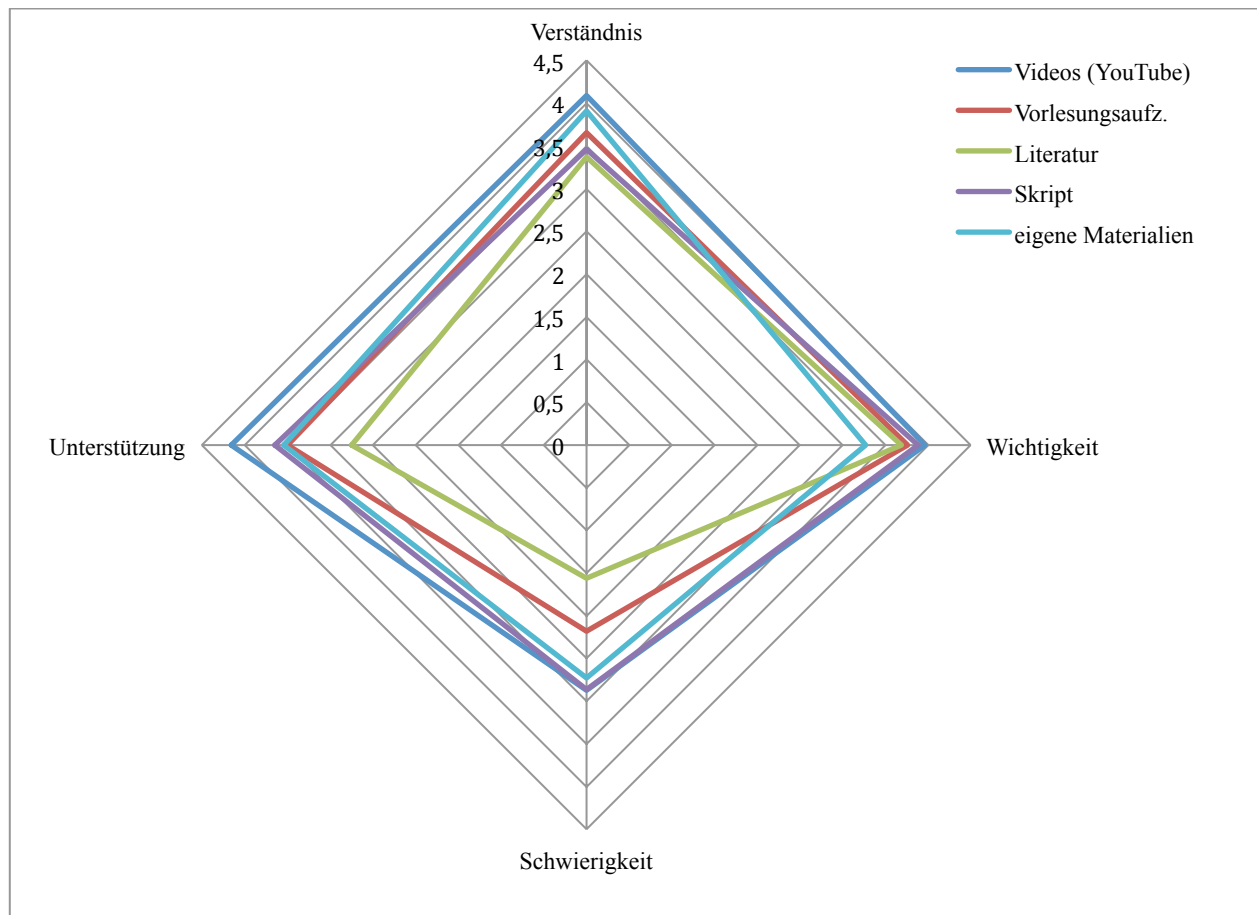


Abbildung 4 Bewertung der Lernmaterialien

Die YouTube Videos wurden durch die Studierenden mit einem hohen Grad an Unterstützung – mit recht deutlichem Abstand zu den übrigen Materialien - und einem hohen Verständnis bewertet. Abgesehen von der Literatur weisen alle weiteren Materialien einen ähnlichen Grad der Unterstützung auf. Die Materialien die durch die Studierenden selbst gesucht wurden erzielen ein ähnlich hohes Verständnis wie die bereitgestellten YouTube Videos. Die übrigen Materialien folgen mit einer geringeren Wertung in der Kategorie Verständnis. Als wichtig empfinden die Studierenden alle durch den Lehrenden zur Verfügung gestellten Materialien. Die eigens gesuchten Materialien fallen in diesem Punkt deutlich zurück. Als besonders schwierig wurden die YouTube Videos, das Vorlesungsskript und die eigenen Materialien bewertet.

In Abbildung 5 ist die Angabe von Gründen in Abhängigkeit der Lerntagebucheinträge dargestellt. Die Abbildung zeigt deutlich, dass die Studierenden als Angabe zur Nutzung vor

allen eine Vor-/und Nachbereitung angegeben haben. Nur selten erfolgte die Nutzung aufgrund einer Übung.

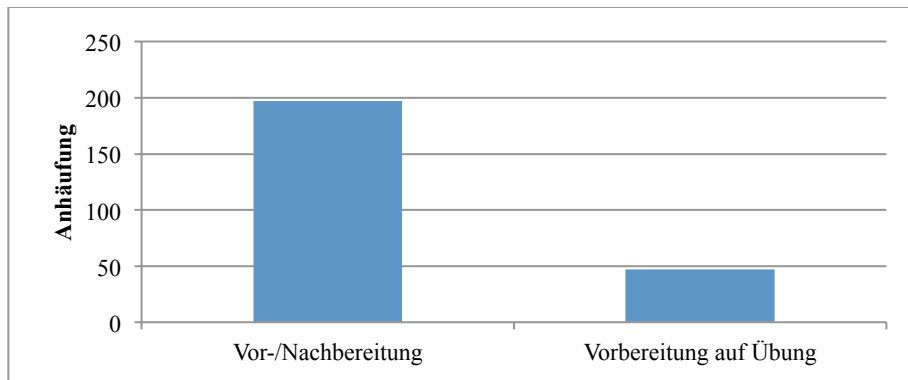


Abbildung 5 Grund des Lernens in Abhängigkeit der Lerntagebucheinträge

2.4.2 Fragebogen

Die Auswertung der abschließenden Fragebögen ergab die nachfolgenden Ergebnisse. Jeder Studierende, der an der Durchführung eines Lerntagebuches teilgenommen hat, wurde zum Ausfüllen des Fragebogens angehalten. Die Teilnehmerzahl bleibt gegenüber dem Lerntagebuch Szenario somit gleich. Die Ergebnisse der ersten beiden Fragen sind in Abbildung 6 zusammengefasst.

	Ja	Nein
Hast Du dir neue Materialien zum Lernen gesucht?	60%	40%
Hast Du dich mit den bestehenden Materialien genügend unterstützt gefühlt?	80%	20%

Abbildung 6 Auswertung Frage 1 und 2

Es ist zu sehen, dass der Großteil der Studierenden sich mit den zur Verfügung gestellten Materialien genügend unterstützt gefühlt hat. Dennoch gibt es eine große Schnittmenge von denen, die sich gut vorbereitet gefühlt haben und derer die sich dennoch neue Materialien gesucht haben. Auf die Frage nach ggf. fehlenden oder gewünschten Informationen in den gegebenen Materialien wurden nachfolgende Antworten gegeben: Zusätzliche Beispiele, ausführlichere Beispiele, neue Aufgaben mit Lösungswegen, ausführlichere Lösungswege, Beispiele mit zusätzlicher Erklärung, alternative Erläuterung, zusätzliche Erläuterung, verbesserte eigene Mitschrift.

In Abbildung 7 ist die Dauer der Nutzung von neu gesuchten Materialien zu sehen. Der Großteil der Studierenden gab an zwischen 30 und 60 Minuten pro Woche mit eigens gesuchten Materialien zu lernen. Nur ein geringer Teil gab an sich länger als eine Stunde mit neuen Materialien auseinanderzusetzen.

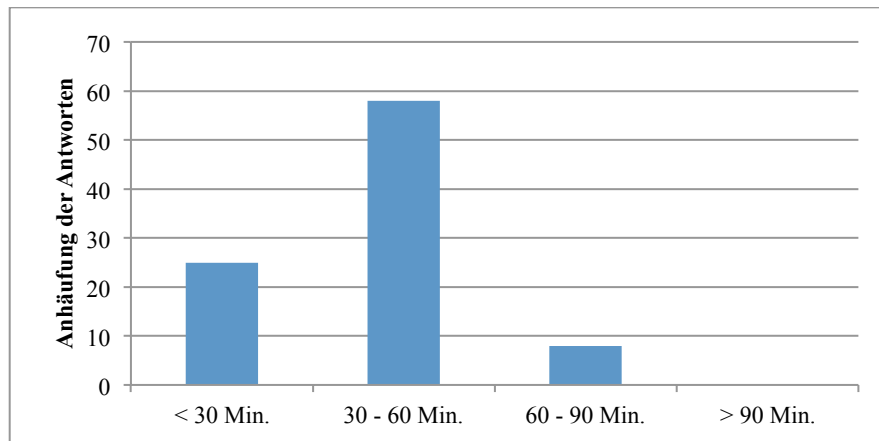


Abbildung 7 Dauer der Nutzung neuer Lernmaterialien (in Prozent)

Abbildung 8 und Abbildung 9 legen die Suchgewohnheiten der Studierendengruppe nach neuen Lernmaterialien unabhängig von der Lehrveranstaltung „Theoretische Informatik“ dar. Die primäre Anlaufstelle für die Suche nach neuen Materialien ist das Internet. Die Studierenden konnten ihre Antworten frei wählen und benannten konkret die Suchmaschine Google und den Videodienst YouTube als primäre Quelle. Erst danach wurde der allgemeine Begriff Internet und nur von einigen die Bibliothek als Anlaufstelle genannt. Abbildung 9 zeigt nach welcher Art von Materialien gesucht wird. Die Antwortmöglichkeiten waren vorgegeben und es konnten mehrere Antworten gegeben werden. In erster Linie wird nach Videomaterial und Internetseiten allgemein gesucht, obgleich Skripte und Vorlesungsaufzeichnungen in ähnlicher Form von Interesse sind. Nach Literatur wird im Allgemeinen weniger gesucht.

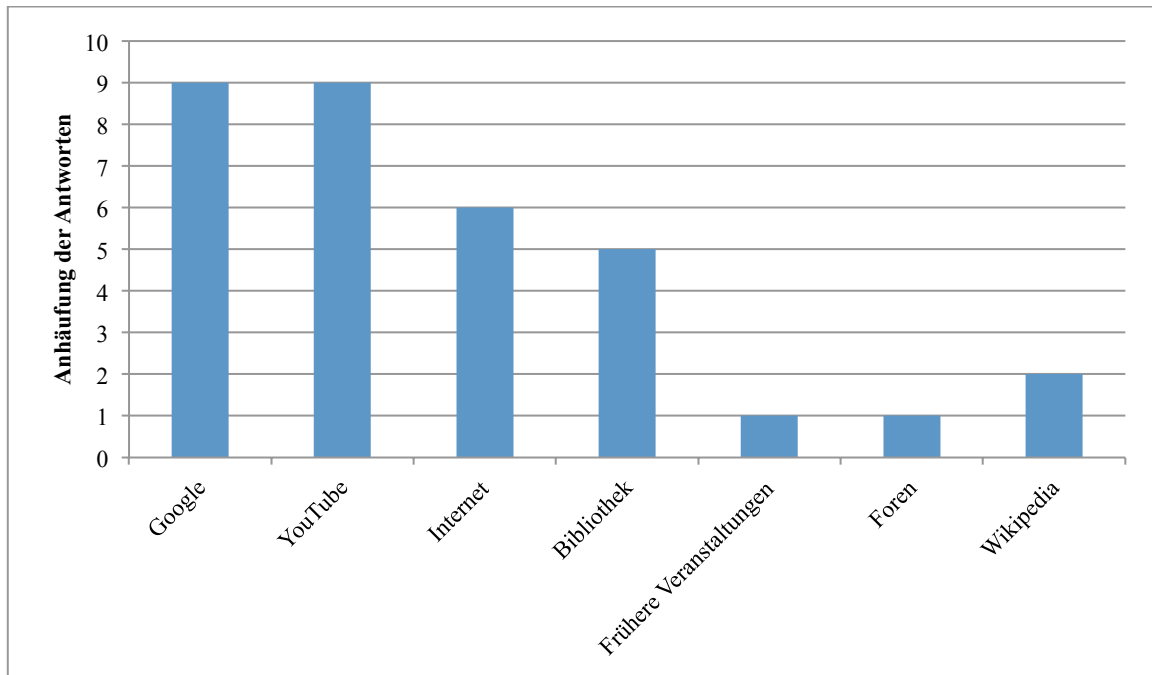


Abbildung 8 Anlaufstelle zur Suche neuer Lernmaterialien

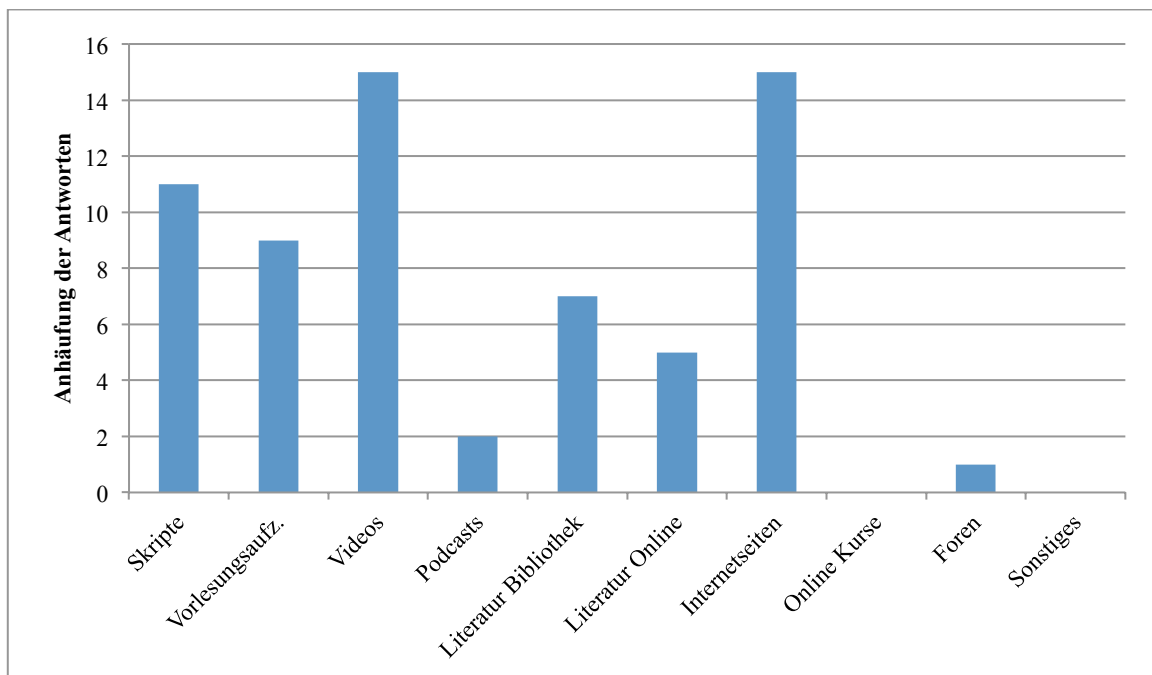


Abbildung 9 Suche im Internet nach Lernmaterialtypen

2.5 Diskussion der Ergebnisse

Kapitel 2.4 hat die erhobenen Ergebnisse präsentiert, die im nachfolgenden Abschnitt zur Diskussion gestellt werden. Betrachtet man zunächst die Nutzungshäufigkeit der Lernmaterialien

abgeleitet aus den Lerntagebüchern, so ist das Skript als zentraler Ausgangspunkt zum Lernen bestimmt worden. Die Bewertung des Skriptes bestätigt die Aussage, dass die Wichtigkeit des Skriptes besonders hoch bewertet wurde. Interessant ist in diesem Zusammenhang, dass das Skript ein recht geringes Verständnis sowie Unterstützung aufweist. Das bereitgestellte multimediale Material ist vor allem im Punkt Verständnis besser bewertet. Ohnehin wurden den YouTube Videos der höchste Grad an Verständnis und Unterstützung zugeordnet. Dies scheint insofern ersichtlich, als dass die YouTube Videos direkt mit dem Inhalt aus dem Skript verknüpft sind und somit als ergänzende Erläuterung zum schriftlichen Inhalt dienen. Die Verknüpfung erfolgt, wie in Abbildung 10 zu sehen in Form eines QR-Tags. In der digitalen Form des Skriptes kann der QR-Tag auch angeklickt werden, um den dahinterstehenden Link zu öffnen.

3.1. Grammatiken

Viele Sprachen aus der Informatik werden durch Grammatiken beschrieben. Eine Grammatik definiert auf Basis von Symbolen und Zeichen eine Menge von Umformungsregeln, mit deren Hilfe eine Sprache erzeugt wird. Dabei wird ausgehend von einem Startsymbol oder einer Startvariable durch sukzessive Anwendung der Regeln ein Wort erzeugt, welches aus sogenannten Variablen oder Terminalen besteht. Die Terminalen werden nach ihrer Erzeugung üblicherweise nicht weiter ersetzt und die Umformung endet, sobald man ein nur noch aus Terminalen bestehendes Wort erhält.

Definition 3.1.1 (Grammatik) Eine Grammatik G ist ein Vier-tupel $G = (V, T, R, S)$. Sie besteht aus

- V eine endliche Menge von Variablen (Nonterminals)
- T eine endliche Menge von Terminalen
- R eine endliche Menge von Regeln. Eine Regel ist ein Element (P, Q) aus

$$((V \cup T)^* V (V \cup T)^*) \times (V \cup T)^*$$

Für eine Regel $(P, Q) \in R$ schreibt man üblicherweise $P \rightarrow Q$.

- S das Startsymbol, $S \in V$.



Grammatik

Abbildung 10 Verweis auf YouTube Video im Skript

Aber auch die Kürze der YouTube Videos scheint hier ein entscheidender Erfolgsfaktor zu sein. Die Vorlesungsaufzeichnungen sind so z.B. als deutlich weniger schwierig bewertet. Sie weisen im Vergleich zu den Videos allerdings ein geringeres Verständnis sowie eine geringere Unterstützung auf. Zieht man nun die Nutzungsdauer der beiden Materialtypen heran, so ist ersichtlich, dass die Vorlesungsaufzeichnungen nur selten vollständig betrachtet wurden. Viele der Studierenden haben sogar angegeben die Aufzeichnungen weniger als 30 Minuten pro Woche zu nutzen. Es liegt somit die Vermutung nahe, dass aus den Aufzeichnungen nur bestimmte Teile als Ergänzung angeschaut werden. Der bereitgestellten Literaturhinweise wird

insgesamt nur eine geringe Bedeutung beigemessen, da sowohl die Nutzungshäufigkeit als auch die Bewertungen darauf hinweisen.

Die Rolle der eigens gesuchten Lernmaterialien scheint auf den ersten Blick nicht ganz nachvollziehbar. Diesen wurde ein hohes Maß an Verständnis und Unterstützung bestätigt, betrachtet man allerdings die geringe Ausprägung der Wichtigkeit, so scheinen die selbst gesuchten Materialien nur eine untergeordnete Rolle zu spielen. Und auch wenn die Nutzungszeiten der eigenen Materialien diesen Eindruck bestätigen, investieren die Studierenden dennoch Zeit für die Suche nach neuen Materialien. Die Intention der Suche liegt in den Aussagen der Studierenden, welche Informationen in den bestehenden Materialien fehlen oder gewünscht werden. Es werden vor allem zusätzlich Beispiele (ggf. mit Lösung), alternative oder zusätzliche Erklärungen sowie ausführlichere Lösungsansätze gefordert. Studierende erwarten diese Informationen in weiteren Lernmaterialien zu finden. Die Materialtypen die von Studierenden zum Lernen gesucht werden entstammen zum Großteil aus dem universitären Lernkontext - so z.B. Skripte, Vorlesungsaufzeichnungen und Lernvideos, und decken den gesuchten Bedarf potentiell ab. Auch die Hauptquellen für die Suche – Google und YouTube, präsentieren eine Vielzahl möglicher Lernmaterialien aus verschiedensten universitären Lehrinstitutionen unter dem Begriff „Theoretische Informatik“. Dies wurde in gleicher Weise bereits zu Beginn der Arbeit motiviert. Die entscheidende Frage an dieser Stelle ist nur, ob Studierende einen tatsächlichen Nutzen aus den alternativen Materialien ziehen. Dieser Nutzen oder zumindest ein angemessener Zeitaufwand zur Suche darf wie zu Beginn der Arbeit in Frage gestellt werden. Aufgrund der Anhäufung an vielen verschiedenen Materialien, die im Netz zur Verfügung stehen erscheint die Suche und der damit einhergehenden Sichtung und Bewertung von Materialien sowohl schwierig als auch zeitintensiv. So sind z.B. die Skripte und Foliensätze, die man über Google finden kann, zumeist sehr umfangreich und didaktisch in der Regel unterschiedlich zum bestehenden Skript aufgearbeitet. Neben Variation von Sprache und Ausprägung der inhaltlichen Erläuterungen ist das Niveau häufig unterschiedlich. Im Kontext der theoretischen Informatik sind des Weiteren formale Unterschiede in den Lernmaterialien auszumachen. Studierende müssten zumindest diese Unwägbarkeiten überwinden, um die gewünschten Informationen zu extrahieren. Offen bleibt in diesem Zusammenhang, in wieweit eine etwaige Suche – z.B. bei Misserfolg oder falscher Beurteilung neuer Materialien – zu

negativer Lernmotivation oder Prokrastination führt. Dies kann durch die vorliegenden Ergebnisse auch nicht bewertet werden. Durch die Arbeit in [3] kann allerdings bestätigt werden, dass zumindest ein Überangebot von Lernmaterialien zu Desorganisation und mentale Überforderung führt.

2.6 Schlussfolgerung

Welche Schlussfolgerungen können abschließend für die Entwicklung der vorliegenden Arbeit abgeleitet werden? Sicher erscheint, dass Studierende nach neuen Lernmaterialien suchen und dies unabhängig davon, ob bereits ausreichend unterstützende Materialien zur Verfügung stehen oder nicht. Die in der Untersuchung präsentierten und genutzten Lernmaterialien boten bereits eine große Vielfalt und neben vielerlei Übungen und Beispielen direkte Verknüpfungen zwischen Lernmaterialien. Dennoch suchten Studierende nach weiteren, alternativen Beispielen und Erläuterungen die zusätzlich zum bestehenden Material genutzt werden konnten. Es kann vermutet werden, dass dieses Vorgehen in einem Lernsetting mit weniger Hilfestellungen noch ausgeprägter auftritt. Sicher scheint auch, dass Studierende mit einer direkten inhaltlichen Verbindung zu ergänzenden Materialien (z.B. den YouTube Videos) eine größere Unterstützung und eine vereinfachte Wissensergänzung erfahren. In diesem Zusammenhang ist auch die Kürze der Videos von Bedeutung. Kurze und inhaltlich verknüpfte Videos ermöglichen eine sehr genaue und zielgerichtete Vermittlung von ergänzenden Informationen und erleichtert den Konsum für die Studierenden. Dies wäre im Kontext von Vorlesungsaufzeichnungen ebenfalls sinnvoll da Studierende Aufzeichnungen auch als Ergänzung und zum Wiederholen von dedizierten Themenbereichen verwenden [59]. Das für Vorlesungsaufzeichnungen eingesetzte System Opencast Matterhorn bietet die Möglichkeit über den Kontext von Präsentationfolien Themenbereiche schnell und zielgerichtet auszuwählen. Dies stellt für Studierende bereits eine erhebliche Vereinfachung beim Lernen dar. Die direkte Verknüpfung zu Inhalten anderer Lernmaterialien könnte ein weiterer Mehrwert für Studierende darstellen. Dies gilt in gleicher Form auch für Lernmaterialien aus dem Internet. Studierende verbringen regelmäßig Zeit mit der Suche nach zusätzlichen Materialien, obgleich nur bestimmte Bereiche relevant erscheinen. Sinnvoll wäre auch hier eine Verknüpfung zwischen relevanten Teilbereichen der Materialien, um Studierenden einen direkten Zugang zu ermöglichen. Im Idealfall würde dies für Materialien gelten die bereits einer Prüfung unterzogen und als hilfreich oder unterstützend ausgemacht

wurden. Für die durchzuführende Entwicklung ergibt sich somit die Anforderung, Verknüpfungen zwischen Inhalten bestehender und neuer, alternative Lernmaterialien zu generieren, so dass Studierende einen einfachen und direkten Zugang zwischen zugehörigen Lerninhalten erhalten.

Was an dieser Stelle nicht vernachlässigt werden sollte ist die Tatsache, dass bereitgestellte Materialien – davon sollte zumindest ausgegangen werden – eine genügende Grundlage für die inhaltliche Bearbeitung der Veranstaltung bieten. Die besondere Bedeutung von Lernmaterialien des Lehrenden der Veranstaltung konnte zumindest im Kontext der Untersuchung durch die primäre Nutzung der bestehenden Materialien und deren bewertete Wichtigkeit belegt werden. Es ist deshalb erforderlich die bestehenden Lernmaterialien als zentralen Ausgangspunkt zum Lernen zu betrachten und die Studierenden mehr an die bereits bestehenden Materialien zu binden sowie die Nutzung stärker zu befördern. Dies kann gelingen, indem elementare Inhalte innerhalb der Materialien gekennzeichnet und Informationen über die Schwierigkeit und der Bedeutung von Inhalten (z.B. in Form von Empfehlungen innerhalb eines Empfehlungsdiensteszenarios) bereitgestellt werden. Bei offensichtlichen Verständnisproblemen oder Schwierigkeiten könnten alternative Inhaltsverknüpfungen das Material ergänzen und den Bedarf einer Suche minimieren. Hierdurch ergibt sich als Anforderung an die bevorstehende Entwicklung eine Möglichkeit zur inhaltlichen Bewertung durch die Studierenden, so dass vorhandene Verständnisprobleme und Schwierigkeiten aufgedeckt werden können.

3 Grundlagen und Stand der Forschung

Der Begriff *Empfehlungsdienst* fiel im Kontext der vorliegenden Arbeit bereits einige Male und ist bislang sehr allgemein als ein assistierendes Softwaresystem definiert. Das nachfolgende Kapitel wird sich mit der Vermittlung von Grundlagen beschäftigen, so dass von einer allgemeinen zu einer technisch formalen Betrachtungsweise übergegangen werden kann. Es wird sich zeigen, dass das Forschungsfeld der Empfehlungsdienste um ein Vielfaches komplexer ist als die erste Darstellung vermuten mag. Aus verschiedenen Anwendungskontexten oder Anwendungsgebieten ergeben sich ebenso verschiedene Anforderungen an einen Empfehlungsdienst, weshalb eine Vielzahl an Ansätzen existiert. Vor allem die zugrundeliegende Datenbasis und die zu erreichenden Ziele erheben entscheidenden Einfluss auf die Art des Ansatzes. Dieser Tatbestand wurde in den ersten Kapiteln der vorliegenden Arbeit bereits kurz thematisiert. Aus diesem Grund liegen dem übergeordneten Begriff Empfehlungsdienst eine Vielzahl weiterer Forschungsfelder zugrunde, die weitestgehend als Grundlage für den Bereich der Empfehlungsdienste zu sehen sind und deren Techniken in verschiedenster Form miteinander kombiniert werden. Hierzu gehören z.B. der Bereich des *Data Mining* und des *Information Retrieval*, aber auch die Bereiche *Statistik* und *maschinelles Lernen*. Die prägendste Rolle für Empfehlungsdienste der heutigen Zeit nimmt wohl der Forschungsbereich des Data Mining ein (vgl. [129] S. 21). Abschnitt 3.2 zeigt auf welche Verfahren für Empfehlungsdienste existieren und in wieweit diese mit den genannten Forschungsfeldern zusammenhängen. Vor der Betrachtung konkreter Verfahren bedarf es allerdings der Vermittlung einiger grundlegender Sachverhalte über den Aufbau, die übergeordneten Ziele und die vorherrschende Terminologie des Forschungsfeldes. Abschnitt 3.1 wird sich der Vermittlung dieser Sachverhalte annehmen.

Des Weiteren wird in Abschnitt 3.3 das Themen- bzw. Forschungsfeld *Social Tagging Systems (STS)* betrachtet. Das Forschungsfeld hat zunächst keinen direkten Bezug zum Bereich der Empfehlungsdienste und mag deshalb deplatziert wirken. STS nehmen allerdings einen immer größeren Stellenwert im Forschungsfeld der Empfehlungsdienste ein, da sie sich zur Datenerhebung und -herleitung in verschiedenen Ansätzen als wertvolle Ergänzung erwiesen haben. Es wird sich zeigen, dass auch die vorliegende Arbeit einen Nutzen aus STS zieht (vgl. hierzu Abschnitt 4). Abschnitt 3.3 erläutert zu Beginn wie STS mit der vorliegenden Arbeit in Zusammenhang steht. Des Weiteren sind in Abschnitt 3.3 Grundlagen für den Forschungsbereich STS thematisiert.

Die Abschnitte 3.1, 3.2 und 3.3 sind für die Vermittlung von Grundlagen zu verstehen, auf denen die Abschnitte 3.4 und 3.5 aufbauen. Abschnitt 3.4 versucht das Feld der Empfehlungsdienste im Kontext E-Learning zu eröffnen und eine Abgrenzung zu weiteren Forschungsbereichen durchzuführen. Abschnitt 3.5 dient dann der Erarbeitung einer Argumentationsgrundlage zur Herleitung eines Verfahrens für die vorliegende Arbeit, in dem relevante Ansätze von Empfehlungsdiensten der Anwendungsdomäne E-Learning präsentiert und diskutiert werden.

3.1 Empfehlungsdienste Grundlagen: Terminologie und Ziele

Der Begriff Empfehlungsdienst¹⁵ tritt im allgemeinen Sprachgebrauch nur selten auf und ist wohl eher in entsprechenden Fachgruppen ein gängiger Begriff. Dabei beeinflussen Empfehlungsdienste in den verschiedensten alltäglichen Anwendungen die Entscheidungen von Personen, ohne dass bei diesen Personen ein Bewusstsein dafür existiert. Als einer der bekanntesten Empfehlungsdienste sei der des online Versandhandels Amazon.com¹⁶ genannt [94]. Der Benutzer des Versandhandels sieht sich einer enormen Auswahl an potentiell interessanten Artikeln gegenüber, die für einen Kauf durch den Kunden gesichtet und bewertet werden müssten. Der im Onlineshop integrierte Empfehlungsdienst wirkt hierbei als assistierendes System, welches dem Kunden solche Artikel vorschlägt, die zum Zeitpunkt des

¹⁵ Im deutschen werden häufig auch die englischen Begriff Recommender System (RS) oder Recommendation System (RS) synonym verwendet. Empfehlungssystem ist ebenfalls ein gängiger Begriff.

¹⁶ Offizielle Internetseite: <http://www.amazon.com>

Besuchs im Shop – und darüber hinaus – interessant sein könnten. So werden E-Mails mit potentiell interessanten Produkten versendet oder Artikel werden mit vergleichbaren Produkten in Zusammenhang gebracht. Der Käufer erhält somit die Möglichkeit, personenbezogene Produktempfehlungen zu erhalten, die im besten Fall den Geschmack oder den Nutzen des Käufers treffen und so ggf. zum Kauf führen. Bis heute hat sich eine Vielzahl weiterer Einsatzgebiete von Empfehlungsdiensten erschlossen, die sich gemäß ihrer produzierten Empfehlungen typisieren lassen (vgl. [129] S. 14 ff):

- *Entertainment*: Recommendations for movies, music and IPTV
- *Content*: personalized newspapers, recommendations for documents, recommendations of Web Pages, e-Learning applications, and e-mail filters
- *E-Commerce*: Recommendations for consumers of products to buy such as books, cameras, PCs etc.
- *Services*: Recommendations of travel services, recommendation of experts for consultation, recommendation of houses to rent, or matchmaking services.

Es sei an dieser Stelle der Versuch einer Definition erlaubt, um sich bestehenden Techniken und Verfahren anzunähern. In [128] ist ein Empfehlungsdienst als Benutzer unterstützendes System beschrieben, welches auf dem *Word of Mouth*¹⁷ – auch *Mund-zu-Mund-Kommunikation* oder auch *Mundpropaganda* – Prinzip beruht:

„It is often necessary to make choices without sufficient personal experience of the alternatives. In everyday life, we rely on recommendations from other people either by word of mouth, recommendation letters, movie and book reviews printed in newspapers [...] Recommender systems assist and augment this natural social process.“

Interessant ist die Tatsache, dass der zitierte Artikel bereits vor dem Beginn des Informationszeitalters einen starken Einfluss durch Informationssysteme im alltäglichen Leben

¹⁷ Word of Mouth kann prinzipiell als Marketingkonzept verstanden werden, welches die informelle Kommunikation zwischen Kunden über Produkte und Dienstleistungen vorsieht und daraus Nutzen zieht (vgl. [95] S. 5).

vorhergesagt hat. Eine stark am Konsum ausgerichtete Definition stammt von Jannach im Jahr 2008 (vgl. [70]):

„Recommender systems are interactive software applications that support the online customer in his/her decision making and buying process.“

Eine ebenfalls sehr treffende und verhältnismäßig aktuelle Definition, die sich bereits der Funktionsweise von Empfehlungsdiensten nähert, stammt von Anand und Ullmann aus dem Jahr 2011 (vgl. [123] S. 307):

„There is an extensive class of Web applications that involve predicting user responses to options. Such a facility is called a recommendation system.“

Es gibt eine Vielzahl weiterer Definitionen für Empfehlungsdienste, allerdings treffen alle eine grundlegend gleiche Aussage: Empfehlungsdienste helfen bei der Aufbereitung einer zumeist großen Menge an Daten oder Informationen, um sie in Abhängigkeit von Interessen oder eines bestimmten Verhaltens leichter für den Endbenutzer zugänglich zu machen. Herlocker und Konstan beschreiben in [61] insgesamt elf *Ziele* oder auch *Funktionsweisen*¹⁸, die ein Empfehlungsdienst zur Verfügung stellen kann und berücksichtigen in erster Linie solche, die dem Endbenutzer zuträglich sind. Folgende Ziele eines Empfehlungsdienstes werden dabei genannt:

- *Find Some Good items*: Empfiehlt einem Benutzer einige nach Nutzen gewichtete Objekte.
- *Find all good items*: Empfiehlt einem Benutzer alle Objekte, die einen Nutzen enthalten können.
- *Annotation in context*: Stellt Informationen innerhalb eines bestimmten Kontextes bereit.
- *Recommend a sequence*: Empfiehlt einem Benutzer sequentiell aufeinander aufbauende Objekte bereit z.B. themenverwandte Bücher.

¹⁸ Auch wenn Herlocker und Konstan in [61] ihre Untergliederung als Funktionsweisen beschreiben, werden diese nachfolgend ebenso als Ziele betrachtet und bezeichnet.

- *Recommend a bundle*: Empfiehlt einem Benutzer eine Ansammlung von Objekten, die in einem Zusammenhang stehen.
- *Just browsing*: Der Benutzer durchsucht einen Katalog an Objekten, wobei der Empfehlungsdienst als eine Art Wegweiser durch den Katalog fungiert.
- *Find credible recommender*: Bereitstellung einer Funktionalität, um die Funktionsweise eines Empfehlungsdienstes nachvollziehbar zu machen. Dies soll die Vertrauenswürdigkeit eines Empfehlungsdienstes steigern.
- *Improve the profile*: Erhebung benutzerbezogener Interessen, um personalisierte und benutzerbezogene Empfehlungen herleiten zu können.
- *Express self*: Bietet für Benutzer die Möglichkeit Aussagen – z.B. Bewertungen – über Objekte zu treffen.
- *Help others*: Bietet für Benutzer die Möglichkeit anderen Benutzern durch eigene Aussagen eine Art „Hilfe“ zukommen zu lassen. Hilfe bedeutet in diesem Zusammenhang einen Mehrwert für andere Benutzer durch eigene Aussagen – z.B. Bewertungen – zu erbringen.
- *Influence others*: Bietet die Möglichkeit andere Benutzer durch eigene Aussagen – z.B. Bewertungen – zu beeinflussen z.B. beim Kauf von Produkten.

Denkbar ist ebenso eine Betrachtungsweise aus Sicht eines Empfehlungsdiensteanbieters, wie z.B. der Betreiber eines online Shops, wo kommerzielle Aspekte stärkere Berücksichtigung finden. Im Kontext einer Forschungsarbeit ist die Ausführung in [61] allerdings als sinnvoller einzuschätzen. Zum einen spielen kommerzielle Aspekte in Forschungsfragen oftmals eine untergeordnete Rolle, was zumindest für solche gilt, die sich in erster Linie mit der technischen Umsetzung von Empfehlungsdiensten beschäftigen. Des Weiteren wird sich zeigen, dass die detailreiche und weitestgehend technisch geprägte Darstellung von möglichen Zielen zu einem späteren Zeitpunkt eine sinnvollere Eingliederung von Forschungsansätzen ermöglicht.

Welche Aufgabe ein Empfehlungsdienst schlussendlich erfüllen soll, hängt auch von den strukturellen Gegebenheiten ab, die für die Einbettung eines etwaigen Systems zur Verfügung stehen. Mit strukturellen Gegebenheiten sind im Grunde die Fragestellungen nach der Benutzergemeinde und den zu empfehlenden Objekten (z.B. Artikel in einem Onlineshop) gemeint. Darüber hinaus ist die Relation zwischen einem Benutzer und einem Objekt von

Bedeutung. In diesem Zusammenhang müssen auch unweigerlich die zuvor genannten Ziele eines Empfehlungsdiensteanbieters berücksichtigt werden. Burke betrachtet hierzu in [22] einen Prozess, der bei der Konzeption eines Empfehlungsdienstes betrachtet werden sollte. So hat ein System zunächst Informationen, die vor dem Empfehlungsprozess vorliegen (z.B. Informationen über im System vorliegende Objekte). Im zweiten Schritt interagiert der Benutzer mit dem System und gibt für den Empfehlungsprozess Informationen über sich preis. Der abschließende Empfehlungsprozess kombiniert beide Informationen und generiert auf Basis der vorliegenden Daten Empfehlungen. Über den skizzierten Prozess können drei für die Konzeption eines Empfehlungsdienstes zu betrachtende Gruppen oder Mengen abgeleitet werden. In [129] sind diese als *Item*, *User* und *Transaction* bezeichnet (vgl. S. 7 ff.). Betrachtet man die zu Beginn des Abschnitts zitierten Definitionen zu Empfehlungsdiensten, so erkennt man, dass die Terminologie einem gewissen Standard entspricht und in vielen Artikeln und Fachbüchern in gleicher Form genutzt wird. Auch in der vorliegenden Arbeit seien in den nachfolgenden Abschnitten und Kapiteln die englischen Begriffe *Item*, *User* und *Transaction* genutzt, um diese gemäß der Forschungscommunity zu bezeichnen. Nachfolgend sind die Begriffe *Item*, *User* und *Transaction* definiert.

Item: Ein *Item* stellt ein für den *User* sichtbares Systemobjekt dar, welches abhängig vom Anwendungskontext variiert. Ein *Item* ist in der Regel immer das durch den Empfehlungsdienst zu empfehlende Objekt. Welche Informationen für ein *Item* vorliegen und in welcher Form diese zur Verfügung stehen, variiert je nach Anwendungskontext. Bücher in einem Onlineshop werden so z.B. durch ihre Attribute wie Titel, Autor oder Genre beschrieben. Dies gilt für eine Vielzahl an Anwendungen, sofern sich die jeweiligen *Items* durch Attribute beschreiben lassen. Deutlich schwieriger wird die Deklaration, wenn sich ein *Item* erst aus kausalen Zusammenhängen ergibt. Dies gilt z.B. bei Versicherungen oder Geldanlagen (vgl. z.B. [107]). Ein *Item* ist in diesen Fällen vorab nicht vollständig definiert, sondern hängt von situationsabhängigen Informationen ab. Die Auflistung der Attribute genügt hier nicht mehr, so dass es einer Verknüpfung z.B. durch vordefinierter Regeln bedarf. Die Vergleichbarkeit von *Items* ist in diesen Fällen somit deutlich aufwändiger.

User: Ein User ist in der Regel der Endanwender des Systems, in dem ein Empfehlungsdienst implementiert ist. Durch die Interaktion mit dem System wird der Empfehlungsprozess gestartet. Dies geschieht nicht zwangsläufig durch ein bewusstes Einwirken durch den User, sondern oftmals aufgrund gewöhnlicher Transaktionen im System. Man unterscheidet hierbei auch zwischen *implizitem* (z.B. der Kauf eines Artikels) und *explizitem* (z.B. der Bewertung eines Artikels) Feedback, welches im System erfasst werden kann [2, 53]. Vielfach wird versucht beide Informationsgruppen abzufragen, um detailliertere Informationen über einen User zu erlangen und bessere Empfehlungen ableiten zu können [72]. Über das implizite oder explizite Feedback eines Benutzers wird ein Benutzerprofil aufgebaut, welches als Grundlage für den Empfehlungsprozess dient. In der Regel besteht ein Benutzerprofil aus einer Ansammlung von Items, die aus impliziten und expliziten Informationen abgeleitet werden. Aber auch die Speicherung der User Interaktionen (z.B. Bewertungen) selbst ist denkbar, sofern diese mit dem jeweiligen Item verknüpft sind. Dem Benutzerprofil kommt eine besondere Bedeutung zu, da je nach Umfang bzw. Detailgrad des Benutzerprofils die Empfehlungen in ihrer Qualität variieren können. Der Grad an Details ist hierbei ein entscheidendes Kriterium. Umso mehr Informationen über den User bekannt sind, umso präziser können Präferenzen oder Geschmäcker abgeleitet werden. Da sich ein Benutzerverhalten von Zeit zu Zeit ändern kann, bedarf es allerdings einer regelmäßigen Anpassung der Benutzerinformationen. Die Bewertung eines bestimmten Buches muss so z.B. nach einigen Jahren nicht unbedingt die aktuelle Meinung des Nutzers widerspiegeln. Ähnlich wie bei Items in einem System ist die Darstellung von Usern oder vielmehr das Profil eines Users je nach Anwendungsfall sehr unterschiedlich. Die Bewertung von Artikeln in einem Onlineshop ist als Benutzerprofil somit eher trivial abbildbar. Es existieren aber auch Anwendungsfälle, bei denen entscheidend ist, wie sich ein User in einer bestimmten Situation entschieden hat oder in denen eine zeitsensitive Abbildung von Daten notwendig ist. Dann sind wiederum komplexere Modelle notwendig, die eine derartige Abbildung von Benutzerdaten ermöglichen. Klahold unterscheidet noch zwischen einem *flüchtigen* und einem *persistenten* Benutzerprofil, das entweder eine begrenzte Zeitspanne oder langfristige Gültigkeit besitzt (vgl. [82] S. 22). Der Artikel von Höhfeld gibt zu bedenken, dass ein Empfehlungsdienst grundsätzlich auch ohne eine personalisierte Betrachtung auskommt, wie z.B. bei Suchmaschinen [64]. Etwaige Systeme liefern Informationen unabhängig vom Benutzer,

der die Information angefragt hat. Nachfolgend liegt der Fokus allerdings auf personalisierten Systemen.

Transaction: Wie die vorherigen zwei Definitionen der User bzw. Item-Gruppe bereits implizieren, ist eine Transaction eine Interaktion zwischen Benutzer und System oder in der Regel sogar direkt zwischen User und Item. Welche Art von Interaktion vorliegt und ob eine Interaktion überhaupt zur Verfügung steht, ist wiederum abhängig vom Anwendungskontext. Logischerweise bietet nicht jedes Anwendungsgebiet z.B. den Kauf eines Items an, so dass nicht jede Interaktion in jedem Anwendungskontext zur Verfügung steht. Eine generalisierte Form von Transactions – in besonderer Weise für implizite Rückmeldungen – lässt sich deshalb nur schwer ableiten. Dennoch kommt Transactions eine zentrale Rolle in Empfehlungsdiensten zu. Die über Transactions gesammelten Informationen dienen als Grundlage zur Herleitung von (personalisierten) Empfehlungen, weshalb der Bestimmung von geeigneten impliziten und expliziten Transactions eine genügende Bedeutung beigemessen werden sollte. Wo es für die Bestimmung von impliziten Informationen der genaueren Betrachtung des jeweiligen Anwendungsfalls bedarf, hat sich in vielen Systemen als explizite Rückmeldung durch den Benutzer der Einsatz von Bewertungen bewährt. Diese können in numerischer (z.B. von 1 bis 5) oder in binärer Form (z.B. gut oder schlecht) vorliegen [138]. Einer der Hauptvorteile von Bewertungen liegt in der einfachen Interpretation durch das System, obgleich die Aussagekraft eines numerischen Wertes nur bedingt hoch ist. Sofern eine inhaltliche Beschreibung durch den Benutzer erwartet wird, haben sich Tags als probates Mittel der Interaktion erwiesen [163].

Abbildung 11 zeigt zusammenfassend den generellen Aufbau einen Empfehlungsdienstes.

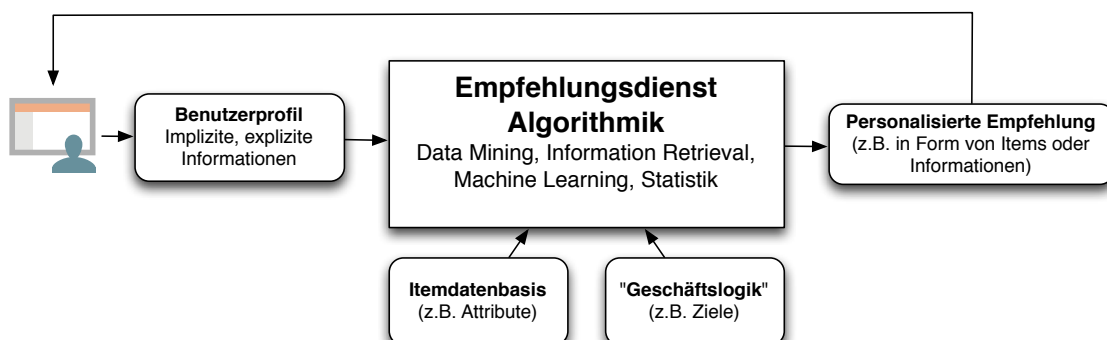


Abbildung 11 Aufbau Empfehlungsdienst

Ausführlich diskutiert wurde bereits die Bedeutung von User- und Iteminformationen, wozu auch die Relation zwischen beiden Gruppen, also die der Transactions gehört. Beides ergibt schlussendlich die Datenbasis des Systems, von der eine zentrale Bedeutung zur Herleitung von Empfehlungen ausgeht (vgl. hierzu Abschnitt 3.2). Auch der personalisierte Output des Systems, welcher in direktem Zusammenhang mit der vorherrschenden Geschäftslogik steht, wurde bereits thematisiert. Der Begriff Geschäftslogik suggeriert an dieser Stelle die Implementierung eines kommerziellen Systems, was allerdings nicht der Fall sein muss. Prinzipiell ist die Umsetzung der geforderten Ziele auch auf Grundlage der zur Verfügung stehenden Datenbasis gemeint, die nicht zwangsläufig kommerzielle oder geschäftsmäßige Absichten vorhalten müssen. Zuletzt bleibt der Bereich der Empfehlungsdienst Algorithmik offen, der wohl der umfassendsten Erläuterung bedarf. Der Bereich ist für die Bündelung aller zur Verfügung stehenden Informationen verantwortlich, um auf dieser Grundlage neue Informationen herzuleiten. Der nachfolgende Abschnitt 3.2 wird sich dieser Thematik annehmen.

3.2 Empfehlungsdienste Grundlagen: Verfahrensarten

Der vorherige Abschnitt hat in einer allgemeinen Form die Thematik der Empfehlungsdienste eröffnet. Neben einem einführenden Beispiel und der Bereitstellung einer geeigneten Definition wurde kurz skizziert, in welcher Form Empfehlungsdienste von Nutzen sind und welche Faktoren auf die Entwicklung eines Empfehlungsdienstes einwirken. In Kapitel 3.1 wurde bereits ein dreistufiger Prozess von Burke zitiert, der von einem Vorwissen über die im System befindlichen Items ausgeht, eine Benutzerinformation erwartet und beide Informationen bzw. Informationsgruppen kombiniert, um Empfehlungen ableiten zu können (vgl. [22] S. 2). Das aktuelle Kapitel beschäftigt sich mit dem letzten Schritt des Prozesses, der die Verfahren zur Herleitung von Empfehlungen beschreibt. Bislang ist nämlich noch ungeklärt, wie ein Empfehlungsdienst eine Entscheidung darüber trifft, ob ein Item für einen User empfehlenswert ist oder nicht. Hierfür sind Bewertungsverfahren notwendig, die in Abhängigkeit von zur Verfügung stehenden (Item-, User-) Informationen eine Aussage über die Nützlichkeit, den

Nutzen oder auch die Wichtigkeit¹⁹ eines Items für einen bestimmten User trifft. Abschnitt 3.5 adaptiert diese Thematik für den Anwendungskontext E-Learning. Nach Adomavicius und Tuzhilin ist die Aufgabe eines Empfehlungsdienstes eine Vorhersage darüber zu treffen, ob ein unberücksichtigtes Item für einen Benutzer nützlich ist oder nicht und kann formal wie folgt definiert werden (vgl. [2] S. 734 ff.):

Sei U die Menge aller User und sei I die Menge aller möglich zu empfehlenden Items im System. Sei R eine reellwertige Funktion zur Bestimmung der Nützlichkeit eines Items i für einen User u . z.B. über $R: U \times I \rightarrow V$, wobei V eine Menge an Wertigkeiten aus einem vordefinierten Wertebereich darstellt. Für jeden User $u \in U$ wird schlussendlich versucht die Items $i'_u \in I$ auszuwählen, die eine maximale Nützlichkeit oder maximalen Mehrwert für einen jeweiligen Nutzer ergeben:

$$i'_u = \arg \max_{i \in I} R(i, u) \quad (1)$$

Zur Verdeutlichung sei in Tabelle 1 ein Beispiel gezeigt, in dem Personen (User) Filme (Items) von 1 (schlecht) bis 5 (gut) bewertet haben. Filme die nicht bewertet wurden, sind mit der Wertigkeit 0 gekennzeichnet.

	Star Wars	Der Pate	Avatar	Forrest Gump
Alice	5	2	4	0
Bob	0	5	3	1
Cindy	3	4	2	5
David	2	0	0	4

Tabelle 1 Beispiel Bewertungstabelle für Filme

Die Idee eines Empfehlungsdienstes besteht nun darin eine Bewertung für die noch unberücksichtigten Items (im Beispiel Filme) auf Grundlage bestehender Bewertungen zu ermitteln bzw. zu schätzen und solche Items (Filme) zu präsentieren, die den größten Nutzen für einen User darstellen. So existiert für etwaige Systeme eine Zuordnung zwischen den Mengen User und Item, für die sich eine Zuordnungsmatrix wie in Tabelle 1 bewährt hat. Eine formale

¹⁹ Im englischen Sprachgebrauch ist der Begriff *utility* gebräuchlich. Nachfolgend wird in dieser Arbeit von Nützlichkeit gesprochen.

Darstellung einer User-/Item-Matrix sei in Abbildung 12 dargestellt, in der jede Zelle der Matrix einer Bewertung eines Users $u \in U$ für ein Item $i \in I$ entspricht.

	i_1	i_2	...	$i_{ I }$
u_1				
u_2		$R(i, u)$		
...				
$u_{ U }$				

Abbildung 12 User-Item-Matrix

Die Herausforderung für einen Empfehlungsdienst liegt nun in der Umsetzung der Funktion $R(i, u)$ zur Bewertung der Nützlichkeit eines Items i für einen User u , sofern für das entsprechende Item noch keine Bewertung vorliegt. Verschiedene Lösungsansätze haben sich in diesem Zusammenhang etabliert, wobei die Literatur eine Einteilung in *Inhaltsbasierte*-, *Kollaborative*²⁰- und *Hybride*-Verfahren vorschlägt (vgl. z.B. [2, 9, 82]). Die Betrachtungsweise der zur Verfügung stehenden Informations- oder Datenquellen variiert zwischen den Verfahrensarten, so dass sich auf Grundlage der zur Verfügung stehenden Datenbasis unterschiedliche Problemlösungen ergeben. *Inhaltsbasierte Verfahren* vergleichen bereits als nützlich eingestufte Items eines Users mit unberücksichtigten Items auf Basis von Itemeigenschaften. Hierüber können ähnliche Items ermittelt und an einen User empfohlen werden. *Kollaborative Ansätze* versuchen Interessensgruppen auf Seiten der Usergruppe zu bilden, um Items innerhalb der Interessensgruppe weiter zu empfehlen. *Hybride Verfahren* vereinen beide Verfahrensarten, um hierüber bestehende Nachteile gegeneinander aufzuwiegen. Oft werden die Verfahren noch um demografische- oder standortbasierte-Verfahren ergänzt (vgl. z.B. [22, 129]), die nachfolgend allerdings keine Relevanz besitzen.

Wie zu Beginn des aktuellen Kapitels bereits erwähnt, werden in Empfehlungsdiensten verschiedene Techniken aus ebenso verschiedenen Forschungsbereichen (z.B. Data Mining, Information Retrieval, Statistik, Maschinelles Lernen) genutzt und miteinander kombiniert. Die

²⁰ Gängig sind auch hier die englischen Bezeichnungen *Content-based*- und *Collaborative-Filtering*

Techniken kommen dabei für unterschiedliche Zwecke zum Einsatz, weshalb sich das Forschungsfeld Empfehlungsdienste häufig unübersichtlich darstellt. Ein Empfehlungsdienstverfahren kann allerdings grundlegend in drei Schritte eingeteilt und entsprechende Methoden dem jeweiligen Schritt zugeordnet werden. Um ein besseres Verständnis für die Folgeabschnitte zu erwirken, illustriert Abbildung 13 die drei Schritte der *Datenaufbereitung*, *Datenanalyse* und *Empfehlung*. Des Weiteren sind in Abbildung 13 die in Abschnitt 3.2.1, 3.2.2 und 3.2.4 diskutierten Methoden dem jeweiligen Verfahrensschritt zugeordnet²¹.

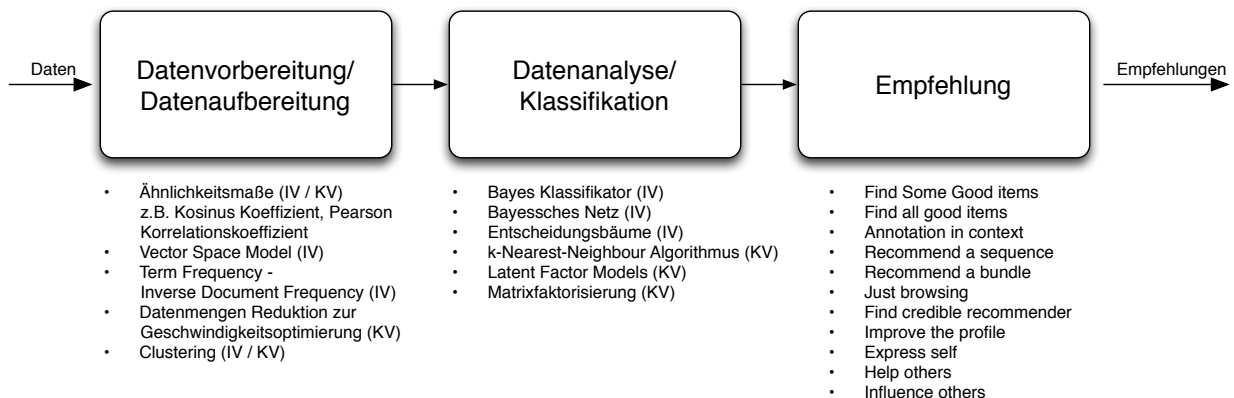


Abbildung 13 Schritte eines Empfehlungsdienst Verfahrens

Eine Datenvor- oder Datenaufbereitung ist je nach Verfahrensart zur Vereinheitlichung einer Datenrepräsentation (vorrangig inhaltsbasierte Verfahren) oder zur Vorberechnung von Ähnlichkeiten in Datenmengen vorgesehen. Der Schritt der Datenanalyse kann grundsätzlich als Schritt zur Umsetzung der Funktion $R(i, u)$ zur Bewertung der Nützlichkeit von Items gesehen werden. Der Schritt der Empfehlung geschieht in Abhängigkeit der in Abschnitt 3.1 angeführten Ziele eines Empfehlungsdienstes und beschreibt die Präsentation eines Items für einen User. In diesem Zusammenhang eine allgemeingültige Menge von Methoden zu präsentieren ist nur schwer möglich, da verschiedene Einflussfaktoren (z.B. Inhalt eines Items, Zeit, Kontext, Emotion, etc.) bei der Präsentation einer Empfehlung zu berücksichtigen sind. Es gibt zwar für einige Anwendungskontexte Methoden die sich z.B. in einem Webshop etabliert haben, die allerdings nicht für jeden Anwendungskontext gleichwohl sinnvoll nutzbar sind. In Abschnitt 3.5

²¹ IV=Inhaltsbasiertes Verfahren, KV=Kollaboratives Verfahren

soll diese Thematik im Kontext E-Learning Berücksichtigung finden, weshalb vorrangig die Schritte und Methoden der Datenaufbereitung und -analyse in den nachfolgenden Teilabschnitten 3.2.1 bis 3.2.4 diskutiert sind.

3.2.1 Inhaltsbasierte Verfahren

Wie in der Einleitung des Abschnittes bereits angedeutet, versuchen inhaltsbasierte Verfahren durch den Vergleich von Itemeigenschaften bzw. Attributen²² weitere nützliche Items herzuleiten. Die Nützlichkeit eines Items i für einen User u ergibt sich aus der Bewertungsfunktion $R(i, u)$. Bei der Ermittlung dieser Größe werden in der Regel die bereits bekannten Nützlichkeiten²³ $R(i_j, u)$ von Item $i_j \in I$ betrachtet, die eine möglichst hohe Ähnlichkeit zu $i \in I$ besitzen. Hierzu werden die Itemattribute genutzt, um eine *Ähnlichkeit* zwischen Items zu ermitteln, die wiederum für eine Analyse und Empfehlung von Bedeutung ist. In einem Beispiel gemäß Tabelle 1 könnten z.B. Titel, Schauspieler, Genre oder Erscheinungsjahr als Attribute eines Items definiert sein. Zur Darstellung der Item Attribute bedarf es einer geeigneten *Datenrepräsentation*, die für einen Vergleich herangezogen werden kann.

Das *Vector Space Model (VSM)* ist in diesem Zusammenhang eine gängige Form der Repräsentation. Im VSM betrachtet man jedes Item als Vektor im N -dimensionalen Raum, in dem jede Dimension ein Attribut aus einer Menge von mehreren Items darstellt, sowie N der Gesamtzahl an Begriffen in der Menge entspricht. Das VSM findet z.B. für die Empfehlungen von Internetseiten [93, 118] oder Nachrichten [4, 11, 88], deren Attribute aus Textbausteinen extrahiert wurden, Verwendung. Formal sei $I = \{i_1, i_2, \dots, i_{|I|}\}$ die Menge an verfügbaren Items und $T = \{t_1, t_2, \dots, t_N\}$ die Menge an Attributen, die aus allen Items der Menge I extrahiert sind. Jedes Dokument $i_j \in I$ wird repräsentiert als N -dimensionaler Vektor mit $i_j = \{w_{1j}, w_{2j}, \dots, w_{Nj}\}$, wobei w_{kj} dem Gewicht des Attributes t_k in Item i_j entspricht [2]. Zur Errechnung des Gewichts w_{kj} eines Attributes t_k in Item i_j ist *Term Frequency – Inverse*

²² Oder auch *Feature oder Keyword*

²³ z.B. durch bereits getätigte Bewertungen

*Document Frequency (TF-IDF)*²⁴ eine gängige Methode. Die Annahme des Verfahrens besteht darin, dass ein Attribut, welches in einem einzelnen Item häufig (TF), in den übrigen Items allerdings weniger häufig vorkommt (IDF), von größerer Bedeutung ist. Eine ausführliche Betrachtung des Verfahrens ist z.B. aus [8] S. 29 ff. zu entnehmen. Für die Errechnung einer Nützlichkeit bedarf es für den User einer äquivalenten Vektordarstellung. Dies geschieht in Form eines Userprofils $UP(u)$ für einen User $u \in U$, welches ebenfalls als Vektor von Attributgewichten aufgefasst wird und in dem jedes Gewicht $w_{u,k}$ die Bedeutung eines Attributes t_k für einen User u dargestellt [2]:

$$UP(u) = (w_{u1}, \dots, w_{uN}) \quad (2)$$

Die Nützlichkeit $R(i, u)$ eines Items i für einen User u kann dann über die Errechnung eines Ähnlichkeitsmaßes erfolgen. In der Literatur findet sich eine Vielzahl an Ähnlichkeitsmaßen, die an dieser Stelle Verwendung finden können, so z.B. der Dice- oder Jaccard-Koeffizient (vgl. [82] S. 73 bzw. S. 74). Das am häufigsten vorgeschlagene Ähnlichkeitsmaß ist der Kosinus-Koeffizient, welcher für zwei Gewichtsvektoren \vec{w}_u und \vec{w}_i ermittelt wird durch

$$R(i, u) = \cos(\vec{w}_u, \vec{w}_i) = \frac{\vec{w}_u * \vec{w}_i}{\|\vec{w}_u\|^2 * \|\vec{w}_i\|^2} = \frac{\sum_{j=1}^N w_{j,u} * w_{j,i}}{\sqrt{\sum_{j=1}^N w_{j,u}^2} * \sqrt{\sum_{j=1}^N w_{j,i}^2}} \quad (3)$$

Im Fall des Kosinus-Koeffizienten errechnet die Funktion $R(i, u)$ den eingeschlossenen Winkel zwischen den Vektoren \vec{w}_u und \vec{w}_i und ermittelt einen Wert im Wertebereich $[-1, 1]$. Ein Wert von 1 weist dabei auf eine hohe Ähnlichkeit bzw. ein Wert von -1 auf eine hohe Unähnlichkeit hin. Der ermittelte Wert kann ebenso das Maß an Nützlichkeit repräsentieren, sofern man davon ausgeht, dass ähnliche Items einen Nutzen für den User ergeben. In vielen Empfehlungsdienstanwendungen ist vor allem die Empfehlung der nützlichsten Items von Interesse und somit solche, die eine hohe errechnete Wertigkeit für $R(i, u)$ besitzen. Eine gängige Variante ist die Betrachtung einer nach Wertigkeit sortierten Liste, wobei der

²⁴ Ursprünglich vor allem für die Gewichtung von Begriffen in Textdokumenten genutzt, was dem Bereich des Information Retrieval entstammt [8] S. 27 ff.

Empfehlungsdienst die N -nützlichsten Items aus der Liste präsentiert. Man spricht in diesem Zusammenhang auch vom *Top-N* Prinzip [32].

Alternativ zu den beschriebenen heuristischen Methoden sind *Modell-basierte* Ansätze gebräuchlich, die auf Grundlage einer zur Verfügung stehenden Datenbasis ein Modell ableiten. Es wird in diesem Zusammenhang versucht eine Klassenzuordnungen (z.B. interessant, nicht interessant) für Items herzuleiten, so dass das Vorgehen einem Klassifikationsproblem entspricht. Der Ansatz eines *Naive Bayes Klassifikators* aus dem Bereich des maschinellen Lernens findet zur Lösung des Klassifikationsproblems häufig Verwendung. Die Methode ist hinsichtlich der Klassifizierung von Textdokumenten z.B. zur Klassifizierung von Spam Mails, ein gängiges Verfahren [104], dient aber ebenso in der Anwendungsdomäne Empfehlungsdienste als probate Methode zur Empfehlung für jegliche Form von Items, die durch Attribute beschreibbar sind z.B. bei Büchern [108] oder Filmen/Fernsehinhalten [41, 42]. Naive Bayes Klassifikatoren errechnen die *a posteriori Wahrscheinlichkeit*, dass ein bestimmtes Item einer bestimmten Klasse zugeordnet werden kann. So bestimmt ein Naive Bayes Klassifikator unter Verwendung des Bayes Theorems die bedingte Wahrscheinlichkeit $P(c|i)$, dass ein Item $i \in I$ einer Klasse $c \in C$ zugeordnet werden kann. Genauer ermittelt der Klassifikator die Wahrscheinlichkeit des Auftretens von Itemattributen innerhalb einer Klasse mit $P(c|t_{1,j} \& \dots \& t_{N,j})$ ²⁵, wodurch sich die Wahrscheinlichkeit für ein übergeordnetes Item $i_j \in I$ ergibt. Für eine ausführliche Darstellung sei an dieser Stelle auf die Literatur verwiesen (z.B. [129] S. 91). Die Wahrscheinlichkeit $P(c|f_{1,k} \& \dots \& f_{N,k})$ wird für jede Klasse $c_j \in C$ errechnet und das jeweilige Item der Klasse mit der höchsten Wahrscheinlichkeit zugeordnet:

$$\operatorname{argmax}_{c_j} P(c_j) \prod_x P(f_{x,k} | c_j). \quad (4)$$

Als eine Weiterentwicklung des Naive Bayes Klassifikators kann das *Bayes'sche Netz* gesehen werden, in dem die Abhängigkeiten zwischen Attributen als bedingte Wahrscheinlichkeiten in einem gerichteten azyklischen Graphen dargestellt sind (vgl. [82] S. 80). Die Verbreitung in inhaltsbasierten Verfahren ist allerdings gering. Verbreiteter zur Klassifizierung von Items ist die

²⁵ Wobei N die Gesamtanzahl an Attributen darstellt.

Nutzung von Entscheidungsbäumen, die auf einer Baumstruktur basieren. In einem Entscheidungsbaum repräsentieren die Blätter eine Klasse, die inneren Knoten ein Attribut und die Kanten ein Gewicht vom Vorgängerknoten, welches binär, numerisch oder auch textuell sein kann. Ein zu klassifizierendes Item durchläuft den Baum und wird an jedem Knoten zu einem Attribut geprüft. Je nach Attributwert des Items wird der nächste Knoten geprüft oder im Fall eines Blattes ist das Item final klassifiziert. Die Herausforderung eines Entscheidungsbaumes liegt im Aufbau des Baumes, der in der Regel durch eine Menge von Trainingsdatensätzen realisiert wird. Der automatisierte Aufbau oder auch die *Induktion* von Entscheidungsbäumen erfolgt in der Regel in rekursiver Form nach dem Top-Down-Prinzip. Gängige Verfahren zum automatisierten Aufbau von Entscheidungsbäumen sind *Classification and Regression Trees (CART)*, der *Iterative Dichotomiser Version 3 (ID3)*- oder der *C4.5*-Algorithmus (vgl. z.B. [142] S. 142 ff.). Entscheidungsbäume sind ähnlich wie Bayes Klassifikatoren zur Klassifikation von Dokumenten z.B. auf Textbasis ein wertvolles Instrument, aber ebenso auch in Empfehlungsdienstszenarien zur Empfehlung von Filmen oder im Kontext von Ontologien anwendbar [16, 92].

Für inhaltsbasierte Verfahren sind verschiedene Probleme bekannt, die grundlegend übergeordnet betrachtet werden können. Die Kernprobleme können in einer *geringen Inhaltsanalyse*, *Überspezialisierung* und dem „*New User*“-Problem zusammengefasst werden [2, 9]. Die Verwendung von Attributen zur Inhaltsanalyse von Items ist dabei das wohl schwerwiegendste Problem. Zum einen bedarf es einer genügenden Beschreibung von Items, die in der Anzahl und Qualität von Attributen mündet. Das Auffinden geeigneter Attribute ist dabei ein nicht immer trivialer Prozess. Wie bereits erläutert können Attribute manuell beschrieben oder aus einem Textsegment extrahiert werden. Der Vorgang Attribute manuell zu deklarieren ist aufwändig und eher statisch, da es einer dauerhaften manuellen Pflege der Attribute bedarf. Die Extraktion von Attributen, wie zuvor dargestellt z.B. mit dem TF-IDF Ansatz, ist für Textdokumente oder Textbausteine ein probates Mittel und ist in heutigen Verfahren immer noch ein fester Bestandteil. Allerdings bleibt z.B. Bild- und Videomaterial bei diesem Ansatz unberücksichtigt. Des Weiteren bedarf es bei der Extraktion aus Textdokumenten Hürden zu umgehen, die nicht

direkt auf technischer Ebene entstehen. So müssen z.B. Spracheigenschaften wie Polysemie²⁶ und Synonymie²⁷ mit berücksichtigt werden (vgl. [82] S. 32 ff.). Des Weiteren können inhaltsbasierte Verfahren i.d.R. keine Aussage über die Qualität eines Items geben. Vergleicht ein inhaltsbasierter Ansatz z.B. zwei Kinofilme miteinander und errechnet auf Grundlage von Schauspielern, Genre und Regisseur eine hohe Gleichheit, so beinhaltet dies noch keine Aussage darüber, ob der Film tatsächlich gut ist, ob er spannend, komisch ist, ob er eingängig oder schwer verständlich ist.

Der Begriff der Überspezialisierung thematisiert das Problem Empfehlungen auf Grundlage eines Userprofils abzuleiten, welches lediglich aus Präferenzen des Users selbst besteht. Der User bewertet oder kauft i.d.R. Items, die ihr oder ihm bekannt sind. Das Userprofil füllt sich dadurch mit Items, die für den User ohnehin eine gewisse Relevanz besitzen, so dass alternative Items, die für sie oder ihn ebenfalls einen Nutzen haben, unberücksichtigt bleiben. Über die Thematik des Userprofils ergibt sich auch das „New User“-Problem²⁸. Über einen neuen Benutzer im System sind zum Startzeitpunkt keine Informationen bekannt. Ohne oder auch mit nur wenigen bekannten Präferenzen kann ein inhaltsbasiertes System keine genügenden Empfehlungen ableiten. Einige Verfahren umgehen dieses Problem durch die Nutzung eines initialen Profils, welches Präferenzen vom User zum Startzeitpunkt erfragt (vgl. z.B. [42]).

3.2.2 Kollaborative Verfahren

Im Vergleich zu inhaltsbasierten Verfahren bestimmen kollaborative Verfahren eine Nützlichkeit eines Items für einen User nicht auf Grundlage von Itemattributen, sondern vielmehr auf Aussagen bzw. Bewertungen anderer User. Die Grundannahme besteht darin, dass nützliche Items eines Users ggf. auch für weitere User nützlich sein können, sofern eine Ähnlichkeit zwischen den Usern besteht. Formal beschreibt sich der Ansatz wie folgt:

²⁶ Unter Polysemie versteht man Begriffe, die mehrdeutige Inhalte beschreiben

²⁷ Unter Synonymie versteht man Begriffe, die den gleichen oder einen ähnliche Bedeutungsumfang besitzen

²⁸ Wird häufig auch als *Kaltstartproblem* bezeichnet.

Die Nützlichkeit $R(i, u)$ eines Items $i \in I$ für einen User $u \in U$ wird über die Vielzahl an Nützlichkeiten $R(i, u_j)$ errechnet, die für die User $u_j \in U$ existieren, die *ähnlich* zum User u sind [2].

Im Kontext kollaborativer Verfahren bedeutet der Begriff der Ähnlichkeit somit das Auffinden ähnlicher Geschmäcker oder ähnlicher Interessen und weniger der Vergleich statischer Attribute wie in inhaltsbasierten Verfahren. Verwendet man wiederum das Beispiel einer Filmdatenbank, so würden in einem kollaborativen Empfehlungsdienst z.B. die Bewertungen von Filmen verglichen und Empfehlungen aus einer Usergruppe mit gleichen Bewertungen abgeleitet. Diese Grundannahme ist in unterschiedlichen Formen adaptiert worden und so zeigt Abschnitt 3.2.2.1, dass eine Bewertung der Nützlichkeit für unberücksichtigte Items ebenso über den Vergleich ähnlicher Items resultieren kann. Die Herleitung von Empfehlungen auf Grundlage ähnlicher User- bzw. Item-Gruppen bezeichnet man auch als *memory-basierte kollaborative Verfahren*²⁹. Des Weiteren zeigt Abschnitt 3.2.2.2 eine Möglichkeit zur Herleitung von Empfehlungen, die auf Grundlage eines angelernten Modells basieren. In diesem Zusammenhang spricht man auch von *model-basierten kollaborativen Verfahren*. Die Untergliederung nach memory- und model-basierten kollaborativen Verfahren folgt einem gängigen Standard in der Literatur [2, 33].

3.2.2.1 Memory-basierte kollaborative Verfahrensarten

Ähnlich zu den heuristischen Ansätzen der inhaltsbasierten Verfahren versuchen memory-basierte kollaborative Verfahren auf Grundlage einer User-Item Beziehung die vom User unberücksichtigten Items zu bewerten. Dies kann entweder ausgehend von einem Vergleich der Usergruppe oder ausgehend von einem Vergleich der Itemgruppe erfolgen. Man unterscheidet hierbei zwischen *User-basiertem* und *Item-basiertem kollaborativen Filtern*³⁰ [33, 82].

Beim User-basierten kollaborativen Filtern wird die Bewertung $R(i, u)$ für einen User $u \in U$ und für ein Item $i \in I$ auf Grundlage der Bewertungen für das Item i der k ähnlichsten User zu u errechnet. In diesem Zusammenhang spricht man auch von der *Nearest-Neighbour* oder auch k -

²⁹ z.T. werden für memory-based auch die Begriffe neighborhood- oder heuristic-based synonym verwendet.

³⁰ Der englische Begriff *Collaborative Filtering* ist ebenso gebräuchlich.

Nearest-Neighbour (kNN) Methode. Das Attribut k kann als eine Art Schwellwert gesehen werden und ist abhängig vom jeweiligen Anwendungskontext auszuwählen. Grundsätzlich ist es sinnvoll eine Schwelle festzulegen, da die Errechnung über alle Nutzer bei einer genügend großen Menge an Usern zu Lasten der Leistung des Systems und der Qualität der Empfehlungen geht [33]. Formal lässt sich das Vorgehen wie folgt beschreiben:

Sei \widehat{U} die Menge der k ähnlichsten User, wobei k zwischen 1 und der maximalen Anzahl an Usern der Menge U liegen kann, für die eine Bewertung für Item $i \in I$ existiert. $r_{u,i}$ kann im einfachsten Fall als aggregiertes Mittel über alle Bewertungen der Menge \widehat{U} ermittelt werden mit

$$r_{u,i} = \frac{1}{k} \sum_{u' \in \widehat{U}} r_{u',i} \quad (5)$$

Denkbar ist auch die Errechnung einer gewichteten Summe wie in Gleichung (6) oder einer normalisierten Summe wie in Gleichung (7), so dass die höher Ähnlichkeit zwischen Usern bzw. die unterschiedliche Interpretation von Bewertungsskalen mit berücksichtigt werden.

$$r_{u,i} = \frac{\sum_{u' \in \widehat{U}} sim(u, u') * r_{u',i}}{\sum_{u' \in \widehat{U}} sim(u, u')}, \quad (6)$$

$$r_{u,i} = \overline{r_u} + \frac{\sum_{u' \in \widehat{U}} sim(u, u') * (r_{u',i} - \overline{r_{u'}})}{\sum_{u' \in \widehat{U}} sim(u, u')} \quad (7)$$

mit $\overline{r_u} = (1/|I_u|) \sum_{i \in I_u} r_{u,i}$, wobei $I_u = \{i \in I | r_{u,i} \neq \emptyset\}$ gilt.

Zur Umsetzung der Funktion $sim(x, y)$ gibt es verschiedene Ansätze. Der in Kapitel 3.2.1 beschriebene Kosinuskoeffizient kann auch in den Gleichungen (6) und (7) Verwendung finden. Hierzu wird ein User u als Vektor $w_u \in \mathbb{R}^{|I|}$ betrachtet, wobei $w_{u,i} = r_{u,i}$, sofern eine Bewertung für Item i existiert, sonst 0. Die Gleichheit zwischen zwei Usern u und v wird dann errechnet über

$$sim(u, v) = \cos(\vec{w}_u, \vec{w}_v) = \frac{\sum_{i \in I_{uv}} r_{u,i} * r_{v,i}}{\sqrt{\sum_{i \in I_u} r_{u,i}^2} \sqrt{\sum_{j \in I_v} r_{v,j}^2}} \quad (8)$$

wobei $I_{u,v}$ die Menge an Items darstellt, für die eine Bewertung durch User u und v existiert. Eine häufig verwendete Alternative in kollaborativen Verfahren ist der Pearson Korrelationskoeffizienten³¹, der die Varianz in Bewertungen mit berücksichtigt [2].

Die Methode für das Item-basierte kollaborative Filterverfahren ist grundsätzlich analog zum User-basierten kollaborativen Filtern durchzuführen. Das Verfahren geht zur Errechnung einer Bewertung $r_{u,i}$ allerdings nicht von Bewertungen ähnlicher User, sondern von Bewertungen ähnlicher Items aus. Formal sei \hat{I} die Menge an Items mit der größten Ähnlichkeit zu Item i , welches durch User u bewertet wurde. Die Vorhersage der Bewertung $r_{u,i}$ kann z.B. wiederum als gewichtete Summe errechnet werden

$$r_{u,i} = \frac{\sum_{i' \in \hat{I}} sim(i, i') * r_{u,i'}}{\sum_{i' \in \hat{I}} sim(i, i')} \quad (9)$$

Die Errechnung kann ebenfalls analog zu den Gleichungen (5) und (7) erfolgen.

Ob bei der Entwicklung ein User- oder Item-basierter Ansatz gewählt wird, ist abhängig von der zugrundeliegenden Datenbasis. Eine Datenbasis mit einer hohen Anzahl an Items und einer geringeren Anzahl an Usern profitiert i.d.R. mehr von einem User-basierten Ansatz, da eine Vielzahl an nicht bewerteten Items hoch erscheint und somit in einer verminderten Vorhersagekraft münden würde [55]. Bei einer hohen Userbasis mit einer geringeren Anzahl an Items ist somit ein Item-basierter Ansatz plausibel. Des Weiteren muss betrachtet werden, in welcher Gruppe – Items oder User – sich häufiger Veränderungen in Datensätzen in Form von Neubewertungen ergeben. Theoretisch müsste für jede z.B. neue Bewertung eines Users für ein Item eine Neuberechnung aller davon abhängigen Bewertungen erfolgen. Zwar gibt es verschiedene Ansätze die Effizienz der Algorithmen auf großen Datenmengen durch

³¹ auch Person Correlation Score

Bestimmung von Standardwerten, das Auslassen unbedeutender Datensätze oder das Vorfiltern von Usergruppen [33], [17] zu steigern, dennoch ist jede Berechnung – z.B. bei Amazon mit über 300 Millionen aktiven Benutzern³² – zeit- und kostenintensiv. Ein deutlich schwieriger abzuwägendes Problem beschreibt die Literatur als *Zufälligkeit*³³ [33]. Dieses Problem tritt vor allem bei User-basierten Vergleichen auf, bei denen es aufgrund einer zu geringen Datenbasis zu unerwünschten bzw. *zufälligen* Empfehlungen kommen kann. Der Grund hierfür ist, dass bei einer geringen Userdatenbasis eine Ähnlichkeit bereits auf wenigen übereinstimmenden Bewertungen erfolgt. So kann es passieren, dass Empfehlungen von zwei sehr unterschiedlichen Usern abgeleitet werden und diese dem jeweiligen Geschmack nicht entsprechen. Für den jeweiligen Anwendungskontext ist abzuwägen, wie schwer eine falsche Empfehlung wiegt und ob diese aufgrund anderer Vorteile des Ansatzes akzeptiert werden kann.

Die Obergruppe der Memory-basierten Verfahren zeichnet sich vor allem durch die Einfachheit der Implementation, sowie der Effizienz auch auf sehr großen Datenmengen aus [33]. Dabei versprechen Item-basierte Ansätze im Vergleich zu User-basierten kollaborativen Filterverfahren qualitativ bessere Empfehlungen, sofern sich die zur Verfügung stehende Datenbasis für einen Item-basierten Ansatz eignet [136]. Ein eingangs bereits beschriebener Ansatz ist der Empfehlungsdienst von Amazon, der einen Item-basierten kollaborativen Ansatz verwendet [94, 105]. Das GroupLens Projekt zur Empfehlung von Newsgroup Inhalten ist ein frühes Beispiel für einen User-basierten kollaborativen Ansatz und stellte eine Alternative zu den sonst vor allem bei der Empfehlung von Textstrukturen üblichen inhaltsbasierten Verfahren dar [84, 127]. Weitere Ansätze, für die genügende Forschungsergebnisse existieren folgten z.B. für die Empfehlung von Musik [143], Website Informationen [157] oder Videos [62]. Ein umfassender Überblick von kollaborativen Ansätzen in allgemeiner Form, aber auch aus Sicht verschiedener Anwendungskontexte kann der Literatur entnommen werden [2, 153].

³² Abgerufen am 15. April 2016 von <http://www.amazon.com>

³³ auch Serendipity

3.2.2.2 Model-basierte kollaborative Verfahrensarten

Anders als Memory-basierten Verfahrensarten, wo Empfehlungen auf Grundlage von eher statistischen Ähnlichkeiten abgeleitet werden, versuchen Model-basierte kollaborative Verfahrensarten ein Modell auf Grundlage der gegebenen Datenbasis anzulernen. Das Modell dient zur Errechnung von Vorhersagen. Frühe Ansätze in diesem Bereich erweitern Verfahren aus dem Bereich des maschinellen Lernens, wo z.B. Klassifikatoren und Bayesche- oder Neuronale Netze als Model-basierte kollaborative Verfahren untersucht wurden [12, 17].

Latent-Factor-Models haben sich im Bereich der Model-basierten Verfahrensarten, bedingt durch eine sehr hohe Genauigkeit Empfehlungen abzuleiten, zu einem quasi Standard entwickelt [85]. Die grundlegende Idee für Latent-Factor-Models ist die Charakterisierung von Usern und Items durch die automatisierte Ableitung von *Einflussfaktoren*, die sich durch das explizite/implizite Feedback von Usern ergibt. In einem Empfehlungsdienst für Filme könnte es z.B. offensichtliche Faktoren wie der Vergleich eines Genres geben, allerdings auch weniger offensichtliche Faktoren wie inhaltlicher Tiefgang eines Films und ebenso nicht zu definierende Faktoren. Ein Latent-Factor-Model kann somit eine Aussage darüber treffen, warum ein User ein Item gut oder schlecht bewertet hat. Die Methode der *Matrix-Faktorisierung* ist dabei die vielversprechendste Form der Umsetzung eines Latent-Factor-Models. Formal kann die Methode der Matrix-Faktorisierung wie folgt dargelegt werden:

Sei U eine Menge an Usern, sowie I eine Menge an Items, so ergibt sich zunächst eine User-Item Bewertungsmatrix R , die der Dimension $|U| \times |I|$ entspricht mit $R_{u,i} = r_{u,i}$, sofern eine Bewertung zwischen User u und Item i existiert, sonst 0. Durch Zerlegung der Bewertungsmatrix R sollen die zuvor angesprochenen Faktoren abgeleitet und quantifiziert werden. Dies erfolgt durch die Zerlegung von R in zwei Matrizen Q und P , die eine jeweilige Dimension bei k zu schätzenden Faktoren mit $|U| \times k$ bzw. $|I| \times k$ annehmen. Q bestimmt dabei das Maß der Ausprägung der Faktoren für die Menge der User bzw. P das Maß der Ausprägung der Faktoren für die Menge der Items. Ziel ist die Annäherung von P, Q durch $P \times Q^T$

$$\widehat{R} = P \times Q^T \approx R \quad (10)$$

bzw.

$$\hat{r}_{u,i} = \sum_{j=1}^k p_{u,j} q_{i,j} \approx r_{ui} \quad (11)$$

wo $p_{u,j}$ die Elemente von P und $q_{i,j}$ die Elemente von Q darstellen. $\hat{r}_{u,i}$ stellt final die geschätzte Bewertung eines Users u zu einem Item i dar [171]. Zur Schätzung von P und Q werden beide Matrizen z.B. mit Zufallswerten gefüllt und der Fehler zwischen vorhandenen Bewertungen, sowie den geschätzten Wertungen minimiert. Gradientenverfahren sind an dieser Stelle eine häufig angewendete Methode zur Minimierung [85, 171]. Verschiedene Alternativen zur Anlernung eines Modells, die sich z.B. zur Nutzung gering gefüllter Matrizen oder zur parallelen Berechnung eignen, sind in der Literatur zu finden [85].

Im Vergleich zu Memory-basierten Verfahren erreichen Ansätze auf Basis eines Latent-Factor-Models bedingt durch die Abbildung verschiedener Einflussfaktoren und Zusammenhänge eine zumeist verbesserte Vorhersagekraft und somit besser Qualität bei der Herleitung von Empfehlungen. Sowohl empirische Studien als auch der Erfolg eines Latent-Factor-Model Ansatzes im Netflix Prize Wettbewerb belegen dies (vgl. [129] S. 151 ff.). Der Aufbau eines etwaigen Modells ist allerdings zeitintensiv und nicht jede Implementation unterstützt eine inkrementelle Erweiterung des Modells. Wo kollaborative Filter-Ansätze eine quasi direkte Neuberechnung von Empfehlungen erlauben, bedarf es in einigen Implementationen von Latent-Factor-Model Ansätzen der Neuberechnung des Modells, was gegen die Nutzung sprechen kann.

3.2.2.3 Nachteile kollaborativer Verfahren

Die Nachteile, die in Abschnitt 3.2.1 für inhaltsbasierte Verfahren beschrieben wurden, können durch kollaborative Verfahren in der Regel aufgewogen werden. So ist es in kollaborativen Verfahren z.B. möglich, jegliche Art von Items inhaltlich zu bewerten. Auch die Form der Überspezialisierung tritt in kollaborativen Verfahren nicht auf. Dennoch ergeben sich auch für kollaborative Verfahren verschiedene Nachteile. Das bereits bei inhaltsbasierten Verfahren diskutierte „New User“-Problem bzw. Kaltstartproblem ist auch in kollaborativen Verfahren von Bedeutung. Über einen neuen User im System sind zumeist keine Informationen – wie z.B. Bewertungen von Items – bekannt. Ein Vergleich zu weiteren Benutzern und somit der

Eingruppierung in eine Usergruppe ist nicht möglich, was den Anstoß eines sinnvollen Empfehlungsprozesses ausschließt. Gleiches gilt analog für neue Items im System. Das „New User“-Problem existiert in gleicher Form auch für Model-basierte Verfahren wie z.B. das Latent-Factor-Model. Verschiedene Verfahren haben sich zur Lösung dieses Problem etabliert. Die bereits in Abschnitt 3.2.1 diskutierte Lösung zur Erstellung eines initialen Benutzerprofils könnte dem Problem entgegen wirken. Ebenfalls möglich ist die Bereitstellung von Empfehlungen auf Grundlage systemweiter Informationen, wie z.B. welche Items sind über die gesamte User Gruppe am beliebtesten [124].

Ein weiteres Problem ist die geringe Anzahl an Bewertungen eines Users im Vergleich zu den Bewertungen, die durch das System zu errechnen sind. Das Problem wird häufig auch unter dem Begriff *Sparsity* beschrieben und tritt vor allem dann auf, wenn das Verhältnis von Items zu Usern deutlich größer ist. In der Regel gibt es nur für eine geringe Stückzahl an Items eine Bewertung durch einen User, so dass die User-Item Matrix sehr große Datenlücken enthält. Möglichkeiten diesem Problem entgegenzuwirken ist die Errechnung einer Ähnlichkeit von Usern nicht ausschließlich auf der User-Item Beziehung, sondern auf Basis User abhängiger Faktoren wie z.B. Geschlecht, Alter, demografische Informationen oder Bildungsstand [119]. Auch die Nutzung von Kontextinformationen³⁴ wie z.B. Zeit oder Ort kann diesem Problem entgegenwirken [1]. Die bereits in Abschnitt 3.2.2.2 beschriebenen Verfahren der Matrix-Faktorisierung dienen in diesem Zusammenhang als probates Mittel zur Schließung der Datenlücken [2, 137].

3.2.3 Hybride Verfahrensarten

Die Abschnitte 3.2.1 und 3.2.2 präsentierten mit inhaltsbasierten- und kollaborativen Verfahren grundlegende Verfahrensweisen für Empfehlungsdienste. Es hat sich gezeigt, dass beide Verfahrensarten unterschiedliche Vor- und Nachteile aufweisen. Abhängig vom jeweiligen Einsatzszenario wiegen bestimmte Nachteile weniger schwer und können in einigen Fällen somit vernachlässigt werden. In der Regel versucht man allerdings Nachteile der jeweils gewählten Verfahrensart aufzuwiegen, so dass eine größtmögliche Qualität für Empfehlungen entsteht. Vor

³⁴ In diesem Zusammenhang spricht man auch vom Forschungsfeld *Context-Aware Recommender Systems*

allem fehlende Daten führen häufig zu entscheidenden Problemen (z.B. Kaltstartproblem), die es zu überbrücken gilt. In diesem Zusammenhang hat es sich etabliert hybride Verfahren einzusetzen, die die Kombination aus inhaltsbasierten und kollaborativen Verfahren darstellen [2, 22]. Nach [2] ergeben sich vier Möglichkeiten Verfahren zu einem hybriden System zu kombinieren:

- *Combining Separate Recommenders*: Inhaltsbasierte und kollaborative Verfahren separat implementieren und deren Empfehlungen kombinieren z.B. durch Aggregation der errechneten Wertigkeiten [119], [29] oder durch Auswahl der Empfehlungen eines Verfahrens, die gemäß einer Strategiegrundlage sinnvollere Empfehlungen erschlossen haben [160], [13].
- *Adding Content-Based Characteristics*: Inhaltsbasierte Methoden in ein kollaboratives Verfahren integrieren z.B. durch die Herleitung von Empfehlungen auf Grundlage inhaltsbasierter Methoden, sofern keine Informationen für das kollaborativen Verfahren (Bewertungen) zur Verfügung stehen [10].
- *Adding Collaborative Characteristics*: Kollaborative Methoden in ein inhaltsbasiertes Verfahren integrieren z.B. durch die Herleitung von Empfehlungen auf Grundlage kollaborativer Methoden, sofern keine Informationen für das inhaltsbasierte Verfahren (Itemattribute) zur Verfügung stehen [152].
- *Unifying Model*: Ein für beide Verfahrensarten gebräuchliches Modell entwickeln z.B. durch Kombination von Bewertungen und Itemattributen [96], unter Verwendung probabilistischen oder statistischen Methoden [7, 63] oder der Nutzung übergeordneter, domänenspezifischer Informationen, die unabhängig von einer inhaltsbasierten oder kollaborativen Methoden verwertbar sind [23].

Eine pauschale Aussage, welche hybride Variante in welchem Anwendungsfall zu verwenden ist, kann nicht getroffen werden. Auch bei hybriden Verfahren gilt es zu betrachten, welche Vorteile für ein jeweiliges Anwendungsszenario inklusive der zur Verfügung stehenden Datenbasis erwirkt werden können. Es gibt dennoch einige Vorüberlegungen, die bei der Integration zu berücksichtigen sind. Die Kombination zweier Verfahrensarten ist recht unproblematisch, da die jeweiligen Verfahrensarten unabhängig voneinander Ergebnisse generieren und diese abschließend aggregieren. Erweitert man allerdings eine Verfahrensart

durch eine zweite, so ist eine direkte Abhängigkeit zwischen den Verfahren gegeben. So basiert z.B. die Ermittlung von Empfehlungen eines inhaltsbasierten Verfahrens zunächst auf den Ergebnissen eines kollaborativen Ansatzes. Die Qualität variiert ggf. je nachdem in welcher Reihenfolge die Verfahren angestoßen werden. Die Literatur gibt an dieser Stelle eine recht genaue Vorstellung davon, welche Kombinationen in welchem Zusammenhang sinnvoll erscheinen und welche ggf. nicht (vgl. hierzu [22]). Ohnehin bietet die Literatur an dieser Stelle eine Vielzahl an bereits existierenden Lösungsvorschlägen, die bei der Integration eines etwaigen Ansatzes wertvoll sind.

Es ist noch zu klären, ob der Einsatz von hybriden Systemen im Vergleich zu einzeln verwendeter Verfahrensarten einen Mehrwert für die Generierung von Empfehlungen birgt. Die vorherigen Abschnitte haben ersichtlich gemacht, dass hybride Verfahren grundlegend von einer veränderten Datenbasis ausgehen. Je nach Methode wird die Datenbasis somit vergrößert oder kombiniert, so dass auf den ersten Blick von einem vergrößerten Informationsgehalt ausgegangen werden kann. Durch diese Annahme kann die These vertreten werden, dass sich ein größerer Informationsgehalt auch auf eine verbesserte Qualität an Empfehlungen bzw. besser geschätzten Bewertungen auswirkt. Verschiedene Studien in diesem Bereich bestätigen diese These (vgl. z.B. [22]), weshalb die Implementation eines hybriden Verfahrens generell sinnvoll erscheint. Dennoch bedarf es vorab einer genauen Prüfung der vorliegenden Datenbasis, welche und ob eine hybride Strategie zweckdienlich ist.

3.2.4 Erweiterte Analyseverfahren: Clustering

Unter Clustering³⁵ versteht man die Unterteilung einer Datenmenge in Gruppen oder auch die gruppenweise Zuordnung von Datensätzen. Die Idee des Clustering ist, „natürliche Gruppen“ in Datenmengen zu finden, und gehört zu den Verfahrensarten des unüberwachten Lernens (vgl. [129] S. 61). Clustering dient im Kontext von Empfehlungsdiensten zur Datenaufbereitung. So bewirkt eine Datenaufbereitung durch Clustering sowohl in inhaltsbasierten als auch in kollaborativen Verfahrensarten die Erhöhung der Skalierbarkeit von Empfehlungsdienstsystemen. Bedingt durch zumeist große zu verarbeitende Datenmengen sind

³⁵ Im deutschen häufig auch als Clusteranalyse bezeichnet.

eine Vielzahl an Vergleichsoperationen zwischen Items oder Usern notwendig, um die ähnlichsten Items bzw. User bestimmen zu können. Mittels Clustering ist eine deutlich vereinfachte Voreinteilung von Daten in sogenannte Cluster möglich, was zu einer kostengünstigen und effektiven Errechnung von Empfehlungen auch auf großen Datenmengen führt. Clustering stellt in vielen Anwendungskontexten eine optimierende Maßnahme für Empfehlungsdienste dar und wird in der Literatur häufig ergänzend zu den grundlegenden Verfahrensarten beschrieben.

Clusteringverfahren können grundlegend in *hierarchische* und *partitionierende* Clusteringverfahrensarten unterschieden werden. Bei partitionierenden Verfahren ist zunächst eine Vorbestimmung von Clustern (oder auch Gruppen) durchzuführen, in die es die Datensätze einzugliedern gilt. Jeder Cluster besitzt dabei ein Clusterzentrum, welches als Ausgangsposition zur Clusterung von Datensätzen gesehen werden kann. Der *k-Means Clustering Algorithmus* ist in diesem Zusammenhang eines der gängigsten Verfahren, welches einen Datensatz in Abhängigkeit von k -Clusterzentren in k disjunkte Teilmengen aufteilt (vgl. [129] S. 61 ff.):

Sei I die Menge an Items und k die Anzahl an vorgesehenen Clustern, dann ergeben sich k disjunkte Teilmengen S_1, \dots, S_k . Beschrieben ist jeder Cluster S_j durch seine Teilmenge an Items I_j und einem jeweiligen Clusterzentrum λ_j . Das Ziel des Verfahrens ist die Minimierung der Distanzen aller Items eines Clusters zum jeweiligen Clusterzentrum:

$$\min E = \sum_{i=1}^k \sum_{n \in S_i} d(x_n, \lambda_i), \quad (12)$$

wobei x_n ein Vektor mit Merkmal- oder Distanzeigenschaften eines n -ten Items darstellt sowie λ_j das Zentrum der Teilmenge S_j und d eine entsprechende Distanzfunktion – gängig ist z.B. die euklidische Distanz – zur Ermittlung des Abstandes zwischen Zentrum und Item repräsentiert. Der Algorithmus terminiert, sofern E bis zu einer vorbestimmten Grenze nicht weiter minimiert werden kann.

Die Auswahl der Clusterzentren erfolgt zu Beginn zufällig. In Abhängigkeit der kleinsten Distanz zum jeweiligen Clusterzentrum ergibt sich die Zuweisung der Items. Die Clusterzentren werden durchweg neu berechnet und ggf. eine Neuordnung von Items zu einem anderen

Cluster mit einer geringeren Distanz durchgeführt. Das Verfahren ist in Abbildung 14 vereinfacht dargestellt.

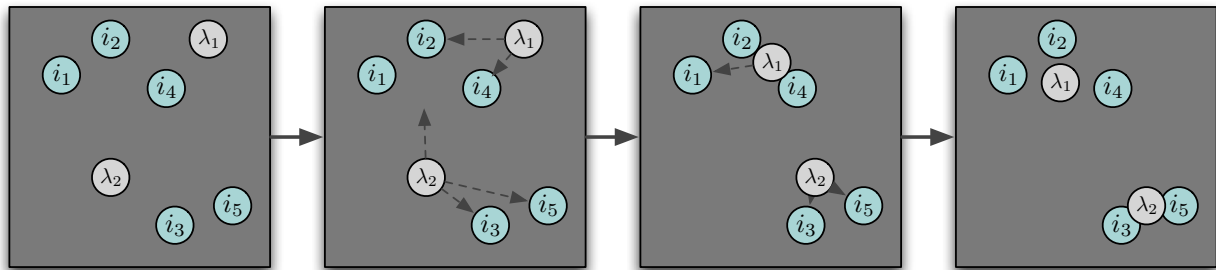


Abbildung 14 k-Means Clustering Algorithmus

Ein häufiger Nachteil bei partitionierenden Clusteringverfahren ist die Vorauswahl von (k -) Clustern, da die Anzahl von Clustern ggf. nicht oder nur schwer zu bestimmen ist. Es darf vorweggenommen werden, dass auch in der vorliegenden Arbeit die Vorauswahl einer bestimmten Anzahl von Clustern wenig sinnvoll ist. Je nach Anzahl an Tags variiert die Anzahl von relevanten inhaltlichen Bereichen möglicherweise, so dass ebenfalls eine Anzahl an Clustern variieren würde. Es ist somit ein Verfahren wünschenswert, welches eine sukzessive Annäherung einer Einteilung von Clustern vornimmt. Dies ist in Form von hierarchischen Clusteringverfahren realisierbar.

Bei hierarchischen Verfahren ist es das Ziel, Datensätze mit einer geringen Distanz zueinander zu einem Cluster zu vereinen. Hierarchische Verfahren können ihrerseits in *divisiven* und *agglomerativen* Clusteringverfahren unterschieden werden. Wo divisive Verfahren eine Top-Down-Strategie verfolgen und ausgehend von einer Gesamtmenge an Datensätzen als Gesamtcluster zur Teilung ausgehen, steht bei agglomerativen Verfahren eine Bottom-Up-Strategie im Vordergrund, in der jeder Datenpunkt einen eigenen Cluster darstellt, die sukzessive zusammengeführt werden. In agglomerativen Verfahrensarten findet somit eine Fusionierung von Datenpunkten zu einem Cluster statt und kann formal wie folgt betrachtet werden (vgl. [169]):

Sei K eine Menge an bereits geclusterten Items sowie I die Menge der Items bzw. Cluster, die noch zu clustern sind. Sei $D(K, I)$ eine Fusionierungsfunktion zur Bestimmung der jeweils einander nächsten Cluster der zwei Clustermengen D, K z.B. definiert durch

$$D(K, I) := \min_{k \in K, i \in I} \{d(k, i)\}, \quad (13)$$

wobei k, i Vektoren eines jeweiligen Items mit Merkmal- oder Distanzeigenschaften darstellen und d einer entsprechende Distanzfunktion³⁶ zur Bestimmung des Abstandes zwischen zwei Items der Clustermengen entspricht.

Die Fusionierungsvorschrift ermittelt den minimalen Abstand zwischen allen Elemententenpaaren beider Clustermengen. Die Variante aus Gleichung (13) folgt einem Ansatz, welcher als *Single-Linkage* bezeichnet wird. Weitere Varianten, in denen der maximale Abstand oder der durchschnittliche Abstand zwischen Elemententenpaaren beider Clustermengen ermittelt wird, sind in der Literatur beschrieben (vgl. z.B. [169]). Abbildung 15 illustriert in stark vereinfachter Form das Vorgehen bei einer (Single-Linkage) agglomerativen Strategie.

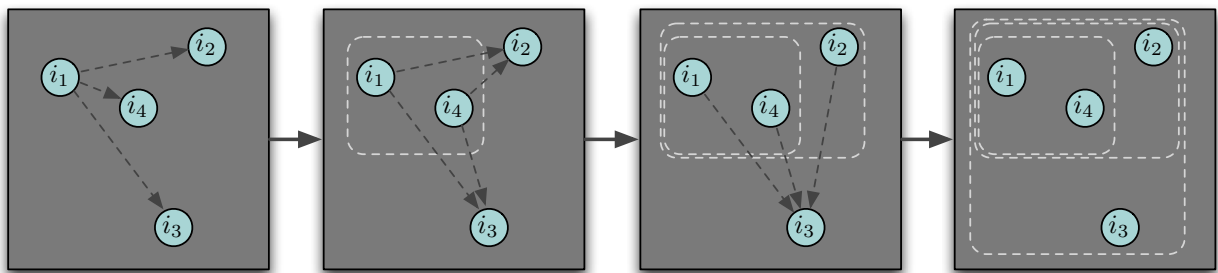


Abbildung 15 Agglomeratives (hierarchisches) Clustering

Bei der Umsetzung eines entsprechenden Ansatzes bedarf es der Auswahl eines geeigneten Distanz-/Ähnlichkeitsmaßes sowie einer Fusionierungsfunktion oder -vorschrift.

3.2.5 Evaluationsmethoden

Für die Evaluation von Empfehlungsdiensten haben sich grundlegend drei Evaluationsformen etabliert (vgl. [129] S. 261 ff.):

³⁶ Die Distanzfunktion kann wiederum verschiedenen Ausführungen folgen, so z.B. der euklidischen Distanz (vgl. Ausführungen zum k-Means-Clustering).

- Offline Experiment
- User Studies
- Online Evaluation, online Simulation, online Experiment (Real Life Testing)

Jede Evaluationsform erhebt dabei den Datenbestand für die Evaluation in unterschiedlicher Form, so dass sich jede Form unterschiedlich stark an eine reale Nutzung eines etwaigen Systems annähert. Ein *offline Experiment* greift in der Regel auf bestehende bzw. angesammelte (große) Datensätze zurück, so dass es sich vielmehr um die Simulation eines Empfehlungsdienstes handelt. Einige wenige genügend große und frei zugängliche Datenbestände existieren, so z.B. der Netflix Datenbestand für Filme³⁷ oder der Bibsonomy Datenbestand für Bookmarks bzw. Publikationen³⁸. Die Nutzung frei zugänglicher Datensätze ermöglicht einen aussagekräftigen Vergleich verschiedener Forschungsansätze. Die Entwicklung eigener Datenbestände kann hingegen zeitaufwändig sein. Für die Evaluation eines kollaborativen Ansatzes bedarf es so einer Vielzahl an Bewertungen für unterschiedliche Items. Für ein inhaltsbasiertes Verfahren bedarf es hingegen einer genügenden Item Datenbasis. Die Möglichkeit eine Evaluation ohne Akquise von Probanden durchzuführen ist ein entscheidender Vorteil eines offline Experiments. Sofern ein Vergleich zwischen Algorithmen auf Grundlage gleicher Datensätze beruht, birgt das Evaluationsergebnis zudem eine gute Vergleichsmöglichkeit zu weiteren Verfahren. Der entscheidende Nachteil für offline Experimente ist die eingeschränkte Vielfältigkeit der möglich zu prüfenden Aussagen. Offline Experimente eignen sich in der Regel für den Vergleich von algorithmischen Verfahren hinsichtlich ihrer Vorhersagequalität und Leistung.

In einer Vielzahl an Empfehlungsdienstszenarien kommt es allerdings auf die Interaktion zwischen User und System an. *User Studies* oder *Benutzerstudien* ermöglichen es in einem experimentellen Testszenario Aufgaben von Probanden innerhalb des Systems durchführen zu lassen und sowohl die Interaktion als auch die Reaktion von Probanden auf Ereignisse im System zu untersuchen. Auch die Genauigkeit von Ereignissen im System, so z.B. die Empfehlung eines Items, kann beim Probanden direkt erfragt werden. In einer Benutzerstudie

³⁷ <http://netflixprize.com>

³⁸ <http://www.kde.cs.uni-kassel.de/bibsonomy/dumps/>

geht es somit vor allem um die Untersuchung eines Benutzerverhaltens in Bezug auf das System. Die Möglichkeit ein Benutzerverhalten zu untersuchen ist einer der entscheidenden Vorteile innerhalb einer Benutzerstudie, da die Ergebnisse tatsächliche Aussage über den Einfluss eines Empfehlungsdienstes auf einen User ermöglichen, was in offline Experimenten nur hypothetisch möglich ist. Ein klarer Nachteil ist der organisatorische sowie zeitliche Aufwand der mit der Evaluationsform einhergeht. Für die Durchführung bedarf es neben der Akquise von Probanden auch der Teilnahme eines Experten, der die Evaluation beobachtet und ggf. protokolliert. Mit der Durchführung der Evaluation selbst geht ebenso ein gewisser Aufwand einher, da es für die Planung des Evaluationssettings einer genügenden Zeitanstrengung bedarf, um Fehlschläge zu vermeiden. Häufig wird deshalb eine Vorstudie durchgeführt, um zu prüfen, ob innerhalb des Settings alle benötigten Daten erhoben werden können.

Bei einer *online Evaluation*, *online Simulation* oder einem *online Experiment* ist eine Nutzung des Empfehlungsdienstes unter realen Bedingungen vorgesehen. Die Evaluationsform sieht somit die Nutzung des Systems mit realen Items in der gewohnten Umgebung des jeweiligen Probanden vor. Für eine Vielzahl an Empfehlungsdiensten ist die gewohnte Umgebung innerhalb einer Evaluation nicht zu vernachlässigen, da Empfehlungen ggf. direkten Einfluss auf die Probanden oder auch auf die Umgebung, in der sich der Proband befindet ausüben. In einer Simulation oder einer Benutzerstudie im Labor wird dies nicht berücksichtigt. Durch eine online Evaluation ergeben sich deshalb aussagekräftige Ergebnisse hinsichtlich des Verhaltens der Probanden in Abhängigkeit der durch das System präsentierten Empfehlungen. Des Weiteren bietet eine online Simulation die Möglichkeit zur Beurteilung der durch die Probanden empfundenen Nützlichkeit für präsentierte Empfehlungen. Bei einer längerfristigen Nutzung des Systems kann eine online Evaluation ebenso Aufschluss über die langfristige Qualität des Systems geben, was in offline Experimenten oder Benutzerstudien i.d.R. nicht der Fall ist.

Für die Evaluation von Empfehlungsdiensten im Kontext E-Learning haben die drei beschriebenen Evaluationsformen ebenso Bestand (vgl. [101] S. 66). Auch die nachfolgenden Ausführungen, die einer generellen Beschreibung folgen, sind ebenso für den Anwendungskontext E-Learning anwendbar. Welche Evaluationsform gewählt wird ist von den Gegebenheiten der jeweiligen Entwicklung abhängig. Für rein algorithmische Entwicklungen ist

z.B. häufig die Vorhersagekraft für Empfehlungen oder der Aspekt der Leistungsfähigkeit von Interesse. In diesem Zusammenhang bedarf es dann nicht der Rückmeldung durch eine Probandengruppe, weshalb ein offline Experiment genügt. Von Bedeutung ist hierbei allerdings eine geeignete Datenbasis, auf der ein offline Experiment durchgeführt werden kann. Wenn eine prototypische Implementierung eines ganzheitlichen Empfehlungsdienstansatzes vorliegt, ist die Rückmeldung durch reale Probanden wünschenswert, um verschiedene Aspekte des jeweiligen Systems zu beurteilen. Welche Aspekte dies sind, hängt wiederum von den Zielen eines Systems ab bzw. welche Ziele ausgesprochene Empfehlungen verfolgen. Verschiedene Merkmale oder auch Attribute sind innerhalb eines Empfehlungsdiensteszenarios messbar, die zur Bestätigung von Zielen eines Systems quantifiziert werden können. [129] schlägt insgesamt 14 messbare Merkmale vor (vgl. S. 271 ff.):

- *User Preference*: Die Präferenz eines Users von Empfehlungen aus unterschiedlichen Ansätzen.
- *Prediction Accuracy*: Qualität von Empfehlungen hinsichtlich ihrer Klassifizierung.
- *Utility*: Das Maß an Nutzen, was sich für einen User oder das System selbst ergibt.
- *Coverage*: Qualität von Empfehlungen auf sehr kleinen Datenmengen.
- *Confidence*: Vertrauen in Empfehlungen durch das System selbst.
- *Trust*: Vertrauen in Empfehlungen aus Sicht eines Users.
- *Novelty*: Neuheit von Empfehlungen aus Sicht eines Users.
- *Serendipity*: Das Maß an Zufälligkeit, dass ein System eine gute Empfehlung gibt.
- *Diversity*: Das Maß an Unterschiedlichkeit zwischen Items oder Users.
- *Risk*: Das Maß an Risiko was mit einer Empfehlung einhergeht (z.B. beim Kauf von Aktien).
- *Robustness*: Die Robustheit eines Empfehlungsdienstes im Zusammenhang mit fehlerhaften oder vorsätzlich falschen Informationen.
- *Privacy*: Das Maß an Privatsphäre im Umgang mit personenbezogenen Daten innerhalb des Systems.
- *Adaptivity*: Die Möglichkeit des Systems auf Trendwenden innerhalb eines Datenbestandes zu reagieren.
- *Scalability*: Skalierbarkeit hinsichtlich größer werdender Datenmengen.

In einer Evaluation werden in der Regel nur einige wenige Merkmale ausgewählt und gemessen. In diesem Zusammenhang bedarf es der Prüfung, welche Merkmale für die sachdienliche Hinterfragung der Ziele des Systems sinnvoll erscheinen. Für die Ziele der vorliegenden Arbeit sind vorrangig die Merkmale Prediction Accuracy und Utility (in der Auflistung unterstrichen) von Interesse, weshalb beide Merkmale nachfolgend gesonderte Betrachtung finden und geeignete Methoden zur Messung vorgestellt werden³⁹.

Prediction Accuracy: Die Prediction Accuracy oder auch die *Genauigkeit*⁴⁰ bei der Herleitung von Empfehlungen ist gemäß der Literatur eines der am häufigsten geprüften Merkmale bei der Evaluation von Empfehlungsdiensten [57]. Dieser Sachverhalt ist nur wenig verwunderlich, da qualitative Empfehlungen als Grundstein zur Nutzung eines etwaigen Systems gesehen werden können⁴¹. Zur Messung der Genauigkeit ergeben sich drei Ansatzpunkte, die je nach gewähltem Ansatz des Systems zu berücksichtigen sind. So wurde in Kapitel 3 dargelegt, dass Empfehlungsdienste Bewertungen für Items schätzen oder eine Empfehlung für das Item selbst aussprechen. Für die Genauigkeit geschätzter Bewertungen gilt es den Fehler zwischen geschätzten und tatsächlichen Bewertungen (z.B. bestimmt durch einen User) zu messen⁴². Eine weitere Möglichkeit ist das Sortieren bzw. Ranken von Items. Die Bewertung erfolgt dann über den Abgleich einer systemeigenen und einer User definierten Liste. Für die vorliegende Arbeit liegt der Fokus allerdings nicht auf der Schätzung von Bewertungen oder der Sortierung von Items, sondern vielmehr auf der Klassifizierung von Items, die ein User nutzen würde. In diesem Zusammenhang gilt es die *Usage Prediction* oder *Nutzungswahrscheinlichkeit* zu messen⁴³. Zur Messung der Nutzungswahrscheinlichkeit wird die durchgeführte Klassifizierung des Systems durch eine Gruppe von Probanden bestätigt bzw. abgelehnt. So ist die systemeigene Klassifizierung den Probanden zu präsentieren, die wiederum eine eigene Klassifizierung – ohne Kenntnis über die Klassifizierung des Systems – durchführen. Der Abgleich kann gemäß einer

³⁹ Die Argumentation weshalb die genannten Merkmale Verwendung finden erfolgt in Kapitel 5 innerhalb der Herleitung eines für die vorliegende Arbeit anwendbaren Evaluationsszenarios.

⁴⁰ Prediction Accuracy ist nachfolgend als Genauigkeit benannt

⁴¹ Die Herleitung wird sich im weiteren Verlauf des Abschnitts im Kontext der Akzeptanzprüfung mittels TAM ergeben

⁴² Als gängige Maße können an dieser Stelle der *Root Mean Squared Error (RMSE)* und *Mean Absolute Error (MAE)* genannt werden

⁴³ Nachfolgend wird die Begrifflichkeit Nutzungswahrscheinlichkeit genutzt.

Klassifikationstabelle (vgl. Tabelle 2) erfolgen, die die Anzahl richtig (blau markiert) und falsch (rot markiert) klassifizierter Items in einer Vierfeld-Matrix (oder auch Kontingenztafel) quantifiziert.

	Empfohlen (System)	Nicht empfohlen (System)
Genutzt (User)	(a) True-Positive (TP)	(b) False-Negative (FN)
Nicht genutzt (User)	(c) False-Positive (FP)	(d) True-Negative (TN)

Tabelle 2 Klassifikationstabelle

Zur Verwendung der Daten bedarf es der Prüfung auf Unabhängigkeit zwischen den Klassen mittels Chi-Quadrat - bzw. χ^2 -Signifikanztest (vgl. [141] S. 525 ff.). Die statistische Signifikanz der Daten trifft Aussage über die Wahrscheinlichkeit, dass die Daten zufällig zustande und die Klassen somit unabhängig voneinander sind (Nullhypothese) oder ein tatsächlicher Zusammenhang zwischen den Klassen existiert (Alternativhypothese) und die Daten somit nicht auf Zufälligkeit beruhen. Die Ablehnung der Nullhypothese ist für die Nutzung der Daten wünschenswert. Die Berechnung der Prüfgröße ergibt sich auf Grundlage der Kontingenztafel aus Tabelle 2 mittels

$$\widehat{\chi^2} = \frac{n(a * d - b * c)}{(a + c) * (b + d) * (a + b) * (c + d)}, \quad (14)$$

wobei n der Gesamtanzahl an Klassifikation mit $n = a + b + c + d$ entspricht. Unter Berücksichtigung des Signifikanzniveaus $(1-\alpha)^{44}$ und eines Freiheitsgrades f^{45} ergibt sich ein Quantil (Q), welches mit der Prüfgröße zu vergleichen gilt. Sofern $\widehat{\chi^2} > Q$ kann die Nullhypothese abgelehnt und von einer Abhängigkeit zwischen den Klassen ausgegangen werden. Nach einer Faustregel bedarf es zur Nutzung des Chi-Quadrattests in 80% der Felder einer Kontingenztafel eine erwartete Häufigkeit von ≥ 5 , sowie einer Gesamthäufigkeit von $n \geq 60$ (vgl. [141] S.259). In Fällen, wo die beschriebenen Bedingungen nicht erfüllt sind, ist

⁴⁴ α entspricht einem Fehler, dass eine Unabhängigkeit trotz Ablehnung der Nullhypothese existiert. Wird in der Literatur als sinnvoller Wert mit 5% bzw. $\alpha = 0,05$ oder 1% bzw. $\alpha = 0,01$ angegeben (vgl. [15, 141]). Die Berücksichtigung eines kleineren Fehlers bedeutet somit eine kleinere Irrtumswahrscheinlichkeit.

⁴⁵ Ergibt sich bei einer 2x2 Kontingenztafel mit $f = 1$

die Nutzung des exakten Tests von Fisher sinnvoll. Anders als beim Chi-Quadrat Test bedarf es nicht der Errechnung und des Vergleichs einer Prüfgröße. Vielmehr errechnet der Test eine direkte Wahrscheinlichkeit (p) unter Berücksichtigung eines Signifikanzniveaus α , dass die Abhängigkeit zweier Klassen gegeben ist (vgl. ebenfalls [141] S.259). Unter der Bedingung, dass $p < \alpha$ gilt, kann die Nullhypothese abgelehnt werden. Die Errechnung von p erfolgt dabei ebenfalls über die Zelhäufigkeiten aus Tabelle 2 mit

$$p = \frac{\binom{a+b}{a} \binom{c+d}{c}}{\binom{n}{a+c}} = \frac{(a+b)! (c+d)! (a+c)! (b+d)!}{n! a! b! c! d!}. \quad (15)$$

Nach der Prüfung der Daten auf statistische Signifikanz, ergeben sich über den Abgleich aus Tabelle 2 die zur Bewertung zu errechnenden Maße *Precision* und *Recall*:

$$Precision = \frac{a}{a+c} \quad (16)$$

$$Recall = \frac{a}{a+b} \quad (17)$$

Die Werte bestimmen das Verhältnis der tatsächlich relevanten (positiven) Items im Verhältnis der insgesamt als relevant klassifizierten Items (*Precision*) bzw. das Verhältnis von positiv klassifizierten Items im Verhältnis zur Menge der Items, die als relevant hätten klassifiziert werden müssen (*Recall*). Auf Grundlage einer Evaluation kann man versuchen die Klassifizierung zu verbessern, indem man die False-Positive bzw. False-Negative Rate minimiert. Beide Werte hängen allerdings grundlegend voneinander ab, so dass die Berücksichtigung eines Wertes den jeweils anderen negativ beeinflusst. So ist es z.B. nicht möglich einem User möglichst viele positive Items zu präsentieren (Erhöhung *Recall*) und gleichzeitig davon ausgehen, dass der User einen Großteil als relevant empfindet (Erhöhung *Precision*). Es gilt somit eine Beurteilung auf Grundlage beider Werte durchzuführen und ggf. je nach Anwendungsfall einen Wert stärker zur Beurteilung heranzuziehen. In diesem Zusammenhang hat sich die Verwendung des F_1 -Scores etabliert, der das harmonische Mittel zwischen *Precision* und *Recall* darstellt:

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall}. \quad (18)$$

Utility: Die Messung der *Utility* oder *Nützlichkeit* kann als das Maß an Nutzen einer Empfehlung für einen jeweiligen User oder auch das System selbst gesehen werden. Bedingt durch den User

zentrierten Ansatz innerhalb der vorliegenden Arbeit besteht der Bedarf den Nutzen für einen User zu betrachten. In diesem Zusammenhang ist es von Bedeutung eine Bewertung für Empfehlungen aus Sicht eines Users durchzuführen. Eine gängige Methode in Systemen, wo eine Bewertung für Items durchgeführt wird, ist ebenfalls die Bewertung von Empfehlungen (z.B. in Form von 1-5 Sternen). Ein solches Vorgehen ist in der vorliegenden Arbeit ebenfalls sinnvoll.

Da die vorliegende Arbeit ein weitgehend neues System hinsichtlich des algorithmischen Verfahrens, aber auch hinsichtlich der Eingliederung in die Lehre darstellt, wird die Betrachtung eines weiteren Merkmals vorgeschlagen, welches in der grundlegenden Literatur zur Evaluation von Empfehlungsdiensten nicht genannt ist. Hierbei handelt es sich um die Prüfung der *Akzeptanz zur Nutzung* eines etwaigen Systems aus Sicht der User. In diesem Zusammenhang kann das *Technology Acceptance Model (TAM)* nach Davis genutzt werden [31]. Für Informationssysteme aus dem Bereich des E-Learning ist die Prüfung der Akzeptanz häufig ein wesentlicher Bestandteil im Kontext einer Evaluation [21]. Auch im Zusammenhang der Empfehlungsdienste ist die Akzeptanz ein regelmäßig berücksichtigtes Merkmal [68, 77]. Dies gilt in gleicher Form auch für Empfehlungsdienste im Anwendungskontext E-Learning [46, 97]. Auch wenn sich die Prüfung der Akzeptanz zunächst als weitgehend psychologisches Problem darstellen mag, so zeigt sich nachfolgend, dass die Prüfung auf quantifizierbaren Merkmalen beruht. Das TAM argumentiert in seiner ursprünglichen Form, dass die Nutzung eines technischen Systems von den Einflussfaktoren *Perceived Usefulness (U)* und *Perceived Ease of Use (E)* abhängt, die wie folgt zu definieren sind:

Perceived Usefulness (U): Perceived Usefulness beschreibt die durch einen User wahrgenommene Nützlichkeit eines Systems.

Perceived Ease of Use (E): Perceived Ease of Use beschreibt den durch einen User empfundenen Aufwand, um das System nutzen zu können.

Abbildung 16 illustriert den Zusammenhang zwischen den beschriebenen Merkmalen und der Akzeptanz für ein System bis hin zur tatsächlichen Nutzung durch einen User.

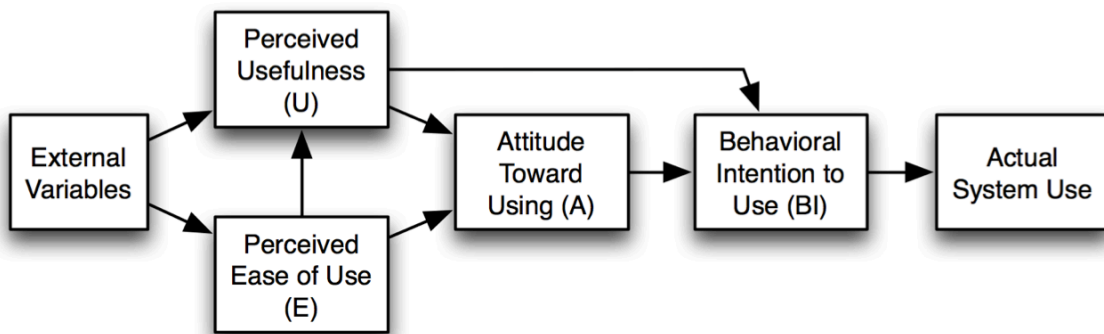


Abbildung 16 Technology Acceptance Model (TAM) nach Davis

Anders als die bislang betrachteten Merkmale für Empfehlungsdienste stellt die Akzeptanz somit kein alleinstehendes Merkmal dar, sondern ergibt sich vielmehr aus zusammenhängenden Merkmalen. Die Adaption des TAM auf Empfehlungsdienste zeigt allerdings, dass Perceived Usefulness in erster Linie in Abhängigkeit zur Qualität der Empfehlungen steht, die sich schlussendlich durch die Merkmale der *Nutzungswahrscheinlichkeit* als auch der *Nützlichkeit* von Empfehlungen ergeben [68, 77]. Des Weiteren bedarf es der Prüfung eines passenden *Kontextes* für Empfehlungen als auch das *Vertrauen* (Trust) in Empfehlungen aus Sicht des Users. Für das Merkmal Ease of Use schlägt die Literatur drei zu prüfende Merkmale vor. Hierbei handelt es sich um die *Einbeziehung* des Users in den Systemablauf, die *Komplexität* des Systems und die genügende *Usability* des Systems. Die Prüfung der Akzeptanz erfolgt häufig in Form eines Fragebogens, der z.B. auf Grundlage der bereits genannten Arbeiten in [68, 77] adaptiert werden kann. Zur Verdeutlichung überführt Abbildung 17 die genannten Merkmale in die präsentierte Darstellung des TAM in Abbildung 16.

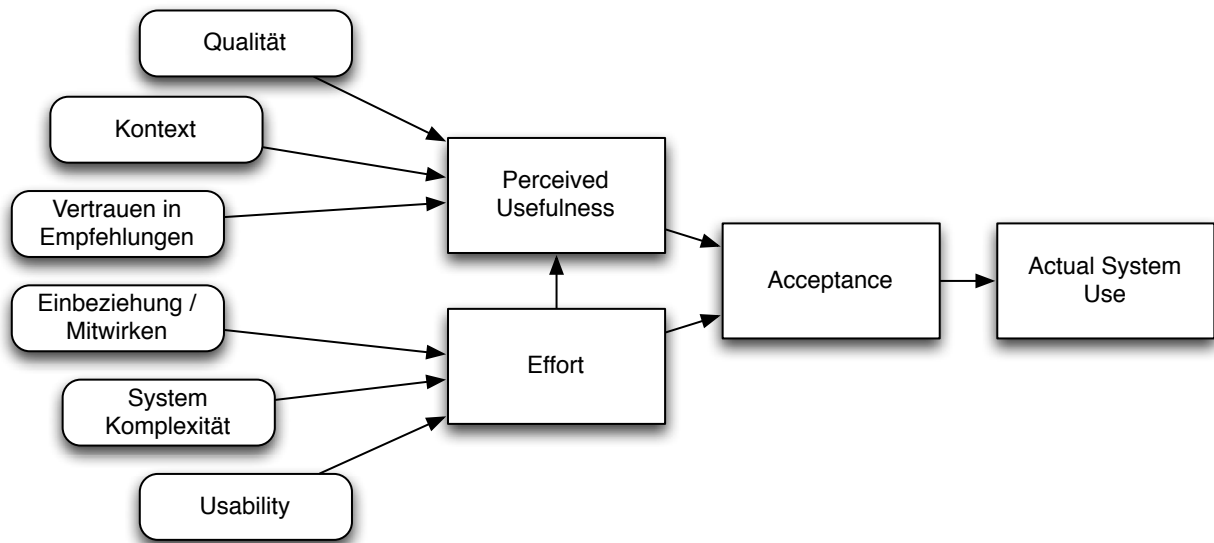


Abbildung 17 TAM Merkmale für Empfehlungsdienste

Das in der genannten Literatur vorgeschlagene Merkmal *Mood* oder auch *Stimmung*, welches Perceived Usefulness mit beeinflusst, wird an dieser Stelle bewusst vernachlässigt. Die Annahme besteht darin, dass die Stimmung innerhalb eines Lernszenarios anders zu beurteilen ist als es für Anwendungsgebiete wie z.B. im Entertainment Sektor der Fall ist. Es wird in diesem Zusammenhang eine neue Betrachtung des Merkmals gesehen, welches durch die vorliegende Arbeit – schlussendlich auch durch die informatische Ausrichtung der Arbeit – nicht abgedeckt werden kann. Das in Abbildung 17 illustrierte Modell reicht dennoch aus, um eine genügende Akzeptanzbefragung durchführen zu können.

3.3 Social Tagging Grundlagen

Bevor Empfehlungsdienste in der Anwendungsdomäne E-Learning genauer untersucht werden, soll vorab in eine weitere Thematik bzw. in ein weiteres Forschungsfeld eingeführt werden. Hierbei handelt es sich um *Social Tagging* oder auch *Social Tagging Systems*. Die Thematik zählt klassischerweise nicht zu dem Bereich der Empfehlungsdienste und ist deshalb unter diesem Begriff auch nicht beschrieben. Allerdings hat sich Social Tagging in vielen Bereichen – auch für den der Empfehlungsdienste – als wertvolle Möglichkeit zur Beschreibung oder Bewertung von Inhalten und Objekten durch Benutzergruppen erwiesen. In Abschnitt 4 wird sich noch zeigen, dass sich Social Tagging auch in der vorliegenden Arbeit als zentrales Hilfsmittel herausstellt. Der nachfolgende Abschnitt beschreibt was Social Tagging genau

bedeutet und schafft alle nötigen Grundlagen, um einen Zusammenhang zum Thema Empfehlungsdienste herstellen zu können.

3.3.1 Grundlagen Social Tagging

Unter dem Begriff Social Tagging (ST) versteht man grundsätzlich die von Personen durchgeführte Kategorisierung von Inhalten durch frei wählbare Begriffe. Diese Begriffe werden auch als *Tags* bezeichnet. Ein Tag ist somit eine benutzerdefinierte Metainformation und könnte in gleicher Form auch als Annotation bezeichnet werden. Ähnlich wie in Empfehlungsdiensten variiert die Form von Inhalten in einem Social Tagging System (STS) in Abhängigkeit der jeweiligen Anwendungsdomäne. Als populärstes Beispiel kann in diesem Zusammenhang die Webanwendung del.icio.us oder Delicious⁴⁶ genannt werden. Der Dienst bietet die Möglichkeit zur Speicherung und Verwaltung von Bookmarks in einem Benutzerkonto, sowie der Möglichkeit zur Verschlagwortung bzw. zum Taggen der Bookmarks. Dies ermöglicht die einfachere Suche in eigenen Bookmarks als auch die Verknüpfung zu weiteren Bookmarks mit ähnlichen Tags. Neben Bookmarks sind aber auch weitere Inhalte wie z.B. Fotos⁴⁷ oder Videos⁴⁸ in STS denkbar. Auch der Hashtag des Mikrobloggingdienstes Twitter⁴⁹ ist ebenfalls ein Tag zur Kategorisierung von Statusmeldungen. Ohnehin sind Tags auf modernen Webseiten ein gängiger Standard und vor allem im Bereich der sozialen Netzwerke ein elementares Werkzeug [52].

Die Beschreibung bzw. Verschlagwortung von Objekten im Internet, um eine nachträgliche Sortierung oder ein vereinfachtes Abrufen und Austauschen von Objekten durchführen zu können, kann als grundlegende Definition eines STS gesehen werden [27]. Es gibt allerdings noch weitere Betrachtungsweisen, die gemäß ihrer Zielvorstellungen variieren. [120] fasst die generellen Ziele eines STS unter den Begriffen *Searching*, *Browsing* und *Retrieving* zusammen, die in erster Linie die Anwendersicht betrachten (vgl. [120] S. 287 ff.). Grundlegend könnten die genannten Ziele auch als Funktionen eines STS verstanden werden. *Searching* meint dabei das aktive Suchen nach Informationen durch einen Benutzer z.B. anhand von Suchbegriffen. Der

⁴⁶ <http://del.icio.us/>

⁴⁷ <https://www.flickr.com/>

⁴⁸ <https://www.youtube.com/>

⁴⁹ <https://twitter.com/>

Begriff Browsing kann grundsätzlich als das Suchen ohne bestimmte Vorgabe – also z.B. ohne expliziten Begriff – gesehen werden. Unter *Retrieving* versteht man das Auffinden von neuen Informationen. [161] bietet eine alternative Betrachtung der Ziele eines STS an und schlägt drei zugrundeliegende Ziele aus Sicht der Forschung vor:

- Auffinden und Herleiten von neuen Informationen oder gleicher Benutzergruppen
- Analyse des Benutzerverhaltens
- Soziale Aspekte in STS

An dieser Stelle sieht man bereits Parallelen zur Thematik der Empfehlungsdienste, da die Ziele beider Anwendungsfelder stark miteinander korrelieren (vgl. auch Abschnitt 3.1). Vor allem für das Auffinden und Herleiten neuer Informationen als auch in der Analyse von Benutzerverhalten bestehen zentrale Anknüpfungspunkte zu Empfehlungsdiensten. Bevor beide Themenfelder weiter in Zusammenhang gebracht werden, bedarf es zunächst der Komplettierung der Grundlagen des Bereichs STS. So wurde der Begriff des Tags bereits als zentraler Bestandteil eines STS ausgemacht. Ebenso sind die Begriffe User und Inhalt benannt worden. Grundsätzlich könnte anstelle von Inhalt auch von Items gesprochen werden, um eine einheitliche Gebrauchsweise von Begrifflichkeiten zu wahren. Die Literatur bezeichnet etwaige Inhalte im Kontext von STS allerdings als *Resource*⁵⁰ [66]. Für STS wird dieser Vorgabe im Nachfolgenden nachgekommen, wobei im Kontext der Empfehlungsdienste der Begriff Item weiter Bestand haben soll. In einem STS werden die drei Gruppen User, Resource und Tag immer als eine trianguläre Verbindung betrachtet, wie Abbildung 18 darstellt (vgl. [150] S. 41):

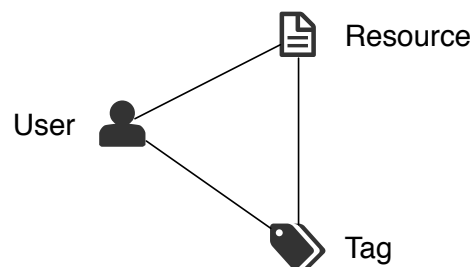


Abbildung 18 Darstellung eines Social Tagging Systems

⁵⁰ Der Begriff Resource wird im vorliegenden Kapitel bewusst mit der in der Literatur üblichen englischen Schreibweise verwendet.

Für den weiteren Verlauf der Arbeit bedarf es einer formalen Darstellung eines STS, die in Abschnitt 3.3.2 hergeleitet wird. Diese ermöglicht in Abschnitt 3.3.3 Zusammenhänge zum Bereich der Empfehlungsdienste herzustellen und gängige Verfahren zu diskutieren.

3.3.2 Der Begriff der Folksonomy

Abschnitt 3.3.1 vermittelt die Grundlagen des Themenbereichs STS. Ein ganz zentraler Begriff wurde innerhalb des Abschnitts allerdings noch nicht definiert. Hierbei handelt es sich um den Begriff der *Folksonomy*⁵¹. Generell kann eine Folksonomy als die Menge von User-generierten Schlagwörtern bzw. Tags gesehen werden (vgl. z.B. [120] S. 153 ff.). Im Besonderen ist eine Folksonomy allerdings die zugrundliegende Struktur eines STS, die nicht nur die Gruppen User, Resource und Tag, sondern auch deren Relation gemäß Abbildung 18 zueinander beschreibt [65, 66, 71]. Formal ergibt sich eine Folksonomy nach [71] wie folgt:

Seien U , T und R endliche Mengen, deren Elemente als User (u), Tags (t) und Ressourcen (r) bezeichnet werden. Die Folksonomy \mathbb{F} stellt sich als Tupel $\mathbb{F} = (U, T, R, Y)$ dar, wobei Y die ternäre Relation zwischen den Mengen U , T und R mit $Y \subseteq U \times T \times R$ ist. Die Elemente aus Y werden als Tag-assignments bezeichnet. Über Y können die User-Tag Matrix M_{UT} mit der Dimension $|U| \times |T|$, die User-Item Matrix M_{UI} mit der Dimension $|U| \times |I|$, sowie die Tag-Item Matrix M_{TI} mit der Dimension $|T| \times |I|$ abgeleitet werden, deren Einträge das Aufkommen der entsprechenden Reihe und Spalte in Y darstellen [125]. Eine alternative Repräsentationsform ist die Darstellung als Tensor (3-dimensionale Matrix) mit $A = (a_{u,t,r}) \in \mathbb{R}^{|U| \times |T| \times |R|}$ (vgl. [129] S. 618 ff.). Ein Eintrag in A kann z.B. einen binären Wert enthalten, der Aussage über die Existenz eines Tags durch einen User auf eine Resource gibt:

$$a_{u,t,r} = \begin{cases} 1, & \text{falls } (u, t, r) \in Y \\ 0, & \text{sonst} \end{cases} \quad (19)$$

Nachfolgende Abbildung 19 verdeutlicht die Interpretation als Tensor.

⁵¹ Der Begriff Folksonomy ist eine Zusammensetzung aus den Begriffen „Folk“ und „Taxanomy“ und beinhaltet somit die Klassifizierung von Ressourcen (Taxanomy) durch ein Benutzereigenes Schema (Folk).

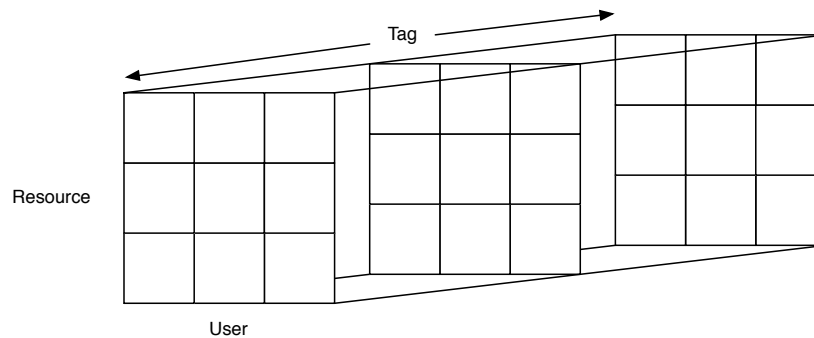


Abbildung 19 Folksonomy als 3-dimensionale Matrix

Eine alternative Betrachtungsweise ist über einen ungerichteten tripartiten Hypergraph $G = (V, E)$ möglich, wobei $V = U \dot{\cup} T \dot{\cup} R$ die Menge an Knoten und $E = \{\{u, t, r\} | (u, t, r) \in Y\}$ die Menge an Kanten im Graphen ist [71]. Ein Hypergraph stellt im Allgemeinen einen Graphen dar, dessen Kanten eine unbegrenzte Anzahl an Knoten verknüpfen kann. Die Struktur eines Hypergraphen ermöglicht durch die Verknüpfung mehrerer Knoten mittels einer Kante die trianguläre Abbildung einer Folksonomy. Bedingt durch ein gesondertes Aufbereitungsverfahren, welches auf der Anhäufung von Relationen in einer Folksonomy beruht und eine schrittweise Betrachtung der Relationen vorsieht, ist für die vorliegende Arbeit die Darstellung in Form von drei separaten Relationsmatrizen von Bedeutung. Für eine ausführliche (formale) Betrachtung der Darstellung als Tensor sowie einer Hypergraphenstruktur wird an dieser Stelle auf die Literatur verwiesen (z.B. [71, 129] S. 618 ff.).

Grundsätzlich stellt die Folksonomy die zugrundeliegende Datenbasis dar, die es im Kontext von Empfehlungsdiensten zu nutzen gilt. Verschiedene Verfahrensarten existieren für die Verwendung eines STS innerhalb eines Empfehlungsdienstes, die im nachfolgenden Abschnitt Betrachtung finden.

3.3.3 Social Tagging und Empfehlungsdienste: Verfahrensarten

Zunächst sei geklärt, weshalb beide Bereiche an dieser Stelle zusammenfinden. Verschiedene Gründe und Möglichkeiten eröffnen sich durch die Kombination der Themenbereiche. So wurde bereits erläutert, dass Probleme aus dem Bereich STS als Probleme aus dem Bereich der Empfehlungsdienste oder zumindest aus dem Bereich des Information Retrieval aufzufassen

sind. Vor allem für das Ziel des Herleitens und das Auffinden neuer Informationen wählt man häufig Empfehlungsdienstansätze. Des Weiteren erweist sich die benutzereigene Klassifizierung von Inhalten in vielen Empfehlungsdienstszenarien als probate Methode zur Erhebung von Metainformationen. Ein Tag im Kontext der Empfehlungsdienste könnte somit als eine Form des expliziten Feedbacks eines Users und somit auch als eine Art Transaction aufgefasst werden. Die klassischen Empfehlungsdienstszenarien sehen in erster Linie binäre (gekauft, nicht gekauft) oder numerische (von 1 bis 5) Bewertungen für Items vor, die in den Anwendungskontexten aus Abschnitt 3.2 sinnvollen Einsatz finden, für diverse weitere Anwendungsszenarien allerdings nicht ausreichen. Häufig bedarf es einer Bewertungsform, die eine inhaltliche Beschreibung von Items vorsieht. Dies ist in der Regel dann der Fall, wenn es einer inhaltlichen Klassifizierung von Items bedarf, die vom System z.B. bedingt durch eine mangelnde Datenbasis nicht ausgeführt werden kann. So könnte man sich z.B. ein inhaltsbasiertes Verfahren vorstellen, für dessen Items keine Attribute existieren. Durch die Verwendung von Tags könnte der Usergruppe die Möglichkeit gegeben werden eigene Attribute für Items zu generieren und so durch das System analysierbar zu machen. Man muss an dieser Stelle beurteilen, ob die Verwendung von Tags im jeweiligen Anwendungskontext sinnvoll erscheint. So wäre ein User in einem Onlineshop vermutlich nicht gewillt vor oder während eines Kaufs Tags für das jeweilige Item zu erstellen. Es hat sich allerdings mehrfach gezeigt, dass es Anwendungsszenarien gibt, in denen Tags als Bewertungsmethode aktiven Einsatz finden und etwaige Systeme hiervon profitieren (vgl. z.B. [58]). Tags können ebenso als Empfehlung ausgesprochen werden und sind dann als eine Form von Items zu betrachten. Ein Tag dient dann nicht der inhaltlichen Klassifizierung durch einen User, sondern vielmehr als inhaltliche Information zu einer Resource zur Unterstützung des Users. Ohnehin besteht die Möglichkeit innerhalb eines STS Kontextes Elemente aus jeder Menge der Tags, User oder Ressourcen zu empfehlen. In diesem Zusammenhang wird auch von *Multi-mode Empfehlungen* gesprochen (vgl. [129] S. 620).

Es zeigt sich an dieser Stelle, dass Empfehlungsdienste als auch STS zur Lösung diverser Problemstellungen im jeweils anderen Anwendungsbereich sachdienliche Mehrwerte erzielen können. So ergeben sich zwei grundlegende Anwendungstypen: *ein Empfehlungsdienst erweitert um STS Funktionalität* und *ein STS erweitert um Empfehlungsdienstfunktionliät*. Ein Empfehlungsdienst kann verschiedene Funktionsweisen bereitstellen und Items somit in

verschiedenen Varianten empfehlen, so z.B. alle sinnvollen Items oder nur einige wenige (vgl. auch [61]). Dies gilt grundsätzlich auch für Empfehlungsdienste im Kontext STS, wobei in diesem Zusammenhang noch einmal der Begriff der Multi-mode Empfehlung und somit der Möglichkeit zur Empfehlung von Tags, Usern und Ressourcen in gleicher Form genannt sei. In diesem Zusammenhang handelt es sich in der Regel um klassische Empfehlungsszenarien, die um Mechanismen der STS ergänzt wurden. Des Weiteren ist in einer Vielzahl an STS die Methode des *Ranking* von Tags oder Ressourcen gebräuchlich, wobei es sich dann um ein STS mit einer Erweiterung aus dem Bereich Empfehlungsdienste handelt. Unter dem Begriff kann die Errechnung einer gewichteten Liste von Tags oder Ressourcen verstanden werden. Um dieses Bild zu schärfen sei an dieser Stelle das populäre Beispiel der Tag-Cloud angeführt. Eine Tag-Cloud ist eine gewichtete Liste der am häufigsten gesetzten Tags auf eine Resource in einer Folksonomy [148]. Abbildung 20 zeigt die Darstellung einer Tag-Cloud der am häufigsten genutzten Tags im Repository BibSonomy⁵² für Bookmarks und Publikationen. Die Begriffe bzw. Tags mit höherem Gewicht sind dabei größer dargestellt. Das Beispiel der Tag-Cloud zeigt eine häufige Problemstellung, die vor allem für den Bereich der STS gilt.



Abbildung 20 Tag-Cloud am Beispiel der häufigsten Tags im System BibSonomy

Um einen Zusammenschluss zwischen den Bereichen Empfehlungsdienste und STS zu ermöglichen, bedarf es bedingt durch die dritte Dimension der Tags, die für die Verfahren aus

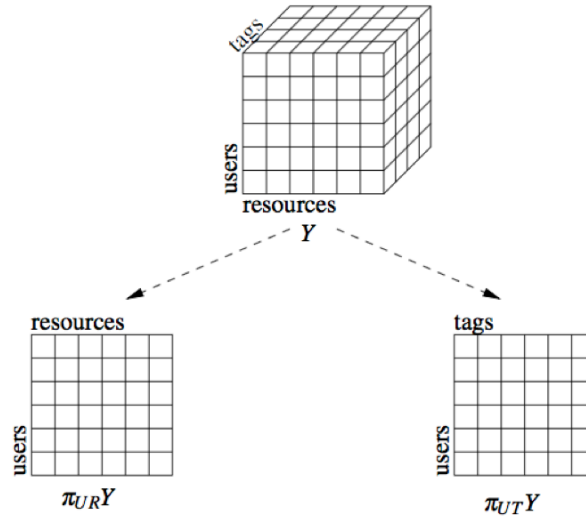
⁵² <http://www.bibsonomy.org/> - Abgerufen am 24.05.2016

Abschnitt 3.2 so nicht existierte, einer entsprechenden Anpassung der bereits vorgestellten Verfahren. Dies wurde durch die Definition einer Folksonomy als Datenbasis grundlegend motiviert. Nachfolgend seien die gängigsten Verfahrensarten für Empfehlungsdienste im Kontext STS beschrieben. Unterschieden wird hierbei zum einen nach adaptiven Empfehlungsdienstverfahren, denen ein Ansatz aus Abschnitt 3.2 zugrunde liegt und den Gegebenheiten eines STS angepasst wurde. Des Weiteren sind in Abschnitt 3.3.3.2 Ranking Verfahren präsentiert, die grundlegend als eine ergänzte STS Methodik zu sehen sind.

3.3.3.1 Adaptive Verfahrensarten

Kollaborative Verfahrensarten wurden in Abschnitt 3.2 bereits als einen der gängigsten Ansätze für Empfehlungsdienste beschrieben. Auch für den Bereich STS existiert ein entsprechender adaptiver Ansatz. Die Idee hinter dem Ansatz ist grundsätzlich äquivalent zu verstehen. Weiterhin ist es das Ziel ähnliche Benutzer innerhalb eines Systems zu finden, um hierüber Empfehlungen ableiten zu können. Nachfolgend wird der Ansatz als User-based Verfahren betrachtet, der allerdings auch analog auf Basis eines Vergleichs zwischen Ressourcen oder Tags durchgeführt werden könnte. Für den Ansatz muss die ternäre Relation zwischen den Mengen aufgelöst werden. Dies kann z.B. durch die Projektion auf zwei Teilmatrizen erfolgen, die formal wie folgt beschrieben sind (vgl. [129] S. 626 ff.):

Sei $\pi_{UR}Y \in \{0,1\}^{|U| \times |R|}$ die Projektion der Mengen U und R mit $(\pi_{UR}Y)_{u,r} := 1$, sofern $t \in T$, sowie $(u, t, r) \in Y$ existiert und 0 sonst. Des Weiteren sei $\pi_{UT}Y \in \{0,1\}^{|U| \times |T|}$ die Projektion der Mengen U und T mit $(\pi_{UT}Y)_{u,t} := 1$, sofern $r \in R$ sowie $(u, t, r) \in Y$ existiert und 0 sonst. Ein Eintrag in Matrix $\pi_{UR}Y$ trifft somit eine Aussage über das Vorhandensein der Relation zwischen U und R bzw. in $\pi_{UT}Y$ zwischen U und T . Abbildung 21 veranschaulicht die Definition:


 Abbildung 21 Projektion der ternären Relation Y auf zwei Teilmatrizen (entnommen aus [129] S.627)

Es kann an dieser Stelle entschieden werden, ob eine Gleichheit auf Grundlage der Relation zwischen User und Resource oder zwischen User und Tag ermittelt werden soll. Hierüber entscheidet sich die Nutzung der Projektion $\pi_{UR}Y$ oder $\pi_{UT}Y$. Analog zum k -Nearest-Neighbour Prinzip aus Abschnitt 3.2.2.1 wird eine Menge N von k Usern bestimmt, die eine möglichst große Ähnlichkeit zu einem User der Menge U besitzt

$$\forall u \in U, \quad N_u^k = \arg \max sim(u, v), \quad (20)$$

wobei gilt $v \in U \setminus \{u\}$ und $k \in \mathbb{N}, k < |U|$ sowie $sim(u, v)$ eine entsprechende Ähnlichkeitsfunktion darstellt, wie z.B. das Cosinus-Ähnlichkeitsmaß oder der Pearson-Korrelationskoeffizient (vgl. hierzu Abschnitt 3.2.2.1). Ausgehend von der Menge N mit der Anzahl von k Elementen können für einen User $u \in U$ Empfehlungen abgeleitet werden. Zur Erinnerung: es gibt an dieser Stelle die Möglichkeit eine Menge an Tags oder eine Menge an Ressourcen zu empfehlen. Für eine Menge an Tags \widehat{T} für einen User u und eine Resource r gilt dann:

$$\widehat{T}(u, r) := \operatorname{argmax}_{t \in T} \sum_{v \in N_u^k} sim(u, v) \delta(v, t, r), \quad (21)$$

wobei für $\delta(v, t, r) := 1$ sofern $(v, t, r) \in Y$ und 0 sonst gilt. Grundsätzlich beschreibt die Gleichung die Menge an Tags, die von anderen, größtmöglich ähnlichen Usern für eine Resource

r genutzt wurden. Der Einsatz eines äquivalenten zu oben präsentierten Verfahrens ist in [73] und [48] beschrieben.

3.3.3.2 Ranking Verfahrensarten

Es wurde bereits erwähnt, dass Ranking Verfahren grundlegend eine nach Wichtigkeit geordnete Liste für die Mengen User, Resource und Tag innerhalb einer Folksonomy erstellen. Ranking Verfahren müssen nicht notwendigerweise dem Zweck der Empfehlung von User, Resource oder Tags dienen und so steht in vielen Ansätzen ein eher genereller Informationsgewinn oder eine Informationsaufbereitung der zugrundeliegenden Folksonomy im Vordergrund. In Abschnitt 3.3.3 wurde bereits eine Tag-Cloud als gängiges Beispiel genannt.

Eines der wenigen erprobten Verfahren zum Ranking in Folksonomies ist das des *Adapted PageRank*⁵³ Algorithmus, welches zum einen ein allgemeines Ranking für die Mengen User, Resource und Tags, aber ebenso das Suchen nach Informationen über die Gewichtung von Präferenzen ermöglicht [65, 66, 71]. Das Verfahren adaptiert die grundlegende Annahme des Google PageRank Algorithmus, dass Internetseiten, die häufig von anderen wichtigen Internetseiten referenziert werden, selbst an Bedeutung zunehmen [18]. Die Annahme für den adaptierten PageRank Algorithmus besteht darin, dass Ressourcen, die häufig von wichtigen Usern mit wichtigen Tags getagged wurden, ebenfalls wichtig sind. Diese Annahme gilt in gleicher Weise für die Mengen der Resource und Tags. So entsteht ein Graph, dessen Knoten sich gegenseitig durch ihr Gewicht beeinflussen bzw. verstärken⁵⁴. Das Ranking selbst wird unter Berücksichtigung des Random Surfer Models durchgeführt, welches im ursprünglichen Ansatz (PageRank) die Wichtigkeit von Web Pages unter der Annahme bestimmt, dass ein Web Surfer Links zu anderen Webseiten folgt, von Zeit zu Zeit allerdings eine Webseite aufruft ohne einem Link zu folgen. Das Model folgt einem wahrscheinlichkeitstheoretischen Ansatz und drückt für eine Folksonomy die Wahrscheinlichkeit aus, dass z.B. ein beliebiger User zu einer Resource oder einem Tag springt. Eine Erweiterung, die die stärkere Berücksichtigung von

⁵³ In angepasster Form auch als FolkRank Algorithmus bekannt, der im weiteren Verlauf noch Erläuterung finden wird

⁵⁴ Als Ergänzung: der Originaltext beschreibt den Graphen als „... graph of vertices which are mutually reinforcing each other by spreading their weights.“ [65]

Präferenzen unter Verwendung eines Präferenzvektors z.B. für einen User vorsieht, ist unter dem Namen *FolkRank* beschrieben [65]. Verschiedene vergleichbare Verfahren, die ebenfalls auf dem Prinzip des PageRank Algorithmus beruhen, existieren z.B. zum Finden der wichtigsten User oder Tags in einer Folksonomy [154], dem Finden von Experten (respektive wichtiger User) [74] oder zum Finden der wichtigsten Ressourcen auf Grundlage der Tag Häufigkeit [170]. Auf eine formale Darstellung des Adapted Pagerank bzw. FolkRank Ansatzes die auf der in Abschnitt 3.3.2 beschriebenen Hypergraphenstruktur basieren wird hinsichtlich der Komplexität und der geringen Bedeutung für die vorliegende Arbeit verzichtet und an dieser Stelle auf die Literatur verwiesen (vgl. z.B. [65]). Die grundlegende Annahme hinsichtlich der Gewichtung von User, Tags und Ressourcen ist im weiteren Verlauf der Arbeit in adaptierter Form allerdings von Bedeutung (vgl. Kapitel 4), so dass eine kurze nicht formale Darlegung sinnvoll ist. Des Weiteren dient die genutzte Darstellung einer ganzheitlichen Übersicht über das Themenfeld der STS.

3.3.4 Nachteile Social Tagging Systems im Kontext von Empfehlungsdiensten

Die vorherigen Abschnitte haben Social Tagging Systems im Kontext der Empfehlungsdienstthematik diskutiert. Es ergeben sich klare Vorteile, die vor allem in einer alternativen Form der Bewertung innerhalb von Empfehlungsdiensten gesehen werden kann. Die freie Anreicherung von Metainformationen für verschiedene Formen von Ressourcen (oder Items) kann allerdings ebenso einen entscheidenden Nachteil darstellen. Die Interpretation von Tags ist nicht immer eindeutig möglich, was in Empfehlungsdiensten aber in der Regel gefordert ist. So wurden bereits die Begriffe Polysemie und Synonymie im Zusammenhang inhaltsbasierter Verfahren genannt, die für Tags – je nach Anwendungsszenario – ebenfalls zu berücksichtigen sind. Zur verbesserten Interpretation von Sprachbestandteilen haben sich häufig Ontologien als probates Mittel zur verbesserten Interpretation etabliert [121, 158]. Auch die Qualität an Tags ist ein entscheidendes Kriterium, welches die Interpretation von Tags beeinflusst. Ein Grund für mindere Qualität ist z.B., dass User ggf. keine eigenen Ideen für die Verschlagwortung generieren können oder auch wollen. Einige Ansätze berücksichtigen deshalb die Bereitstellung von geeigneten Tagbeispielen für eine jeweilige Resource, um einen Kontext klar zu umreißen und einen Anreiz zum taggen zu geben [50, 98].

Das bereits für inhaltsbasierte und kollaborative Verfahren diskutierte Kaltstartproblem existiert in ähnlicher Form ebenfalls für Verfahren auf Grundlage einer Folksonomy [134]. Eine Möglichkeit dieses Problem zu umgehen ist die Hinzunahme von Data Mining Strategien, die Informationen aus den Ressourcen innerhalb einer Folksonomy extrahieren und als Tags zur Verfügung stellen z.B. für Bookmarks [156] oder Musik [37]. Die genannte Literatur spricht dann auch von Tag-Empfehlungen oder Autotagging. Ohnehin ist die Akquise von Tags ein nicht zu unterschätzendes Problem, da eine Folksonomy im Grunde auf die Teilnahme von Usern angewiesen ist. Für kollaborative Systeme ist dies sicherlich in gleicher Form der Fall, allerdings kann die Hürde zum Setzen von Tags durchaus höher eingeschätzt werden, als eine Bewertung zu setzen. Die Bereitstellung von vordefinierten Tags sieht die Literatur als Möglichkeit, um diese Hürde zu überwinden (vgl. wiederum [50, 98]). Eine weitere Alternative könnte das Setzen von *Experten Tags* sein, die durch Fachleute an Ressourcen angeheftet werden (vgl. [129] S. 624).

3.4 Empfehlungsdienste im Kontext E-Learning: Zusammenhänge und Abgrenzung

Abschnitt 3.2 hat ausführlich in die Grundlagen der Empfehlungsdienste eingeführt. Empfehlungsdienste im Kontext E-Learning können grundsätzlich als Anwendungsdomäne von Empfehlungsdiensten betrachtet werden. Annahmen und Methoden aus Abschnitt 3.2 werden dabei in unterschiedlichster Form adaptiert und dienen als eine Art Grundstein für die Verwendung im Anwendungsgebiet E-Learning und schlussendlich auch für jeden weiteren Anwendungskontext. Des Weiteren hat der Abschnitt 3.3 das Themenfeld der Social Tagging Systems (STS) behandelt, welches in Kombination mit Empfehlungsdiensten ebenfalls für das Anwendungsgebiet E-Learning Betrachtung findet. STS werden dabei allerdings als eine Methode im Themenfeld der Empfehlungsdienste eingegliedert, so dass der Begriff Empfehlungsdienst als Oberbegriff zu sehen ist. Der aktuelle Abschnitt beschreibt das Forschungsgebiet *Empfehlungsdienste im Kontext E-Learning* und versucht die vorliegende Arbeit zum einen thematisch noch genauer einzuordnen als auch zu weiteren Forschungsgebieten abzugrenzen. Abschnitt 3.5 diskutiert darauf aufbauend den aktuellen Forschungsstand der Empfehlungsdienste im Kontext E-Learning für die vorliegende Arbeit.

Der Begriff E-Learning im Kontext von Empfehlungssystemen meint in erster Linie ganz generell die Verknüpfung beider Forschungs- oder Themengebiete. Es gibt einige angrenzende Forschungsgebiete, die es an dieser Stelle abzugrenzen gilt. In engem Zusammenhang stehen vor allem die Forschungsfelder *Educational Data Mining (EDM)* und *Learning Analytics (LA)*. EDM kann grundsätzlich als eine Adaption oder Anwendungsdomäne des klassischen Data Mining in Szenarien des Bildungskontextes gesehen werden. LA ist spätestens seit der Nennung im Horizon Reports 2012 als mittelfristig aufkommendes Forschungsfeld ein feststehender Begriff, der für die Analyse und Verwertung von Daten und Informationen in Lehr-/Lernszenarien steht [75]. Nachfolgende Definition sei zur Differenzierung der Bereiche herangezogen:

„Educationa data mining (EDM) is a field that exploits statistical, machine-learning, and data-mining (DM) algorithms over the different types of educational data. Its main objective is to analyze these types of data in order to resolve educational research issues.“ (vgl. [130]).

Learning Analytics wird des Weiteren wie folgt umschrieben:

„Learning Analytics is the use of intelligent data, learner-produced data, and analysis models to discover information and social connections, and to predict and advise on learning.“ (vgl. [38]).

Es zeigt sich bereits die Gemeinsamkeit beider Forschungsbereiche, deren Kernaufgabe oder -intention durchaus gleichwertig zu beurteilen ist. Vergleicht man allerdings verschiedene Ansätze der beiden Disziplinen, so erhalten die vorliegenden Rahmenbedingungen eine unterschiedlich gewichtete Rolle. Ansätze aus dem Bereich EDM verfolgen zumeist die Verarbeitung eines Sachverhaltes und berücksichtigen äußere Rahmenbedingungen oder Informationen nur bedingt. Ansätze aus dem Bereich LA verfolgen hingegen häufig die Verwendung von Daten aus einem ganzheitlichen Lehr-/Lernszenario, wie z.B. innerhalb einer Kursstruktur in einem LMS [20]. Die Analyse von Informationen geht allerdings z.T. über die des Lehr-/Lernszenarios hinaus und verfolgt auch die Nutzung von Informationen auf übergeordneter, institutioneller Ebene [147]. Des Weiteren stehen im Fokus von LA die Analyse von sozialen Verbindungen – häufig auch unter dem Begriff *Social Network Analysis* – und die Visualisierung von Informationen und Daten. Als persönliche Einschätzung wurden die Ziele des Forschungsbereich LA den aktuellen Gegebenheiten, die an einer Vielzahl an

Bildungseinrichtungen vorherrschen, angepasst. Der Forschungsbereich LA ist somit in erster Linie zeitlich geprägt, so dass Nutzen aus aktuellen Trends wie z.B. sozialen Netzwerken, aber auch aus leistungsstärkeren und vielfältigeren Technologien vor allem im Bereich des Webs gezogen wird. Der Trend zu leistungsfähigen webbasierten Systemen zur Unterstützung bei der Suche nach Informationen in der Lehre zeigt sich auch unter dem Begriff *Adaptive Educational Hypermedia (AEH)*. Das Feld Adaptive Hypermedia oder auch *Adaptive Web Systems* thematisiert die automatisierte Anpassung eines Web Systems bei der Suche nach Informationen in einem großen Raum von Informationen. Das Ziel ist sowohl die Auswahl und Präsentation geeigneter Informationen als auch die automatisierte Anpassung von Systemelementen (z.B. Navigation) in Abhängigkeit von Benutzerbedürfnissen. Adaptive Educational Hypermedia kann in diesem Zusammenhang wiederum als eine Anwendungsdomäne für Lehr-/Lernszenarien gesehen werden. Im Vergleich zu EDM und LA bedarf es bei AEH einer wohldefinierten Datenbasis, die zumeist durch Expertenwissen z.B. einem Lehrenden aufgebaut wird und in der Regel einer feststehenden Struktur z.B. einem Curriculum folgt. Hierdurch findet ADH vor allem Verwendung in Lernszenarien, in denen ein formelles Lernen vorgesehen ist (vgl. [101] S. 6 ff.). In [101] wird neben Adaptive Educational Hypermedia noch der Begriff *Learning Networks* als angrenzendes Forschungsgebiet genannt. Die Idee eines Learning Network ist die individuelle Optimierung von Lernprozessen eines einzelnen Lernalters durch das Finden von Bezugspersonen innerhalb einer Domäne (z.B. einem Kurs innerhalb eines LMS). Dies kann auf Grundlage von Aktivitäten innerhalb der Domäne geschehen, so z.B. die Nutzung von Lerninhalten innerhalb eines bestimmten Kontextes. Im Vergleich zu AEH stehen Learning Networks vor allem im Zusammenhang mit informellem Lernen. Die genannten Forschungsbereiche zeigen eine inhaltlich starke Nähe zueinander. Grundlegend wird die Analyse von Daten und Informationen aus Lehr-/Lernszenarien thematisiert, die als Grundlage für die Ableitung von unterstützenden Informationen oder Maßnahmen genutzt werden. Vergleicht man die Intention oder auch Ziele der Forschungsfelder, so stimmen diese ebenso grundlegend mit den generell formulierten Zielen von Empfehlungsdiensten aus Abschnitt 3.1 überein. Prinzipiell wäre eine Zuordnung von EDM, LA, AEH und Learning Networks unter dem Begriff Empfehlungsdienste durchaus denkbar, weshalb Ansätze aus allen Bereichen in Abschnitt 3.5 Betrachtung finden.

An diesem Punkt stellt sich noch die Frage, in wie weit allgemeine Verfahrensarten der Empfehlungsdienste auf die Anwendungsdomäne E-Learning adaptiert werden können. In den vorherigen Abschnitten wurden Empfehlungsdienste als Anwendungssystem gesehen, welches auf Grundlage von Benutzerinformationen (z.B. Bewertungen) sogenannte Items empfiehlt, die für den jeweiligen Nutzer von Interesse oder Nutzen sein könnten. Zwar bleibt im Anwendungskontext E-Learning diese Form ebenso erhalten, obgleich ein Item keinem Buch oder keine DVD, sondern vielmehr einem dem Anwendungskontext entsprechenden Objekt – z.B. ein Lerninhalt in Form eines Materials – entspricht. Allerdings stellt sich der Anwendungskontext E-Learning deutlich komplexer dar als es z.B. für ein kommerzielles System wie einem Webshop der Fall ist [131]. Zur Erläuterung seien die in Abschnitt 3.1 definierten Begriffe oder auch Variablen User, Item und Transaction neu betrachtet.

Die Rolle des Users war bislang in einer eindimensionalen Form definiert. Es handelte sich dabei um einen Endanwender z.B. um einen Käufer in einem Webshop, der auf Grundlage von Bewertungen oder Käufen neue Empfehlungen erhalten soll. Streng genommen wurde in Abschnitt 3.1 die Rolle des Anbieters eines Empfehlungsdienstes implizit mitdefiniert, allerdings im Verlauf des Kapitels nicht gesondert berücksichtigt. Vor allem die Betrachtung von Nutzeraktivitäten für kommerzielle Zwecke – z.B. Steigerung von Item Verkäufen – hätte den eigentlichen Themenschwerpunkt der Arbeit verfehlt. Zu Beginn des Abschnitts hat sich bei der Vorstellung angrenzender Forschungsgebiete allerdings gezeigt, dass die Analyse des Nutzerverhaltens eine zentrale Rolle einnehmen kann. Deshalb gilt es für die Anwendungsdomäne E-Learning zwei Benutzergruppen zu betrachten: *Lehrende* und *Lernende*. In der Regel wird die Unterstützung von Lernenden als zentraler Bestandteil eines Empfehlungsdienstes⁵⁵ gesehen. Welche dies genau sein können, wird sich bei der Betrachtung des Item Begriffs noch zeigen. Ebenso können Informationen für Lehrende die durch explizit oder implizit erhobene Daten Lernender abgeleitet wurden, in einem Lehr-/Lernszenario von Bedeutung sein. Sofern diese Informationen vom Lehrenden in geeigneter Form verwertet und

⁵⁵ Nachfolgend wird der Begriff Empfehlungsdienst auch ohne den Zusatz „im Kontext E-Learning“ synonym für die Anwendungsdomäne E-Learning genutzt.

im Austausch mit Lernenden verwendet werden, könnte ein Nutzen oder Mehrwert für beide Nutzergruppen entstehen [78].

Auch der Begriff des Items stellte sich bislang durch die beschriebenen Anwendungsfälle recht eindeutig dar. Die Deklaration eines Items z.B. in Form eines Produktes in einem Webshop scheint somit offensichtlich, da der Anwendungsfall keine weiteren Möglichkeiten der Variation erlaubt. In der Motivation wurde ein Item in einem E-Learning Szenario – zumindest für die vorliegende Arbeit – bereits als Lernmaterial/-inhalt bzw. Hilfestellungen für Lernmaterialien/-inhalte beschrieben. In diesem Zusammenhang ergeben sich verschiedene Probleme, die ebenfalls in der Motivation kurz diskutiert wurden. Zum einen ist eine Vergleichbarkeit wie bei Produkt-Items i.d.R. nicht möglich. Produkte können als eine Art atomares Objekt gesehen werden, die eine Aufgliederung in weitere Objekte im Grunde nicht erlauben. Lernmaterialien hingegen sind i.d.R. deutlich komplexer und enthalten diverse Lerninhalte, die zu unterschiedlich Zeitpunkten in einem Lernzyklus von Bedeutung sind. So kann z.B. ein Skript als ein möglich zu empfehlendes Item definiert werden, was in einer Vielzahl an Empfehlungsdienst Szenarien allerdings nicht sinnvoll erscheint. Die kontextbezogene Empfehlung von Lernmaterialien oder Lerninhalten scheint in E-Learning Szenarien somit von besonderer Bedeutung zu sein. In diesem Zusammenhang spricht man auch von kontextsensitiven Empfehlungsdiensten [1, 165]. Des Weiteren bedarf es einer Untergliederung in eine feinere Struktur, so dass sinnvolle Items z.B. für kontextbezogene Empfehlungen erst entstehen. In diesem Zusammenhang ergeben sich verschiedene Ansätze dem Problem entgegenzutreten – so z.B. die Verwendung von Learning Objects – auf die in Abschnitt 3.5.2 noch genauer eingegangen wird. Aus den obigen Ausführungen des Abschnitts ergeben sich zudem weitere Möglichkeiten für Empfehlungsdienste, die über die Empfehlung von Items hinausgehen. In [101] sind in diesem Zusammenhang vier mögliche (Forschungs-)Ziele für die Anwendungsdomäne E-Learning definiert (vgl. [101] S. 24):

- Find Novel Items: Empfehlungen von neuen oder passenden Items für bestimmte oder angrenzende Inhalte (für Lernende)
- Find Peers: Empfehlung von Lernpartnern, die gleiche Interessen oder Lerneigenschaften aufweisen (für Lernende)

- Find Good Pathways: Empfehlung von alternativen Lernwegen z.B. Nutzung von alternativen Lernmaterialien (für Lernende und Lehrende)
- Predict Student Performance: Vorhersage der Leistung von Lernenden z.B. im Zusammenhang eines Lernzieles. (für Lernende und Lehrende)

Es zeigt sich somit, dass ein EmpfehlungsdienstszENARIO nicht ausschließlich aus Sicht von vordefinierten Items konstruiert werden muss, sondern im Fall eines Lehr-/Lernszenarios auch die Empfehlung von geeigneten Lernpartnern oder Lernstrategien sinnvoll erscheint.

Auch der Begriff der Transaction erhält im Kontext E-Learning eine neue Bedeutung. Die Begrifflichkeiten des expliziten und impliziten Feedbacks in Empfehlungsdiensten wurde in Abschnitt 3.1 bereits thematisiert. Explizites Feedback wurde bislang in Form von Bewertungen betrachtet und spiegelt ebenfalls eine eher eindimensionale Ausprägung des Begriffes wieder. Im Kontext eines Lehr-/Lernszenarios scheint diese Form unzureichend. Auch in diesem Zusammenhang kann die Komplexität von Lerninhalten und –materialien und auch von Lehr-/Lernszenarios als Grund zur Neubewertung des Begriffes angeführt werden. Die Bewertung in Form von Sternen ist für Produktbewertungen ein probates Mittel und hat in diesem Zusammenhang einen genügenden Informationsgehalt, obgleich einige Fragen offen bleiben, z.B. warum eine bestimmte Bewertung zustande gekommen ist. Dennoch reicht in vielen Ansätzen für Empfehlungsdienste die Information „ein Benutzer findet ein Produkt eher gut oder eher schlecht“ aus. Betrachtet man allerdings ein Szenario zum Finden von neuen Lerninhalten wäre dieser Informationsgehalt nicht ausreichend [146], [97]. Ein Lernender könnte so z.B. eine geringe Bewertung für einen Inhalt vergeben, weil er oder sie den Inhalt als zu schwer oder zu leicht empfindet. Auch könnte eine geringe Bewertung für einen geringen Grad an Details oder eine zu ausschweifende Beschreibung stehen. Des Weiteren variieren die Bewertungen zwischen verschiedenen Benutzern. Das Kernproblem an dieser Stelle ist, dass etwaige Bewertungen für ein System nicht stichhaltig interpretierbar sind. Ein gutes Verständnis für die Entscheidungen oder Aussagen von Lernenden zu erhalten, scheint in der Domäne E-Learning allerdings wichtiger, als in Bereichen, in denen z.B. der Unterhaltungsfaktor im Vordergrund steht. Nachfolgend beschriebene Ansätze greifen größtenteils in den Lernprozess eines Lernenden ein, so dass Entscheidungen des Lerners ggf. mit beeinflusst werden. Falsche Empfehlungen können zum einen dazu führen, dass Lernende beim Lernen nicht gefördert, sondern ausgebremst

werden und des Weiteren das Vertrauen in ein etwaiges System verlieren (vgl. [129] S. 485 ff.). In diesem Zusammenhang ist auch der Begriff des impliziten Feedbacks von zentraler Bedeutung. Bislang lag der Fokus beschriebener Verfahrensarten vor allem auf der Verwendung von expliziten Informationen. In vielen Anwendungsdomänen und auch im Bereich des E-Learning haben sich implizite Informationen allerdings etabliert, um explizite Aussagen besser interpretierbar zu machen oder ableiten zu können [72, 115]. Hierbei wird grundsätzlich der Umgang mit dem System und den im System befindlichen Items oder Usern analysiert. Beispielhaft sei die Häufigkeit oder Dauer der Nutzung von Lerninhalten genannt. Vor allem im Forschungsbereich LA gibt es eine Vielzahl an Ansätzen, die vorrangig die Analyse von User Aktivitäten in Lehr-/Lernszenarien thematisieren.

Es zeigt sich an dieser Stelle, dass ein Empfehlungsdienst einer Vielzahl von Abhängigkeiten unterliegt, die je nach Zielvorgabe ungleich stark zu berücksichtigen sind. Abbildung 22 stellt die zu berücksichtigen Variablen in einem E-Learning Szenario zusammenfassend dar.

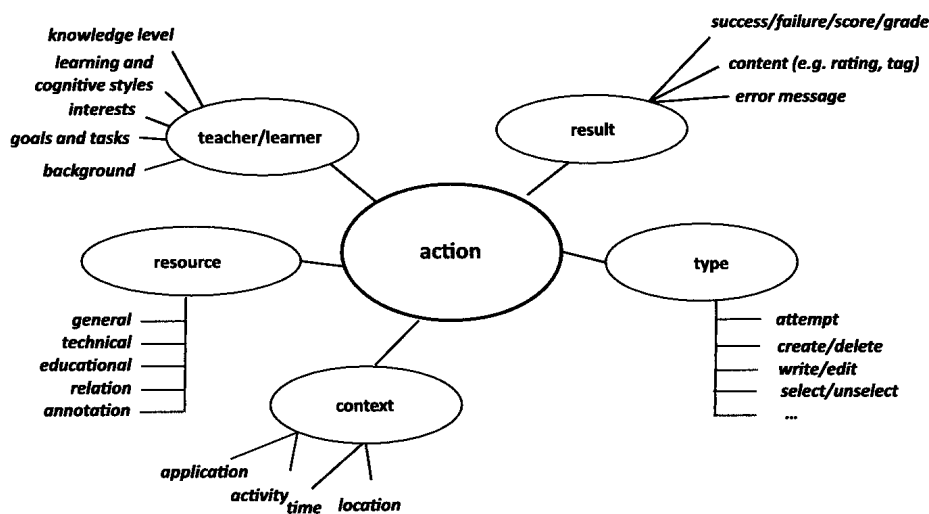


Abbildung 22 Variablen in einem E-Learning Szenario (entnommen aus [101] S. 29)

Bedingt durch die Komplexität eines Lehr-/Lernszenarios und der damit verbundenen Anknüpfungspunkte eines Empfehlungsdienstes – auch bedingt durch die in Abbildung 22 dargestellten Abhängigkeiten – ergeben sich ebenso eine Vielzahl verschiedener Ansätze, die es im nachfolgenden Abschnitt zu untersuchen und in Zusammenhang mit der vorliegenden Arbeit zu diskutieren gilt.

3.5 Empfehlungsdienste im Kontext E-Learning: Überblick relevanter Forschungsansätze

Der nachfolgende Abschnitt thematisiert aktuelle Ansätze von Empfehlungsdiensten in der Anwendungsdomäne E-Learning und erörtert ob bereits Ansätze existieren, die die Forschungsfrage der vorliegenden Arbeit beantworten. Die Form der thematischen Untergliederung weicht von der aus Abschnitt 3.2 aus Gründen einer sinnvolleren und übersichtlicheren Untergliederung ab. In Abschnitt 3.2 wurden die grundlegenden Ansätze in inhaltsbasierte, kollaborative, und hybride Verfahrensarten unterteilt, was einem der Fachliteratur entsprechenden Quasi-Standard nachkommt. Viele der nachfolgenden Ansätze und Verfahren verfolgen allerdings einen hybriden Ansatz, so dass eine Untergliederung wie in Abschnitt 3.2 wenig sinnvoll erscheint. In [46] wird eine nach Evaluationsmetriken getriebene Untergliederung vorgeschlagen, was vor allem zum jetzigen Zeitpunkt der Arbeit eine denkbar ungeeignete Lösung darstellt. Es gilt, zunächst, bestehende aktuelle Ansätze mit Verfahrensarten aus dem Bereich der Empfehlungsdienste zu verknüpfen, was gemäß einer Untergliederung nach Evaluationsmetriken nur unzureichend gelingen würde. Auf Grundlage der definierten Ziele für Empfehlungsdienste im E-Learning (vgl. Abschnitt 3.4) wird in [101] eine Eingliederung von Verfahren gemäß deren Zielvorgabe vorgeschlagen: „Find Novel Items“, „Find Peers“, „Find Good Pathways“ und „Predict Student Performance“. Da sich die vorliegende Arbeit allerdings vorrangig mit der Empfehlung von Lernmaterialien/-inhalten bzw. Items beschäftigt, scheint auch diese Unterteilung nicht optimal. Eine plausible Lösung ist an dieser Stelle die Differenzierung Item-empfehlender Verfahren gemäß ihrer zu empfehlenden Items. Bei eingehender Betrachtung lassen sich diese nämlich in die drei Untergruppen *Lernressourcen*, *Learning Objects* und *Kontextinformation* klassifizieren. Als ein weiterer Teil sollen ebenfalls Empfehlungen für Lehrende Betrachtung finden. Hierbei handelt es sich in der Regel um aggregierte Kurs- oder Nutzerinformationen, die visualisiert und dem/der Lehrenden präsentiert werden. Für diesen Bereich wird deshalb auf eine gesonderte thematische Untergliederung verzichtet und alle zugehörigen Arbeiten werden gesammelt im Abschnitt 3.5.4 diskutiert. Die nachfolgenden Abschnitte diskutieren Methoden und Ansätze, die in Zusammenhang mit der vorliegenden Arbeit gemäß der in Abschnitt 1.3 vorgegebenen Zielvorgabe stehen. Jeder Ansatz wird dabei den in Kapitel 3.2 bzw. auch 3.3 grundlegenden Verfahren zugeordnet. Für jede

Untergruppe wird des Weiteren der Begriff des Items kurz definiert. Jeder Abschnitt schließt mit einer Zusammenfassung, die für die jeweilige Untergruppe Vor- und Nachteile diskutiert und bewertet. Ein quantitativer Überblick über themenverwandte Forschungsansätze die von den Vorgaben der vorliegenden Arbeit abweichen kann der Literatur entnommen werden [46, 101].

3.5.1 Lernressource

Was sich hinter dem Begriff *Lernressource* verbirgt, scheint nach erster Betrachtung nicht wirklich greifbar. Schlussendlich könnte man Begriffe wie z.B. Lernmaterial synonym verwenden, die im Kontext von Empfehlungsdiensten allerdings nicht die tatsächliche Spannweite des Begriffs Lernressource erreichen würden. Lernressource ist in vielen Fällen ganz bewusst generisch formuliert, da viele Ansätze nicht zwangsläufig auf klassische Lernmaterialien zurückgreifen. So stehen z.B. Internetseiten häufig im Fokus eines Empfehlungsdiensteszenarios, was bedingt durch die regelmäßige Nutzung im alltäglichen Lernen durchaus plausibel erscheint. Internetseiten werden an dieser Stelle allerdings weniger als grundständiges Lernmaterial gesehen. Es gibt auch noch weitere Möglichkeiten wie z.B. Tests, Übungsfragen oder ganze Kurse, die als Item in einem Empfehlungsdienst empfohlen werden können. Etwaige Items sind in der Literatur in verschiedenen Arbeiten ebenfalls als Lernressource benannt, wie nachfolgend belegt ist.

Die Arbeit von [126] motiviert den Einsatz von kollaborativen Verfahrensarten zur Empfehlung von neuen Web-basierten Lernressourcen. Innerhalb des Ansatzes erhalten Studierende die Möglichkeit, Internetseiten zum Lernen heranzuziehen und im Kontext des Lernens zu bewerten. Die Arbeit beschreibt in diesem Zusammenhang den Bedarf an neuen Bewertungsmethoden für Lernressourcen und befindet generelle Bewertungen - wie z.B. Bewertungen durch Sterne - als unzureichend (vgl. Abschnitt 3.4). Über den Vergleich von User-Bewertungen werden in einem User-basierten Filteransatz finale Empfehlungen abgeleitet. Die Evaluation in Form einer Benutzerstudie zeigt positive Ergebnisse hinsichtlich der Nützlichkeit⁵⁶, die aus Sicht von Studierenden und Experten erhoben wurde. Die Arbeit erwägt des Weiteren die Möglichkeit zur

⁵⁶ Ein Vergleich zur vorliegenden Arbeit erfolgt innerhalb der Diskussion in Kapitel 8, um den Vergleich zu eigens erhobenen Ergebnissen durchführen zu können.

Empfehlung von Usern als Ergänzung zu Items, da die Empfehlung von Usern in einem Lernszenario ebenfalls einen Mehrwert darstellen könnte. Die Arbeiten in [30, 111, 114] gehen von einer ähnlichen Annahme aus, wobei lediglich in [114] die positiv empfundene Nützlichkeit der Empfehlungen durch Lernende in einer Studie belegt. Die Studie wurde allerdings mit bewusst kleinen und themenbezogenen Lernressourcen durchgeführt, was einem tatsächlichen Lernszenario weniger entspricht. Dass Empfehlungen aus einem sozialen Gefüge sinnvoll sein können, bestätigt auch die Arbeit von [122]. Die Arbeit beschreibt die Erweiterung eines kollaborativen User-basierten Filterverfahrens, um die Möglichkeit Empfehlungen durch selbstdefinierte und vertrauenswürdige Lernpartner im System herleiten zu lassen. Ein Vorteil, der sich hieraus ergibt, ist die Überwindung des Kaltstartproblems. Auch eine durchgeführte Studie belegt die höhere Nützlichkeit für Empfehlungen, die von befreundeten Personen entstammen, als der Empfehlungen die über die gesamte Usergruppe abgeleitet wurden. Die Items, die in diesem Ansatz beschrieben werden, sind allerdings keine direkten Lernmaterialien, sondern Testfragen, die zu Übungszwecken genutzt werden können. Auch in [14] rückt der User in Form eines Lernalters stärker in den Vordergrund, indem Bewertungen von Usern in Abhängigkeit von ihrem Wissenstandes gewichtet werden. Die Arbeit motiviert, dass z.B. bei Produkt- oder Filmbewertungen jeder User gleichermaßen qualifiziert zur Durchführung einer Bewertung ist. Dies sei in Lernszenarien allerdings anders, da der Wissensstand durchaus Einfluss auf eine qualifizierte Bewertung nimmt. So wird die Errechnung von Ähnlichkeiten zwischen Usern durch einen Gewichtungsfaktor beeinflusst. Erhobene Ergebnisse in einem Offline-Experiment zeigen verbesserte Ergebnisse von Empfehlungen unter Berücksichtigung von gewichteten Bewertungen gegenüber gängigen kollaborativen Ansätzen. [26] beschreibt ebenfalls ein kollaboratives Verfahren, allerdings auf Basis eines Item-basierten Filteransatzes. Das Verfahren wurde in der Form adaptiert, als dass die Itemempfehlungen innerhalb eines bestimmten Kontextes ausgesprochen werden. Die Arbeit beschreibt in diesem Zusammenhang zwei Stadien im Empfehlungsprozess: 1.) Errechnung der k ähnlichsten Items durch kollaboratives Filtern und 2.) die kontextbezogene Empfehlung auf Grundlage von Lernsequenzen (*Sequential Pattern Mining* kurz *SPM*). Hierzu wird weitestgehend das Verhalten der Benutzer im Zusammenhang mit Lernressourcen - z.B. Aufrufe - untersucht und mit einer im System definierten Lernsequenz abgeglichen. Die Ableitung von Empfehlungen basiert dann auf Lernressourcen, die innerhalb der Sequenz unberücksichtigt blieben. Der Ansatz berücksichtigt

generell jede Form eines digitalen Lernmaterials, ohne eine explizite Vorgabe zu nennen. In einer ersten Evaluation konnten positive Ergebnisse hinsichtlich der Nutzungswahrscheinlichkeit präsentiert werden.

[97] beschreibt einen Empfehlungsdienst zur Empfehlung von Lernmaterialien, die in Abhängigkeit von Anforderungen Studierender empfohlen werden. Die Arbeit motiviert den Bedarf eines Empfehlungsdienstes bedingt durch die Heterogenität, die eine Studierendengruppe annehmen kann. So sind Grundkenntnisse, Lernstile und persönliche Anforderungen ggf. unterschiedlich. Diese Informationen werden vorab erhoben und in einem vektorbasierten Ansatz für jeden User repräsentiert. Für jedes Lernmaterial in einer vordefinierten Datenbank leitet das System einen User-bezogenen Vektor ab, der die inhaltlichen Voraussetzungen für ein Lernmaterial und die zuvor erhobenen User-Informationen in Zusammenhang bringt. Informationen über User und Lernmaterialien werden in semantischer Form präsentiert, so z.B.: Student hat geringe Grundkenntnisse in Thema A, Lernmaterial erfordert: gute Grundkenntnisse in Thema A. Um eine Vergleichbarkeit zu ermöglichen, beschreibt die Arbeit einen Ansatz auf Fuzzy-Regeln, über die sich finale Empfehlungen ergeben. Auch in dieser Arbeit findet grundlegend jede Form von Lernmaterialien ohne explizite Vorgabe von Typ oder Umfang Berücksichtigung.

Die Arbeit von [99] beschreibt ein hybrides Verfahren auf Basis eines kollaborativen und eines inhaltsbasierten Ansatzes. Als Grundlage dienen Kurse, die jeweils durch eine Menge an Inhalten beschrieben sind. Studierende bewerten diese Kurse und erhalten im Gegenzug jegliche Form von Lernmaterialien aus weiteren Kursen vorgeschlagen. Hierfür wird zum einen ein Verfahren zur Matrixfaktorisierung eingesetzt, um Bewertungen einzelner Studierender durch die Bewertungen der restlichen Studierendengruppe zu schätzen. Des Weiteren ist das Verfahren um einen inhaltsbasierten Ansatz ergänzt, in dem die Anhäufung von Inhalten in den durch User bewerteten Kursen zu einer Gewichtung zusammengefasst ist. Anhand dieser Gewichtung können ähnliche Kurse und somit ähnliche Lernmaterialien abgeleitet werden. Die Herleitung sinnhafter Empfehlungen auf einem bestehenden Datensatz in einer ersten Evaluation gibt Grund zur Annahme, dass das Verfahren ebenfalls nützliche Empfehlungen aus Sicht eines Lernenden herleitet. Die Empfehlung von Kursangeboten ist ebenfalls eine Möglichkeit, die in

verschiedenen Ansätzen Betrachtung findet (vgl. z.B. [47, 86]). Der Fokus ist dabei allerdings weniger die Empfehlung von Inhalten oder Ressourcen als die Hilfe zur Organisation oder das Finden von Ressourcen. Dennoch zeigen die Arbeiten, dass die Organisation von Ressourcen zu einer Unterstützung beim Lernen beitragen kann.

Auch in [19] wird ein kollaborativer Ansatz für die Empfehlung von Lernressourcen in Form von Internetseiten beschrieben. Durch die Bestimmung der ähnlichsten Ressourcen leitet das beschriebene System unberücksichtigte weitere Ressourcen in Form von Internetseiten ab. Anders als zuvor bereits häufig präsentiert, verwendet der Ansatz keine expliziten Bewertungen, sondern analysiert das Benutzerverhalten bzw. implizites Feedback im Zusammenhang mit Lernressourcen und bildet dies in Form einer Ontologie ab. Über ein Metadatenmodell (*Contextualized Attention Metadata* kurz *CAM*) können Lernressourcen schlussendlich mit dem Benutzerverhalten abgeglichen werden. In [81] rückt ebenfalls die Nutzung impliziter Nutzerinformationen in den Vordergrund. Das Userverhalten wird innerhalb des Ansatzes in Form von Sessions abgebildet, die weitestgehend das Nutzungsverhalten mit Lernressourcen⁵⁷ abbildet. Einer Lernressource kann dabei jeglicher digitaler Lerninhalt zugrunde liegen, der über einen Link zur Verfügung steht. Als Empfehlungen werden somit Links zu Lernressourcen ausgesprochen. Des Weiteren sieht der Ansatz die automatisierte Verschlagwortung durch das Indexierungsframework *Apache Nutch*⁵⁸ vor, so dass eine Metadatenebene für Lernressourcen entsteht. Der Empfehlungsprozess besteht aus einem inhaltsbasierten Verfahren auf Grundlage der Verschlagwortung sowie eines TF-IDF Verfahrens. Des Weiteren ergänzt ein User-basiertes kollaboratives Verfahren den Prozess. Grundlegend werden Empfehlungen beider Verfahrensarten - inhaltsbasiert, kollaborativ - separat ermittelt und später zu einer aggregierten Liste zusammengeführt. Für das inhaltsbasierte Verfahren zeigen Evaluationsergebnisse nur bedingt positive Ergebnisse hinsichtlich der Klassifizierung von Ressourcen.

Die Empfehlung von Multimedia Inhalten wie z.B. Videos ist in [172] beschrieben. Ausgehend von einem Tag-basierten Verfahren, in dem User Lernvideos mit Begriffen verknüpfen können,

⁵⁷ Lernressourcen werden innerhalb der Publikation etwas missverständlich als Learning Objects bezeichnet.

⁵⁸ Apache Nutch: kann als Webcrawler Framework bezeichnet werden, welches ähnlich zu Suchmaschinen wie z.B. Google eine Verschlagwortung von Internetseiten durchführt - <http://nutch.apache.org/>.

werden ähnliche Lernvideos bestimmt und empfohlen. Anders als in Tag-basierten Verfahren üblich, basiert der Ansatz nicht auf einer Folksonomy als Datenbasis. Der Grund hierfür ist, dass lediglich eine Zuordnung zwischen Videos (Items) und Begriffen (Tags) besteht. Empfehlungen werden innerhalb eines Item-basierten, kollaborativen Verfahrens auf der semantischen Gleichheit von Begriffen unter Verwendung des Cosinus-Ähnlichkeitsmaßes errechnet. Eine Evaluation in Form einer Expertenanalyse belegt die positive Klassifikation von relevanten Inhalten auf Grundlage von Tags innerhalb des Verfahrens. Ein weiterer Ansatz auf der Grundlage von Tags ist in [106] beschrieben. Als Empfehlungen werden Webseiten präsentiert und Empfehlungen in Form von Videomaterial in Aussicht gestellt. Der Ansatz betrachtet ähnlich wie in [172] die Beziehung zwischen Tag und Item und lässt die Relation zu einem User wiederum unberücksichtigt. An dieser Stelle wird die Verwendung eines „Hit-Counts“ vorgeschlagen, der grundlegend die Anzahl an Aufrufen einer Lernressource bzw. Webseite bestimmt. Eine nach dem „Hit-Count“ geordnete Liste dient als Grundlage für die Ableitung von Empfehlungen. Der Ansatz scheint dabei ausschließlich in einem wohldefinierten, geschlossenen Kontext sinnvoll. Die Empfehlung von Videos im Kontext von MOOCs wird in [3] diskutiert. Die Idee hinter dem Ansatz ist es, Diskussionen im Forum eines MOOCs zu analysieren und bei Fragen eine automatisierte sowie inhaltsbezogene Empfehlung zu einem Videoinhalt bereitzustellen. Hierzu werden Forenposts zunächst in verschiedenen Dimensionen kategorisiert, z.B. existiert eine Frage im Post oder wie groß ist der Grad der Verwirrung des Users. Die im MOOC verfügbaren Videos werden zuvor zeitabhängig mit Schlagwörtern indexiert, wodurch jedes Video in inhaltliche Bereiche gegliedert ist. Durch die Errechnung eines Ähnlichkeitswertes (Cosinus-Ähnlichkeitsmaß) zwischen extrahierten Schlagwörtern in Forenposts und den Schlagwörtern in vordefinierten Videobereichen können inhaltsbezogene Empfehlungen abgeleitet werden. Die positive Klassifizierung hinsichtlich relevanter Inhalte konnte auch in [3] belegt werden. Ein ähnlicher Ansatz zur Empfehlung von Videoinhalten in MOOCs wird in [100] präsentiert. Die Arbeit beschreibt die Empfehlung von Videos in Abhängigkeit von Useraktivitäten, die sowohl aus dem Anschauen von Videos als auch der Durchführung von Online-Tests hergeleitet wird. Hieraus wird ein aktueller Wissensstand modelliert, der als Grundlage für mögliche Empfehlungen dient. Zur Modellierung wird ein Ansatz auf Basis von *Bayesian Knowledge Tracing (BKT)* verfolgt, was grundlegend die Abbildung von Fähigkeiten oder Wissen in Form von binären Wertigkeiten ermöglicht.

Eine Möglichkeit zur Empfehlung von Lernressourcen auch ohne die Nutzung eines inhaltsbasierten oder kollaborativen Verfahrens ist die Nutzung von Lernpfaden [117]. Die Idee hierbei ist es, Empfehlungen in Abhängigkeit eines Lernpfades und in Abhängigkeit von Tests zur Verfügung zu stellen. Hierzu werden verschiedene User Aktivitäten analysiert, die eine Nutzung des Lernpfades sowie die Leistung Lernender repräsentieren. Die Arbeit sieht vor allem die Heterogenität von Lernern als ein Hauptproblem, so dass Lernende z.B. mit unterschiedlichem Vorwissen Lernmaterialien nutzen. Die vorgeschlagene Methode ist in diesem Zusammenhang zwar als durchaus wertvoll einzuschätzen, allerdings bedarf es einer großen Zeitanstrengung ein geeignetes Modell bzw. einen geeigneten Lernpfad mit allen Abhängigkeiten wie z.B. Tests zu modellieren.

Ein semantische Verfahren wird in [135] präsentiert, welches die semantische Verknüpfung von Lernressourcen während der Vorlesungsnach- und Prüfungsvorbereitung vorsieht. Lernmaterialien wie Skripte und PowerPoint Folien sind dabei vom Lehrenden bereitgestellt, wobei der textuelle Inhalt der Materialien als Vergleichskriterium dient. Das beschriebene System beschreibt die semantische Verknüpfung zwischen den Lernmaterialien auf Errechnung eines Ähnlichkeitskoeffizienten (Cosinus Ähnlichkeitsmaß). Als Grundlage dient hierfür die Gewichtung von Termen gemäß des TF-IDF Verfahrens (vgl. Abschnitt 3.2.1). Erweitert wird das Verfahren allerdings um den Faktor *Term Location* (TL). Dieser gewichtet einen Term in Abhängigkeit zu der Position in einem Text. Die Annahme besteht darin, dass Terme an unterschiedlichen Positionen in einem Dokument eine unterschiedliche Bedeutung besitzen. Für Lernende können Empfehlungen in Form ähnlicher Lernmaterialien sowie Kontextinformationen (siehe hierfür auch Abschnitt 3.5.3) präsentiert werden. Der Ansatz beschränkt sich an dieser Stelle logischerweise ausschließlich auf Textdokumente. Eine im Kontext der Arbeit durchgeführte Evaluation betrachtet die verbesserte Leistung und die wahrgenommene Nützlichkeit aus Sicht von Lernenden. Vor allem die Suche nach Inhalten in einem Lernkontext wurde von den Lernenden als einfacher wahrgenommen. In diesem Zusammenhang empfinden die Lernenden eine Steigerung der Effektivität und Effizienz innerhalb des Lernens. Ein weiterer Ansatz, der Nutzen aus inhaltsbasierten Informationen von Textdokumenten zieht, ist in [39] beschrieben. Zum einen wird ein User-basiertes kollaboratives Verfahren auf Basis von User-Bewertungen beschrieben, welches Textdokumente der k-ähnlichsten User empfiehlt. Des

Weiteren beschreibt die Arbeit die Empfehlung der k-ähnlichsten Dokumente zueinander, die auf extrahierten Schlagwörtern basieren. Auch dieses Verfahren berücksichtigt lediglich Textdokumente. Ähnlich wie das Verfahren in [81] können lediglich bedingt positive Ergebnisse hinsichtlich der Nützlichkeit und Relevanz von Lernressourcen festgestellt werden.

Betrachtet man die vorgestellten Verfahren bis hierher, so ergeben sich verschiedene Nachteile, die es zu diskutieren gilt. Zunächst fällt auf, dass kollaborative Verfahrensarten auch im Bereich des E-Learning eine gängige Variante sind. Auffällig ist allerdings ebenfalls, dass ein kollaboratives Verfahren grundlegend in Kombination mit einer weiteren Verfahrensart genutzt wird. Zwar ist die Hinzunahme von inhaltsbasierten Informationen - wie in Abschnitt 3.2.3 beschrieben - ein etablierter Ansatz, allerdings reicht der Informationsgehalt von kollaborativen und inhaltsbasierten Verfahrensstrukturen häufig nicht aus. So rücken vermehrt User-abhängige Informationen in den Vordergrund, um eine genügende Datenbasis aufbauen zu können. Dies regt den Bedarf zur Erhebung impliziter Informationen an, wie im Verlauf des Abschnitts bereits präsentiert. In diesem Zusammenhang fällt auch die grundlegende Veränderung der Bewertungsmethoden für Lernmaterialien im Vergleich zu den bisher beschriebenen Bewertungsmethoden auf. Anders als die sonst übliche Bewertung durch z.B. Sterne, ziehen Verfahren im Kontext E-Learning zur Empfehlung von Lernressourcen Nutzen aus aussagekräftigeren und vor allem kontextbezogenem Informationen. In diesem Zusammenhang ist z.B. die Nutzung von Tags wie in [106, 172] ein probates Mittel, um vor allem inhaltliche Bewertungen zu ermöglichen. Auch die Analyse von Nutzeraktivitäten nimmt einen wichtigen Stellenwert ein. Die Gründe für ein komplexeres Bewertungsschema können in der ebenso komplexeren Struktur von Lehr-/Lernszenarien gesehen werden, die in Abschnitt 3.4 bereits ausführlich diskutiert wurden. So verlangt die Komplexität von Lernressourcen – vor allem für klassisches Lernmaterial wie z.B. Skripte – nach anderen Bewertungs- und Analyseverfahren als diejenigen, die bislang in Empfehlungsdiensten üblich sind. Aus diesem Umstand heraus wird häufig versucht, auf kleinere, inhaltlich geschlossene Ressourcen zurückzugreifen. So ist die Empfehlung von Webseiten oder Videoeinheiten, aber auch die Empfehlung von Benutzern eine häufig erwähnte Variante. Zwei entscheidende Nachteile wurden in diesem Zusammenhang aufgedeckt. Zum einen finden eigene Lernmaterialien von Lehrenden häufig keine Verwendung und so werden eher solche berücksichtigt, die bereits mit genügenden (Meta-)Daten angereichert

sind. Zwar beschreiben die Arbeiten von [97] und [99] die Nutzung eigener Materialien, allerdings bedarf es hierbei immer einer Aufbereitung durch Lehrende oder Experten, so dass sich eine z.T. aufwändige Eigenleistung für Lehrende ergibt. Dieser wird in z.B. [135] zwar entgegengewirkt, allerdings bleiben in diesem Ansatz Multimedia-Inhalte, die immer häufiger zum Repertoire eines Lehrenden gehören, unberücksichtigt. Ein Verfahren zur inhaltsbezogenen Empfehlung von Videos konnte hingegen in [3] gezeigt werden. Der Ansatz erfordert allerdings ebenfalls die Investition von Zeit zur Aufbereitung des Videomaterials. Um Lernmaterialien ohne zusätzliche Aufbereitung nutzen zu können, verfolgen verschiedene Ansätze die Empfehlung ganzer Lernmaterialien, so z.B. [126] oder [97]. Bedingt durch die bereits diskutierte Komplexität von Lernmaterialien sind sinnvolle inhaltliche und kontextbezogene Empfehlungen z.B. in Abhängigkeit eines Leistungsstandes in diesem Zusammenhang weniger möglich. Sicherlich ist die Erschließung alternativer Lernmaterialien gegeben, allerdings bedarf es weiterer zeitlicher Anstrengungen durch den jeweils Lernenden zur Nutzung im eigenen Lernprozess. Der Überblick hat des Weiteren gezeigt, dass Verfahren auf rein semantischer Ebene nur bedingt positive Ergebnisse hinsichtlich der wahrgenommenen Nützlichkeit oder der Klassifikation von relevanten Inhalten liefern (vgl. [81], [39]). Die präsentierten Ergebnisse in [172] konnten allerdings zeigen, dass Tags für die Klassifikation von relevanten Inhalten innerhalb eines bestimmten Kontextes ein probates Mittel darstellen.

Es hat sich gezeigt, dass der häufig fehlende Kontextbezug und die unzureichende Möglichkeit zur Empfehlung von inhaltlichen Bereichen von Lernmaterialien die zentralsten Probleme bei der Empfehlung von Lernressourcen darstellen. Eine Methode, die diesem Umstand entgegenwirkt, ist die Verwendung von Learning Objects die im nachfolgenden Abschnitt 3.5.2 beschrieben sind.

3.5.2 Learning Objects

Ein Lernobjekt oder auch Learning Object (LO) stellt einen wiederverwertbarer, inhaltlicher Baustein dar der einen Text, ein Bild oder auch multimediales Material beinhalten kann [69]. Zwar gibt es unterschiedliche Interpretationsweisen für ein LO, allerdings steht häufig die

Wiederverwendbarkeit in unterschiedlichen Lernkontexten im Vordergrund⁵⁹ [49, 110]. Die Idee ist deshalb, ein LO als atomare oder abgeschlossene Lerneinheit zu betrachten, so dass keine direkten Abhängigkeiten entstehen und LOs als eigenständiger Inhalt nutzbar sind. Jedes LO ist dabei im Idealfall durch Metadaten beschrieben, für deren Verwendung ein eigener Standard existiert [69]. Der existierende Standard ist äußerst umfangreich, so dass die Metadaten ein LO inhaltlich detailreich beschreiben können.

Die Arbeit in [146] beschreibt einen Vergleich von verschiedenen kollaborativen Ansätzen zur Ableitung von Empfehlungen der LOs auf Basis von (numerischen) Userbewertungen. Verwendet wurden Datensätze aus dem LO Repository MERLOT. Ein LO Repository kann grundsätzlich als eine Datenbank gesehen werden, in der inhaltlich gleiche oder ungleiche LOs zur Verfügung stehen. Die Studie deckt Schwächen von kollaborativen Verfahren bei der Empfehlung von LOs auf und zeigt somit, dass sich kollaborative Ansätze nicht direkt auf alle Anwendungsbereiche übertragen lassen. Häufig werden deshalb hybride Verfahren eingesetzt, die einen kollaborativen Ansatz z.B. um ein inhaltsbasiertes Element ergänzen. Es gibt noch weitere Systeme, die sich vor allem mit der Organisation, Suche und Vernetzung von Repositorien beschäftigen (vgl. z.B. [144]). Auf eine gesonderte Betrachtung wird allerdings verzichtet, da etwaige Ansätze außerhalb des Fokus der vorliegenden Arbeit liegen. Die Arbeit in [162, 167] beschreibt die Verwendung von LOs in einem ITS. Studierende können im Kontext einer Veranstaltung, in der Java gelehrt wird, ein Repository mit LOs als zusätzliches Lernmedium nutzen. Hierzu durchsucht ein Lernender das Repository nach Suchbegriffen, auf deren Grundlage neue Empfehlungen hergeleitet werden. Das System verwendet hierfür einen hybriden Ansatz aus semantischen Verknüpfungen mittels einer Ontologie sowie eines memory-basierten, kollaborativen Ansatzes zur Ermittlung von Empfehlungen durch ähnliche LOs. Sowohl das Repository, als auch die Ontologie werden durch einen Experten – also z.B. einen Lehrenden oder Mitarbeiter – generiert. Vor allem wird durch Lernende ein hohes Maß an Nützlichkeit für die hergeleiteten Empfehlungen empfunden, was eine Benutzerstudie mit Studierenden belegt. [24] beschreibt ebenfalls einen Empfehlungsdienst, der auf Grundlage von Suchbegriffen LOs auswählt. Zum einen beinhaltet der Ansatz allerdings die Auswahl von LOs

⁵⁹ In diesem Zusammenhang spricht man auch von Reusable Learning Objects.

aus mehreren verknüpften Repositorien, wodurch eine größere Auswahlmöglichkeit existiert. Zum anderen beschreibt die Arbeit die Verwendung eines Userprofils, das bei der Empfehlung von LOs mit herangezogen wird. Hierzu speichert das System vorherige Suchanfragen und deren präferierten Ergebnisse und entwickelt so ein Usermodell. Zuletzt basieren die Empfehlungen auf einem inhaltsbasierten Vergleich, der bei der Empfehlung von LOs aufgrund der umfangreichen Metadatenbasis ein sinnvolles Einsatzszenario darstellt. Auch die Arbeit in [149] beschreibt ein inhaltsbasiertes Verfahren, in dem ein Userprofil angelernt wird. In diesem Zusammenhang werden implizite Informationen - z.B. die Nutzungshäufigkeit von LOs - als auch explizite Informationen - z.B. Bewertungen von LOs - zur Ableitung eines User-Profils herangezogen. Als Einsatzzweck wird ein MOOC Szenario beschrieben. Eine Evaluation zeigt positive Ergebnisse hinsichtlich der empfundenen Nützlichkeit durch eine Usergruppe. Vor allem die Kombination aus impliziten und expliziten Informationen scheint in diesem Zusammenhang ein probates Mittel zur Herleitung darzustellen.

LOs eignen sich aufgrund ihrer sehr detaillierten Metadatenstruktur vor allem in inhaltsbasierten Verfahren, für die sich verschiedene Möglichkeiten der Vergleichbarkeit ergeben. Da LOs in der Regel thematisch abgegrenzt sind, ist an dieser Stelle auch die inhaltliche und kontextbezogene Empfehlung möglich. Theoretisch könnte so z.B. ein Lernmaterial inhaltlich auf eine Vielzahl an LOs abgebildet werden, die dann kontextbezogen verwertbar sind. Ohnehin scheinen LOs denkbar geeignet, um die Komplexität eines Lehr-/Lernszenarios aufzulösen und dedizierte Inhalte für die Verwendung in Empfehlungsdiensten nutzbar zu machen. So ist z.B. auch die Bewertung durch Lernende vereinfacht und ebenfalls in inhaltlicher/kontextbezogener Form realisierbar. [149] zeigt in diesem Zusammenhang, dass vor allem die Kombination von impliziten und expliziten Informationen deutliche Mehrwerte im Empfehlungsprozess darstellen kann. Der größte Vorteil von LOs erweist sich allerdings häufig auch als entscheidender Nachteil. Die Metadaten und Repositorien für LOs müssen wie z.B. in [162, 167] manuell durch Expertenwissen angelegt werden. Wie in [146] gibt es zwar große Repositorien, die einen freien Austausch von bereits deklarierten LOs erlauben, allerdings bedarf es eigener Anstrengungen durch den/die Lehrenden/Lehrende, sofern die Vermittlung der Inhalte in eigenen Materialien angestrebt ist. Dass Lehrende diese Anstrengung selbst überwinden, kann in der Regel nicht angenommen werden [112]. Die Beschreibung von LOs durch den Kreis der User ist an dieser

Stelle denkbar. So kann das Nutzungsverhalten von Usern eine Möglichkeit zur Ermittlung von inhaltlich zugehörigen LOs sein [113]. Des Weiteren ist zur besseren Beschreibung von LOs, z.B. die Verwendung von automatisierten Tags möglich, die über das Tagging- und Nutzungsverhalten von LOs abgeleitet werden [112]. Ohnehin erweisen sich Tags, wie auch schon in Abschnitt 3.5.1 beschrieben, als wirkungsvolle Methode zur Beschreibung von Lerninhalten jeder Art. Tags können auch selbst als Empfehlungen präsentiert werden (vgl. auch Abschnitt 3.3.3) und dienen dann als eine Art Kontextinformation, die innerhalb eines Lernmaterials oder für ein LO als Zusatzinformation für den jeweils Lernenden dient. In wieweit und in welcher Form Kontextinformationen als Empfehlungen im E-Learning verwendbar sind und welcher Mehrwert aus Kontextinformationen entsteht, wird im nachfolgenden Abschnitt 3.5.3 thematisiert.

3.5.3 Kontextinformationen

Bislang wurde die konkrete Empfehlung von Lernressourcen oder LOs thematisiert und somit die Empfehlung von weiteren Inhalten, die für einen Lernenden potentiell von Nutzen sind. Des Weiteren ist allerdings auch die Möglichkeit der Empfehlung von inhalts- oder kontextbezogenen Informationen denkbar. Hierunter kann grundsätzlich ebenfalls die Präsentation von Lernressourcen oder LOs verstanden werden, allerdings geht es nachfolgend vielmehr um die Bereitstellung von Zusatzinformationen, die das Lernen z.B. für bestehende Lernmaterialien erleichtert. Auch die Sortierung oder Kategorisierung wird an dieser Stelle als Zusatzinformation gewertet. Insofern beinhaltet der nachfolgende Abschnitt auch die Organisation von Lernressourcen. Wie nachfolgend zu sehen ist, hat sich für einen Großteil der genannten Problemstellungen der Einsatz von Tag-basierten Verfahren bewährt.

[121] beschreibt die Empfehlung von Metainformationen⁶⁰, die Lernenden für Lernmaterialien zur Verfügung stehen. Die grundlegende Annahme des Systems besteht darin, dass ein Lernmaterial durch ein Thema oder einen Begriff beschreibbar ist, es häufig allerdings alternative Begrifflichkeiten oder verwandte Themenbereiche gibt, die von Studierenden nicht erkannt werden. Zur Umsetzung beschreibt die Arbeit die Extraktion und Gewichtung von

⁶⁰ In der genannten Quelle als Annotation bezeichnet.

Schlagwörtern aus Lernmaterialien. Hierfür wird im Wesentlichen ein TF-IDF Verfahren genutzt. Auf Grundlage einer Ontologie ist es möglich, Lernenden alternative Begriffe oder Themen im Kontext eines Lernmaterials zur Verfügung zu stellen. Eine Evaluation zeigt, dass das Verfahren vor allem bei fehlenden Begrifflichkeiten mit einer ungenügenden Genauigkeit bei der Herleitung von Metainformationen aufwartet.

In [133] ist sowohl die Organisation von und die Suche nach Videoinhalten auf Grundlage von semantischen Verfahren als auch auf der Nutzung von Tags beschrieben. Das System OSOTIS verwendet zum einen automatische Analyseverfahren, um Schlagwörter aus Präsentationsfolien innerhalb eines Vorlesungsaufzeichnungsszenarios mit Dozentenbild, Audio und Präsentationsfolien zu extrahieren. Zum anderen bietet das System die Möglichkeit Inhalte innerhalb der Vorlesungsaufzeichnung durch User-generierte Tags zu versehen. Hierdurch ergibt sich für die Vorlesungsaufzeichnungen eine umfangreiche Metadatenbasis, die von Studierenden zur Suche nach Inhalten herangezogen werden kann. Für die Suche sieht das System einen Ansatz auf Grundlage des TF-IDF Verfahrens vor. Des Weiteren erstellt das System für jede Aufzeichnung eine Tag-Cloud, die entweder zur visuellen Anzeige der am häufigsten genutzten Tags für einen Videoabschnitt Verwendung findet oder Tag-abhängige weitere Inhalte präsentiert.

Innerhalb der vorliegenden Arbeit wurde bereits des Öfteren argumentiert, dass sich Tags vor allem für die inhaltliche Beschreibung eignen. Dies bestätigt ebenfalls die Arbeit in [25]. Entwickelt wurde ein Empfehlungsdienst, der zum besseren Verständnis von Textinhalten in Fremdsprachen dienen soll. Die Idee besteht in einem Tag-basierten Ansatz, um zusammenhängende Sätze und Absätze miteinander zu verknüpfen. Lernende erhalten Zugriff auf verschiedene Typen von Tags, so dass eine bessere Analyse durch das System möglich ist. Neben normalen Tags gibt es somit z.B. Topic Tags, die (Ab-)Sätze einem Thema zuordnen. Auf der Grundlage von gesetzten Tags können Bereiche mit gleichen Inhalten verknüpft und empfohlen werden. Eine positive Effektivität beim Lernen konnte vor allem in einer durchgeführten Benutzerstudie mit Studierenden gezeigt werden. Eine ähnliche Idee ist in [34] präsentiert. Die Arbeit beschreibt das Annotieren bzw. Taggen von Mathematikinhalten auf einer Webplattform. Die Lerninhalte gilt es dabei in webbasierter Form zur Verfügung zu stellen.

Lernende erhalten die Möglichkeit, Inhalte innerhalb eines Tags miteinander zu verknüpfen oder Erklärungen zu Inhalten abzugeben. Für weitere Lernende können die Tags anderer für ihr eigenes Lernen genutzt werden. Auch das Abrufen von gesammelten bzw. kategorisierten Inhalten, die unter einem bestimmten Begriff/Tag zu finden sind ist möglich. Das Abrufen von kategorisierten Informationen wird über eine Ontologie realisiert. Für die Suche nach Begrifflichkeiten ist ein semantischer Vergleich vorgesehen.

[102] beschreibt die Bereitstellung von Ratschlägen oder Hinweisen. Dies geschieht während des Lernens mit Lernressourcen, allerdings auch auf übergeordneter Ebene, z.B. bei der Durchführung eines Tests. Des Weiteren erhält ein Lernender Ratschläge zur zeitlichen Durchführung, beispielsweise von Aufgaben. Die Arbeit präsentiert somit einen ganzheitlichen Systemansatz. Um die genannten Maßnahmen durchführen zu können, werden sowohl Lernressourcen als auch Userprofile durch Kompetenzen beschrieben. Dies geschieht in Form von Annotationen und einer Ontologie. Somit besteht die Möglichkeit zu beschreiben, was mit welcher Kompetenz erreicht werden kann, in welchem Kontext die Kompetenz wertvoll ist und auf welchem Leistungsniveau sich die Kompetenz bewegt. So kann z.B. während der Nutzung einer Lernressource geprüft werden, ob und in welchem Umfang eine Kompetenz zur Nutzung der Lernressource existiert.

In [35] ist die Verwaltung von Lernressourcen und die Möglichkeit der Selbstorganisation einer Gruppe von Lernenden beschrieben. Die Arbeit beinhaltet die Verwaltung von Lernressourcen in Form von Weblinks. Interessanter ist an dieser Stelle allerdings die Idee, durch verschiedene Webtechnologien ein selbstreguliertes Lernen durch eine Gruppe Lernender zu ermöglichen. Die Idee ist somit, dass das System Expertenwissen vermittelt, Orientierungshilfen für neue Lernressourcen bereitstellt und Beurteilung von Aufgaben/Tests durchführt und somit Tätigkeiten eines Lehrenden abbildet. Grundlegend basiert der Ansatz ähnlich wie viele bereits präsentierte Arbeiten auf der Zusammenführung und Bewertung von Weblinks, auf deren Grundlage Empfehlungen (Expertenwissen, Orientierungshilfe) abgeleitet werden. Ein Mehrwert zu bereits beschriebenen Ansätzen entsteht durch die Möglichkeit der Diskussion über empfohlene Ressourcen. Ein kollaboratives Prüfungssystem vervollständigt den Ansatz, in dem Lernende selbst Aufgaben erstellen und anonym bewerten.

[60] beschreibt die kontextbezogene Empfehlung von Usern, Ressourcen und Tags⁶¹ unter Verwendung eines Graphen-basierten Ansatzes. User erhalten innerhalb des Systems die Möglichkeit Metainformationen in Form von Bewertungen oder semantischen Informationen an eine Ressource anzuheften. Bei der Suche nach Informationen durch eine Suchmaske erhält ein Lernender Zugriff auf eine nach Wichtigkeit geordnete Liste aller drei Mengen (User, Ressource, Tag). Der in der Arbeit beschriebene Vorteil besteht darin, dass ein Lernender bei der Suche nach Informationen neben ähnlichen Usern und Ressourcen auch Zugriff auf Useraktivitäten in Form von Tags erhält. Der Ansatz verwendet eine adaptierte Variante des Google Pagerank Algorithmus. Die Ergebnisse einer durchgeführten Evaluation belegen, dass das Verfahren vor allem hinsichtlich der Klassifikation von Usern, Ressourcen und Tags innerhalb eines bestimmten Kontextes gute Ergebnisse erzielt.

Kontextinformationen in Lernmaterialien können dazu dienen, ein besseres Verständnis für die präsentierten Inhalte zu erwirken. Der aktuelle Abschnitt zeigt verschiedene Möglichkeiten auf, wie Kontextinformationen geeignet in Empfehlungsdiensten genutzt werden können. Der Nachteil, dass Lernmaterialien des Lehrenden in den Hintergrund rücken oder es einer Aufbereitung der Lehrende bedarf, ergibt sich für Kontextinformationen nicht. Die Ermittlung von Zusatzinformationen geschieht grundlegend auf den zur Verfügung stehenden Materialien, so dass sich Lernende in einer Art „vertrauten“ Umgebung befinden und sich mit bekannten Materialien auseinandersetzen können. Was in den präsentierten Ansätzen allerdings fehlt, ist die Präsentation alternativer Lernressourcen. Sicherlich wurde dies z.T. bewusst vernachlässigt bzw. nicht betrachtet, allerdings genügt die Konzeption der Ansätze in einigen Fällen nicht, um einen geeigneten Empfehlungsprozess anstoßen zu können. Vor allem mangelt es an interpretierbaren Bewertungsverfahren und einem Konzept, um alternative Materialien dem System zugänglich zu machen.

3.5.4 Empfehlungen für Lehrende

Bislang wurde vor allem die Herleitung von Empfehlungen für Lernende (Lernressourcen, Learning Objects, inhaltsbezogene/kontextbezogene Informationen) betrachtet. Abschnitt 3.4

⁶¹ In der Publikation als Actor, Assests und Tags bezeichnet.

motiviert bereits, dass aus Empfehlungen für Lehrende ebenfalls ein Mehrwert für Lernende resultieren kann. Der Begriff „Empfehlung“ ist in diesem Zusammenhang allerdings nicht vollkommen optimal gewählt. Wie sich im Verlauf des nachfolgenden Abschnittes zeigen wird, erhalten Lehrende im Grunde keine Empfehlungen, sondern vielmehr einen Überblick über das Verhalten oder die Leistung Lernender. Wie der Lehrende diese Informationen nutzt, liegt dann im eigenen Ermessen. Dennoch können Lernende davon profitieren, sofern der Lehrende diese Informationen nutzt, um z.B. auf erkannte Schwächen von Einzelnen oder einer Gruppe einzugehen [78].

Häufig ist die Entwicklung von zusätzlichen Plattformen für die Nutzung in einem Empfehlungsdiensteszenario nicht zwangsläufig notwendig, da Daten, die vor allem für Lehrende von Interesse sind, bereits in bestehenden Systemen wie einem LMS von Lernenden generiert werden. Die Arbeit in [36] beschreibt die Verwendung von etwaigen Daten zur Aufbereitung für Lehrende. So erhält der Lehrende einen aggregierten Überblick über die Nutzungshäufigkeit von Lernressourcen, Verhalten von Usern sowie Leistungsinformationen - z.B. in Tests und Übungen - einzelner Lernender. Diese lassen sich nach verschiedenen Eigenschaften - z.B. Geschlecht, Kurs oder Zeitspanne - visualisieren. Ohne explizite Empfehlungen zu präsentieren, können Lehrende dennoch einen konkreten Einblick in die Aktivitäten von Studierenden gewinnen. [56] beschreibt einen ähnlichen Ansatz, um das Nutzungsverhalten von LOs in einem LMS zu untersuchen. Hierfür wird eine Art Query Browser zur Verfügung gestellt, mit dem die Abfrage von Datensätzen – ähnlich, aber vereinfacht zu Datenbankabfragen – möglich ist. Die Interpretation der Ergebnisse ist wiederum durch den Lehrenden durchzuführen. Eine ganze Reihe weiterer Arbeiten beschäftigt sich mit der Thematik zur Auswertung und Analyse von Nutzerdaten (vgl. z.B. [109]). Eine Ergänzung findet häufig in Form von gesonderten Visualisierungstechniken⁶² statt, um die Analyse der Daten für Lehrende zu vereinfachen. So beschreibt [103] z.B. für jeden User im System eine aggregierte Übersicht, in der Testergebnisse und Kursfortschritte in Zusammenhang mit Aktivitäten wie Materialzugriffe und der Teilnahme an Diskussionen über den gesamten Kurszeitraum aufgelistet sind. Diese Zuordnung wird über

⁶² Gesonderte Visualisierungstechniken sind solche, die über die gängige Form der Visualisierung – z.B. mit Diagrammen – hinausgehen.

eine zweidimensionale visuelle Matrix hergestellt. Lehrende empfinden innerhalb des Ansatzes vor allem den Überblick von Aktivitäten innerhalb einer Gesamtübersicht als nützlich, was eine durchgeführte Studie mit Lehrenden belegt. Auch die Visualisierung des Verhaltens von Lernenden in Form eines Graphen ist möglich [76, 145]. Es ist die Abbildung von Domains beschrieben, wobei eine Domain die Abfolge von Schritten vorsieht, um ein Problem - z.B. eine Aufgabe - zu lösen. Das Nutzerverhalten wird über einen Graphen abgebildet, in dem Knoten als Zustand und Kanten als eine Aktion zu sehen sind, die ein Lernender für den Wechsel zwischen zwei Zuständen getätigt hat.

Die vorgestellten Ansätze innerhalb dieses Abschnittes beschränkten sich vor allem auf die Visualisierung und Analyse von Benutzerverhalten. Die Interpretation kann je nach Menge an Kursen und Studierenden aufwändig sein. An dieser Stelle ist es ggf. wünschenswert, Automatismen zur Verfügung zu stellen, die Lehrende gezielt auf auffällige Informationen aufmerksam machen. Ein Beispiel gibt die Arbeit in [6], die eine automatisierte Alarmierung für Lehrende bereitstellt, sofern bestimmte Leistungen durch Studierende nicht erfüllt werden. In erster Linie beschreibt der Ansatz dabei kollaborative Aktivitäten innerhalb eines MOOC-Szenarios. Dass problematische Situationen durch das System aufgedeckt werden, konnte eine Evaluation innerhalb eines Kurses bestätigen. Auch die Arbeiten in [79, 166] bestätigen die Mehrwerte für Lehrende die aus der Analyse von (Lern-)Aktivitäten innerhalb von MOOC Szenarien hervorgehen. [173] beschreibt die Analyse von Nutzeraktivitäten innerhalb eines LMS und die automatisierte Intervention durch das System. So erhalten z.B. Lernende eine E-Mail, sofern sie bestimmte Lernmaterialien innerhalb eines Lernkontextes nicht genutzt haben. Auch Lehrende erhalten eine Einsicht in die durch das System getätigten Aktivitäten.

Die Vorhersage der Lernleistung Studierender wäre ein weiterer Schritt, der etwaige Systeme ergänzen würde [87, 101]. Da das Themenfeld allerdings nicht im Fokus der vorliegenden Arbeit liegt, wird an dieser Stelle auf die Literatur verwiesen.

Empfehlungen für Lehrenden können grundlegend als eine Art optionaler Zusatzinformation gesehen werden. Sie erbringen wie beschrieben einen Mehrwert für Lehrende, richtig zurückgekoppelt aber auch einen Mehrwert für Studierende. In diesem Zusammenhang könnten etwaige Informationen somit als Grundlage für Expertenempfehlungen – durch den Lehrenden –

dienen. Dass eine etwaige Systemfunktionalität als nützlich durch Lehrende empfunden wird, konnten z.B. die Arbeiten [103] und [6] auf Grundlage durchgeführter Studien belegen.

3.6 Zusammenfassung

Die vorangegangenen Abschnitte haben die Grundlagen für den Bereich der Empfehlungsdienste geegnet sowie im Anwendungskontext E-Learning betrachtet. Das nachfolgende Kapitel soll dazu dienen, Lücken im aktuellen Stand der Forschung unter Berücksichtigung der Ziele der vorliegenden Arbeit zu lokalisieren und zu benennen. Hierzu dient eine sachgerechte Zusammenfassung der bis hierher diskutierten Ansätze aus dem Anwendungskontext E-Learning.

Der aktuelle Forschungsstand wurde für die vorliegende Arbeit in Empfehlungen von *Lernressourcen*, *Learning Objects* und *Kontextinformationen* untergliedert und betrachtet. Des Weiteren sind Empfehlungen für Lehrende diskutiert worden, die allerdings eine gesonderte Stellung einnehmen und keinen direkten Vergleich zu der übrigen Untergliederung ermöglicht. Es hat sich gezeigt, dass bereits verschiedene Ansätze zur Empfehlung von Lernmaterialien, Inhalten und Informationen existieren, die ihrerseits mit verschiedenen Vor- und Nachteilen⁶³ aufwarten. Für eine genügende Übersichtlichkeit sind diese auf Grundlage der Abschnitte 3.5.1, 3.5.2 und 3.5.3 in Tabelle 3 zusammengefasst.

⁶³ Vor- und Nachteile sind dabei auf die Zielvorgaben der vorliegenden Arbeit bezogen. Eine Wertung für vergleichbare Arbeiten ist mit der Begrifflichkeit somit nicht gemeint, da sich Ziele und Intentionen ggf. unterscheiden.

	Vorteile	Nachteile
Lernressourcen	<ul style="list-style-type: none"> • Einfache Implementierung in Lehr-/Lernszenarien [26, 39, 81, 99, 106, 122, 135] • Bestehende Lernmaterialien sind verwertbar [3, 14, 26, 39, 81, 97, 99, 100, 117, 122, 135, 172] 	<ul style="list-style-type: none"> • Empfehlung ganzer Lernmaterialien [3, 14, 26, 81, 97, 100, 106, 117, 126] • Somit i.d.R. keine inhaltlichen Hilfestellungen möglich [3, 19, 26, 81, 97, 100, 106, 117, 126] • Kontextbezug wird häufig vernachlässigt [19, 81, 99, 106, 117, 126] • Alternative Lernmaterialien haben keinen inhaltlichen Bezug [19, 81, 106, 126] • Kontextbezogene oder inhaltliche Bewertungen sind nicht möglich [19, 81, 100, 106, 117, 126] • Bedarf angepasster Bewertungsmethoden [14, 19, 117, 126, 172] • Erweiterung kollaborativer, inhaltsbasierter Verfahren [3, 81, 97, 99, 100, 172] • Ausschließlich textuelle Dokumente [39, 135]
Learning Objects	<ul style="list-style-type: none"> • Inhaltsbezogene Bewertung, Empfehlung [146, 149] • Kontextbezogene Bewertung, Empfehlung [146, 149, 162, 167] • Inhaltsbasierte, kollaborative Verfahren sind ohne Erweiterung nutzbar [146, 162, 167] • Bestehende Metadatenbasis ermöglicht Empfehlungen ohne explizite Userbewertung [24, 149, 162, 167] 	<ul style="list-style-type: none"> • Hohe Eigenleistung zur Erstellung notwendig, sofern bestehende Lernmaterialien bzw. Inhalte verwertet werden sollen [146, 149, 162, 167] • Vorhandene Repositorien decken nicht den Bedarf von Lehrenden ab [24, 146] • Empfehlung alternativer Materialien zumeist in Form von LO [24, 149, 162, 167] • Inhaltliche Hilfestellungen eigentlich möglich, häufig aber nicht berücksichtigt [24, 146, 149, 162, 167]
Kontextinformation	<ul style="list-style-type: none"> • Direkter Kontextbezug (z.B. in Form von Aufgaben) ist vorgesehen [35, 60, 102, 121, 133] • Inhaltliche Hilfestellungen sind möglich [25, 34, 35, 60, 102] • Bestehende Lernmaterialien sind in einigen Verfahren vorgesehen [25, 121, 133] • Die Herleitung erfolgt häufig ohne explizite Userbewertung [121, 133] 	<ul style="list-style-type: none"> • Keine direkte Empfehlung von alternativen Lernmaterialien [25, 34, 35, 60, 102, 121, 133] • Implementierung in Lehr-/Lernszenarien sind z.T. umständlich [35, 121] • Lediglich für webbasierte Inhalte etabliert [35, 60]

Tabelle 3 Vor-/Nachteile aktueller Verfahren

Die Ziele für die vorliegende Arbeit können kurz als maschinell erstellte Empfehlungen von Hilfestellungen für bestehende Lernmaterialien sowie Verknüpfungen zwischen bestehenden und neuen Lernmaterialien zusammengefasst werden. Als bedeutsam wurde innerhalb der Zielvorgabe herausgestellt, dass Empfehlungen inhaltsbezogen erfolgen sollen, so dass sich diese auf einen bestimmten Bereich innerhalb des jeweiligen Lernmaterials beziehen. Des Weiteren konnte im Verlauf der Arbeit herausgestellt werden, dass Empfehlungen oder Hilfestellungen innerhalb eines wohldefinierten Kontextes auszusprechen sind. Auch die einfache Integration in bestehende Lehrkonzepte ist innerhalb der Zielvorgabe von Bedeutung. Betrachtet man in diesem Zusammenhang nun bestehende Forschungsansätze, so existieren grundsätzlich für jeweils unterschiedliche Teilbereiche der Vorgabe Lösungen. Verfahren zur Empfehlung von Lernressourcen ermöglichen in der Regel die Verwendung bereits bestehender Lernmaterialien. Durch die Verwendung von Learning Objects wird eine inhaltliche Verknüpfung ermöglicht. Inhaltliche Hilfestellungen innerhalb eines bestimmten Kontextes ergeben sich durch Empfehlungen bzw. Verfahren die unter der Bezeichnung Kontextinformationen beschrieben wurden. Durch die existierenden Nachteile der jeweiligen Verfahrensarten ergeben sich allerdings Einschränkungen, die Einfluss auf die Umsetzbarkeit aller Teilaspekte haben. Bestehende Lösungsansätze für die Empfehlung von Lernressourcen ermöglichen häufig nur die Empfehlung ganzer Lernmaterialien, so dass ein inhaltlicher Bezug verloren geht. Evaluationen in diesem Bereich zeigen des Weiteren, dass für die Überprüfung der Verfahren häufig bewusst sehr kleine Lernmaterialien ausgesucht werden, um die Bewertung durch Probanden einer Evaluation erst zu ermöglichen (vgl. z.B. [14, 114])⁶⁴. Learning Objects ermöglichen zwar eine inhaltliche Verknüpfung, wobei bestehende Materialien lediglich nach einer Neuaufbereitung Verwendung finden. Kontextinformationen bieten die Möglichkeit inhaltlicher Hilfestellungen an, allerdings sind hiermit keine Verknüpfungen zu neuen Lernmaterialien verbunden. Verschiedene untergeordnete Nachteile existieren für die drei herausgearbeiteten Systemgruppen und so variiert z.B. der Aufwand zur Einbettung in ein Lehr-/Lernszenario stark. Des Weiteren existieren z.B. für Lernressourcen nur ungenügende Bewertungsmethoden, die die Umsetzung der Ziele innerhalb der vorliegenden Arbeit i.d.R. nicht ermöglichen. Es ist allerdings deutlich zu

⁶⁴ Die ausführliche Diskussion von empirischen Ergebnissen vergleichbarer Verfahren erfolgt in Abhängigkeit der Evaluation der vorliegenden Arbeit in Kapitel 8.

sehen, dass sich die Systemgruppen gegenseitig ergänzen, wodurch Nachteile ggf. aufgewogen werden könnten. Es mangelt an dieser Stelle somit an einem kombinierten Lösungsansatz, der eine maschinelle Untergliederung für bestehende Lernressourcen oder Lernmaterialien vornimmt, die wiederum maschinell mit Kontextinformationen angereichert werden. Ausgehend von der aktuellen Forschungsliteratur ist ein vergleichbarer Ansatz bislang nicht bekannt.

Aus der präsentierten Literatur ergeben sich dennoch einige wertvolle Hinweise die als Ausgangspunkt für eine entsprechende Entwicklung dienen können. So scheint die Komplexität von vielerlei Lernmaterialien genügend groß, als dass die Möglichkeit zur inhaltlichen Verknüpfung realistisch erscheinen könnte. Hier bedarf es einer besseren inhaltlichen Untergliederung, wie es z.B. für Learning Objects der Fall ist. Es wäre somit sinnvoll, Lernmaterialien analog zu Learning Objects in inhaltlich zusammenhängende Bereiche zu untergliedern. Wie eine Untergliederung – vor allem in maschineller Form – aussehen kann, ist bis zum jetzigen Zeitpunkt noch ungeklärt. Maschinelle Ansätze existieren in diesem Zusammenhang nur vereinzelt und basieren in der Regel auf semantischen Informationen. Dass rein semantische Verfahren allerdings klare Schwächen in der Klassifikation von relevanten Inhalten aufweisen, hat sich durch die Übersicht ebenfalls gezeigt [39, 81, 121]. Somit wird die Extraktion inhaltlicher Bereiche in maschineller Form als bestehende Forschungslücke gesehen. Alternativ könnte man auch von einer kontextbezogenen Untergliederung ausgehen. So sind inhaltliche Bereiche eines Lernmaterials in unterschiedlichen Lernkontexten - z.B. die Bearbeitung unterschiedlicher Aufgaben - ggf. von unterschiedlicher Bedeutung. Die Literatur zeigt in diesem Zusammenhang ohnehin, dass ein Lernkontext z.B. in Abhängigkeit eines Lernstandes oder eines Wissensstandes von zentraler Bedeutung für ein entsprechendes Empfehlungsdienstszenario ist. Die Empfehlung von Lernmaterialien oder Inhalten ohne direkten Bezug würde eine Eigenleistung von Studierenden zur sachgemäßen Einordnung erfordern, was ggf. zu Mehraufwand führen könnte. Dies sollte in einem unterstützenden System in jedem Fall vermieden werden oder zumindest ohne bedeutsame Anstrengung möglich sein. Des Weiteren zeigt die Literatur, wie bedeutend eine genügende Bewertungsmethode zur Empfehlung von Lernmaterialien oder Learning Objects, aber auch für Kontextinformationen ist. Die Aussage einer Bewertung muss durch ein System klar interpretierbar sein und einen genügenden Informationsgehalt liefern, um Aussagen eines Users zu verstehen und in

Abhängigkeit des jeweiligen Materials oder Inhalts zu verwerten. Die Hinzunahme von impliziten Informationen ist eine sinnvolle Möglichkeit, um die Aussagekraft von einfachen Bewertungstechniken wie z.B. von numerischen Werten zu steigern. Es hat sich allerdings gezeigt, dass die benutzereigene Klassifizierung oder Verschlagwortung durch Tags in diesem Zusammenhang am meisten Aussagekraft besitzt [25, 60, 172]. So ist unter Verwendung von Tags nicht nur eine genügende inhaltliche Bewertung, sondern auch die Generierung von Kontextinformationen - z.B. durch Ableitung von Tag-Empfehlungen - möglich.

Empfehlungen für Lehrende sollen an dieser Stelle nicht aus dem Fokus der Entwicklung geraten, da sich hierüber, wie in Abschnitt 3.5.4 argumentiert, ebenfalls Vorteile für Lernende ergeben können [6, 103]. Bedeutsam ist dabei, dass sich die Herleitung von Informationen für Lehrende nahtlos in die Entwicklung eines Empfehlungsdienstes einfügt, ohne dabei auf bereits beschriebene Vor- bzw. Nachteile einzuwirken. Betrachtet man allerdings die präsentierte Literatur, so ergeben sich die Informationen für Lehrende ohnehin in impliziter Form aus dem Verhalten Lernender innerhalb eines bestimmten Lernszenarios. Die Annahme besteht somit darin, dass die sorgfältige Berücksichtigung eines Lernszenarios aus Sicht eines Lernenden zwangsläufig zur genügenden Erhebung von Informationen für Lehrende führt. Die Abfrage von Informationen für Lehrende geschieht somit häufig ohne direkten Einfluss auf das Entwicklungsszenario. Zwar gilt es diesen Sachverhalt innerhalb der Entwicklung zu prüfen, allerdings kann an dieser Stelle von einer impliziten Abfrage der Informationen ausgegangen werden.

4 Der Empfehlungsdienst LAOs

Kapitel 3 hat dargelegt, inwieweit Verfahren existieren, die den gestellten Zielen aus Abschnitt 1.3 genügen. Es konnte gezeigt werden, dass zwar verschiedene Ansätze als mögliche Problemlösung vorhanden sind, die die gestellten Anforderungen allerdings nicht im vollen Umfang erfüllen. Die Argumentation stellt fest, dass es grundlegend einer Kombination aus Verfahren zur Empfehlung von Lernressourcen, Learning Objects und Kontextinformationen bedarf. Die Empfehlung von Learning Objects muss man an dieser Stelle allerdings relativieren, denn wie bereits in Abschnitt 3.6 dargelegt, geht es nachfolgend nicht zwangsläufig um die tatsächliche Empfehlung von Learning Objects, sondern vielmehr um die Extraktion von inhaltlichen Teilbereichen, was dem Prinzip eines Learning Objects nahe kommt. Das Finden einer geeigneteren Begrifflichkeit in den nachfolgenden Abschnitten soll einer missverständlichen Interpretation entgegenwirken. Zunächst bedarf es allerdings der Klärung von Zielen und Anforderungen für das zu entwickelnde System. Zwar sind bereits Ziele in Abschnitt 1.3 benannt worden, allerdings gilt es diese mit den Sachverhalten aus Kapitel 3 zu prüfen und in Zusammenhang zu bringen. Auch die Schlussfolgerungen aus Kapitel 2 fließen an dieser Stelle in die Überlegung mit ein. Abschnitt 4.3 gibt dann einen Überblick über die Systemimplementation des Empfehlungsdienstes, welcher nachfolgend als *Learning Assistance Osnabrück (LAOs)* bezeichnet wird. Gemäß Abbildung 11 beschreibt der Abschnitt die zugrundeliegende Datenbasis und definiert die Gruppen der Items, User und Transactions. Des Weiteren wird der Zusammenhang oder die Relation zwischen Items, User und Transaction erläutert, was einen zentralen Bestandteil der Systemlogik ausmacht. Der Bereich der Empfehlungsdienst Algorithmik ist gesondert in Abschnitt 4.4 betrachtet, da dies bedingt durch Umfang und Bedeutung des Bereichs sinnvoll erscheint.

Sowohl die Grundlegende Idee als auch die Empfehlungsdienst Algorithmik wurde in verschiedenen Veröffentlichungen publiziert (vgl. [40, 43–45]).

4.1 Ziele und Anforderungen

Abschnitt 1.3 stellt in Zusammenhang mit der gestellten Forschungsfrage bereits einige Ziele bzw. Anforderungen an einen zu entwickelnden Empfehlungsdienst. Die Argumentation in Kapitel 2 stützt diese Anforderungen, in dem einige Schwierigkeiten, die beim Lernen mit digitalen Lernmaterialien auftreten können, aufgedeckt wurden. Kapitel 3 - und in besonderer Weise Abschnitt 3.5 - hat einen aktuellen Überblick über mögliche Lösungsansätze gegeben, die in Abschnitt 3.6 einer Bewertung in Zusammenhang mit den gestellten Zielen unterzogen wurden. Es gilt an dieser Stelle die erarbeiteten Sachverhalte für eine Neuentwicklung zusammenzutragen und entsprechende Ziele/Anforderungen zu definieren bzw. solche aus Abschnitt 1.3 zu erweitern. Mit Zielen ist dabei die Umsetzung einer bestimmten Funktionalität⁶⁵ gemeint. Anforderungen beschreiben hingegen Rahmenbedingungen, die es innerhalb der Entwicklung zu berücksichtigen gilt.

Die Kernprobleme, die in Kapitel 2 ausgemacht wurden, sind in der richtigen Auswahl an und der genügenden Auseinandersetzung mit Lernmaterialien zu sehen. Zum einen ist die große Auswahl an alternativen Lernmaterialien – vor allem im Internet – als ein Grund zu nennen. Die geringe Zeit, die Lernende mit neuen Lernmaterialien verbringen, ist ein weiteres Problem. Wie bereits ausführlich diskutiert, muss an dieser Stelle der Prozess des Auffindens neuer Materialien stark vereinfacht werden. Da bestehende Lernmaterialien Lehrender generell im Fokus von Lernenden stehen erscheint es ebenfalls sinnvoll ein EmpfehlungsdienstszENARIO ausgehend von bestehenden Materialien zu betrachten und aufzubauen. Grundlegend hat sich dies auch in der aktuellen Forschungsliteratur als sinnvoll erwiesen (vgl. Abschnitt 3.5.1 und 3.5.2). In diesem Zusammenhang ist es bedeutsam, Lernende an bestehende Materialien zu binden bzw. die Nutzung bestehender Materialien zu motivieren. Wie die Literatur gezeigt hat, kann dies

⁶⁵ Die vorliegende Arbeit beschreibt die Funktionalität der Systemimplementation auf einer allgemeinen Ebene, ohne den Detailgrad durch Nutzung von Methoden der Softwareentwicklung (z.B. Klassendiagramm) vollständig auszufüllen. Für eine technische Betrachtungsweise auf die Implementation wird an dieser Stelle auf Anhang A verwiesen, der eine geeignete Betrachtung in Form einer Entwicklungsdokumentation bereitstellt.

gelingen, indem bestehende Lernmaterialien mit Kontextinformationen angereichert werden (vgl. Abschnitt 3.5.3). Zu Beginn der Arbeit wurde in diesem Zusammenhang auch von Hilfestellungen gesprochen, die grundlegend die gleiche Bedeutung besitzen. Wie der Begriff Kontextinformation suggeriert, bedarf es zur Herleitung und Bereitstellung einen sinnvollen situativen oder inhaltlichen Zusammenhang. Die Literatur gibt Beispiele für situative Zusammenhänge in Form von z.B. Tests oder Übungen (vgl. z.B. [122]). Denkbar ist ebenso die Verknüpfung von Hilfestellungen mit einem konkreten Inhalt eines Lernmaterials (vgl. z.B. [102, 121]). Beide Varianten scheinen für die vorliegende Arbeit relevant. Der konkrete inhaltliche Bezug zu einer Hilfestellung ist aufgrund der Komplexität oder auch Größe vieler Lernmaterialien - wie z.B. Skripte - bedeutsam. Übergeordnete Hilfestellungen, wie in einigen Lösungsansätzen gezeigt, scheinen vor allem für komplexe Materialien denkbar ungeeignet. Die Herstellung eines übergeordneten situativen Bezugs - z.B. durch Übungen - scheint ebenso bedeutend, da Hilfestellungen mit inhaltlichem Bezug zwischen verschiedenen Situationen ggf. variieren. So kann eine inhaltliche Hilfestellung für eine Übung angemessen sein, für eine weitere Übung aber ggf. nicht. Gleiches gilt grundlegend auch für die Präsentation neuer Lernmaterialien. Bereits zu Beginn der Arbeit wurde eine Verknüpfung zwischen neuen Lernmaterialien und einem Anknüpfungspunkt im bestehenden Lernkontext gefordert. Die inhaltliche Verknüpfung ermöglicht dabei die direkte Nutzung von inhaltlich gleichen Lernmaterialien, wodurch ein leichter Einstieg im Umgang mit neuen Lernmaterialien für Lernende möglich ist. Der übergeordnete Bezug kann wiederum dazu dienen, Empfehlungen von Materialien für unterschiedliche Situationen zu differenzieren. Sowohl für Hilfestellungen, als auch für die Empfehlung neuer Materialien bedarf es der Auffindung von inhaltlichen Bereichen innerhalb bestehender Materialien, die als Anknüpfungspunkte dienen. Die Anknüpfungspunkte können an dieser Stelle auch als relevante inhaltliche Bereiche innerhalb von Lernmaterialien bezeichnet werden, was einer konkreteren Beschreibung nachkommt und sich nachfolgend als noch treffender erweist. Dies kommt grundlegend der bereits ausgesprochenen Forderung zur Extraktion von inhaltlichen Teilbereichen nach, die über die Vorteile von Learning Objects motiviert wurden. Drei Kernziele seien nachfolgend für die Systementwicklung definiert:

- K1. relevante inhaltliche Bereiche in Lernmaterialien aufdecken und herleiten,
- K2. neue Lernmaterialien finden und mit relevanten Bereichen verknüpfen, um die Suche für Lernende zu erleichtern,
- K3. Hilfestellungen für relevante Bereiche bereitstellen, um die Nutzung bestehender Materialien zu motivieren.

Durch die Literatur hat sich des Weiteren gezeigt, dass Informationen für Lehrende in einem ganzheitlichen Lehr-/Lernszenario von Bedeutung sein können. Die Analyse des Nutzerverhaltens mit Lernmaterialien sowie das Aufdecken von Problemen bei der Nutzung von Lernmaterialien können als wertvolle Informationen für Lehrende herausgestellt werden. Der Begriff „Problem“ ist in diesem Zusammenhang sehr allgemein gehalten, bezieht sich aber vorrangig auf inhaltliche Hürden, die sich z.B. aus einem unangemessen Schwierigkeitsgrad oder ungenauen Beschreibungen ergeben. Auch in diesem Zusammenhang bedarf es der Extraktion von inhaltlichen Bereichen, die entsprechend der/dem Lehrenden zur Verfügung gestellt werden. Neben problematischen Bereichen kann es für eine(n) Lehrende(n) in gleicher Form von Interesse sein, welche Bereiche relevant für Lernende sind. Auch die Einsicht in neue Lernmaterialien sollte für Lehrende zur Verfügung stehen. So ergeben sich zusammenfassend nachfolgende Ziele aus Sicht eines Lehrenden:

- L1. inhaltliche Bereiche in bestehenden Lernmaterialien herleiten/präsentieren, die problematisch für das Lernen sind,
- L2. inhaltliche Bereiche in bestehenden Lernmaterialien herleiten/präsentieren, die relevant für Lernende sind,
- L3. neue Lernmaterialien präsentieren,
- L4. generelles Nutzungsverhalten für Lernmaterialien aufbereiten/präsentieren.

Da das Empfehlungsdienstszenario ausgehend von bestehenden Lernmaterialien konstruiert wird, ist es von besonderer Bedeutung, dass bestehende Materialien in gleicher Form Verwendung finden. Wie bereits argumentiert, sollte die Verwendung ohne eine zusätzliche Aufbereitung der Materialien möglich sein, da Lehrende diese ggf. nicht akzeptieren [112]. Es ergeben sich somit gesonderte Anforderungen oder Rahmenbedingungen, die es zusätzlich zu den genannten Zielen zu berücksichtigen gilt (vgl. auch Abschnitt 1.3):

- L5. bestehende Lernmaterialien von Lehrenden müssen im Empfehlungsdienst nutzbar sein,
- L6. eine gesonderte Aufbereitung von Lernmaterialien z.B. durch den Lehrenden ist nicht erforderlich,
- L7. jegliche Form von Lernmaterialien sollen Berücksichtigung finden (z.B. Text, Audio, Video).

4.2 Konzeption

Zur Umsetzung der genannten Ziele und Anforderungen wird an dieser Stelle ein kollaborativer Tag-Ansatz vorgeschlagen, in dem User jeglichen Inhalt eines Lernmaterials mit Tags (Metadaten) anreichern können. Tags dienen in diesem Zusammenhang als explizites Userfeedback oder auch als Bewertung, wodurch die Lernenden ihre Meinung über einen bestimmten Inhalt kundtun können. Es hat sich gezeigt, dass Tags für die Klassifizierung ganzer Ressourcen – nicht notwendigerweise Lernressourcen – eine probate Methode darstellen [25, 54, 60, 172]. Die Annahme besteht darin, dass der Ansatz der Klassifizierung durch Tags auch für die Klassifizierung von inhaltlichen Bereichen innerhalb von Lernmaterialien herangezogen werden kann. Durch die aktive Teilnahme Lernender besteht des Weiteren die Annahme, dass genügend Informationen existieren, um relevante und ebenso problematische Bereiche in Lernmaterialien aufzudecken. In diesem Zusammenhang ist auch die Nutzung von impliziten Informationen – nachfolgend als implizite Tags bezeichnet – vorgesehen, die sich über das Nutzungsverhalten von Lernern ergeben. Dies ermöglicht die Aussagen von Usern auch in der Anwendungsdomäne E-Learning zu stärken bzw. besser interpretierbar zu machen [81, 149]. Abschnitt 4.3 wird die Formen von expliziten bzw. impliziten Tags noch ausführlich beschreiben.

Auch die Herleitung von neuen Lernmaterialien scheint auf Grundlage einer Gruppe Lernender sinnvoll. In einer Vielzahl an Empfehlungsdiensten ist das „word of mouth“-Prinzip ein etabliertes Konzept zur Weiterempfehlung von Items oder Inhalten innerhalb einer Usergruppe. Dies kann in gleicher Weise für Lernmaterialien adaptiert werden (vgl. z.B. [122]). In der vorliegenden Arbeit besteht die Annahme darin, dass ein Lernmaterial, welches für einen einzelnen Lerner wertvoll erscheint, ggf. auch wertvoll für die übrige Gruppe ist. Auch dieses Vorgehen ist in Abschnitt 4.3 noch ausführlich erläutert.

Durch die aktive Einbeziehung der Lerner wird der Vorteil gesehen, dass eine sinnhaftere inhaltliche Bewertung stattfinden kann, die von rein maschinellen Systemen bzw. rein inhaltsbasierten Verfahren nicht zu leisten ist. Dies ist auch durch die Literatur bestätigt [90], die den Bedarf von menschlichen Bewertungen beschreibt, um z.B. Audio-/Videomaterial evaluieren zu können. Ein weiteres Argument, das die Arbeit in [90] darlegt, ist der alternative Bedarf an Metadaten, die i.d.R. kostenintensiv und ggf. unvollständig sind. Bedingt durch die inhaltliche Komplexität von Lernmaterialien bedarf es allerdings einer fachkundigen Beurteilung, um Zusammenhänge herstellen zu können. Da sich Lernende allerdings innerhalb eines Lernprozesses aktiv mit den Inhalten der Lernmaterialien auseinandersetzen, kann diese Beurteilung auch durch Lernende erfolgen. Zudem verlangt die Zielvorgabe die Herleitung von problematischen Bereichen innerhalb eines Materials, die schlussendlich nur durch einen individuellen Lerner selbst bestimmt werden kann.

Zum Abschluss des Systemkonzeptes sei zusammengefasst, welche Aspekte welche Entwurfsziele mitberücksichtigen.

K1: Durch einen Tag-basierten Ansatz besteht die Möglichkeit durch Bereitstellung geeigneter Tagging Werkzeuge eine inhaltliche Bewertung in Lernmaterialien durchzuführen. Die Aggregation der gesetzten Tags führt dann zu einem Meinungsbild, über das die Herleitung relevanter Bereiche sichergestellt werden kann.

K2: Mittels einer Tag Funktionalität, die die Verlinkung zu alternativen Lernmaterialien ermöglicht, ist eine direkte inhaltliche Verknüpfung zwischen bestehendem und neuem Material möglich. Eine Verknüpfung mit einem relevanten Bereich kann dann erfolgen, sofern das verlinkte Material in einem hergeleiteten relevanten Bereich liegt.

K3: Das Meinungsbild der relevanten Bereiche kann zur Herleitung von Kontextinformationen bzw. Hilfestellungen herangezogen werden. Für Lernende ergibt sich somit ein Überblick wie die übrigen Lernenden den Umgang mit den jeweiligen Bereichen empfunden haben.

L1, L2, L3: Das aggregierte Meinungsbild Lernender dient gleichermaßen für Lehrende zur Präsentation sowohl von relevanten als auch schwierigen Bereichen in Lernmaterialien. Die mit

dem jeweiligen Bereich verknüpften Verlinkungen zu alternativen Materialien können in gleicher Form für Lehrende bereitgestellt werden.

L4: Durch das explizite Tagging Verhalten Lernender ergeben sich bereits verschiedene Informationen – wie z.B. Taghäufigkeit, die das Nutzungsverhalten von Lernenden mitbeschreiben. Über die Berücksichtigung von impliziten Nutzungsinformationen, die sich während der Nutzung mit Lernmaterialien ergeben, – wie z.B. Nutzungsdauer von Lernmaterialien – ergibt sich ein ganzheitliches Nutzungsverhalten für Lernmaterialien.

L5, L6, L7: Der Tag-basierte Ansatz bietet den Vorteil, dass die Metainformation in Form eines Tags lediglich mit einer Stelle im Lernmaterial – als eine Art Koordinate – verknüpft werden muss. Dies ermöglicht es das Konzept auf möglichst alle gängigen Lernmaterialien zu übertragen, wobei es der Definition einer Koordinate für jeden Materialtyp bedarf. Für die Systemimplementation besteht somit die Anforderung der Umsetzung eines Taggingwerkzeuges, welches das Setzen von Tags in unterschiedlichen Medien ermöglicht.

4.3 Systemimplementation

Im nachfolgenden Abschnitt ist die Systemimplementation auf Grundlage von Abschnitt 4.1 und 4.2 beschrieben. Das System teilt sich grundlegend in eine Taggingkomponente und eine Analysekomponente. Die Analysekomponente umfasst dabei die Empfehlungsdienstalgorithmik, welche in Abschnitt 4.4 beschrieben ist. Die Taggingkomponente stellt das Userinterface dar und kann somit als Schnittstelle zwischen User und System gesehen werden, die eine Transaction zwischen User und Item ermöglicht. Um der durchgängigen Betrachtung von User, Item und Transaction zu folgen, seien die Begriffe für die Implementation des Systems LAOs nachfolgend definiert.

LAOs Item: Ein Item repräsentiert im System jegliches zur Verfügung stehende Lernmaterial. Unabhängig dabei ist die Form oder der Typ des Lernmaterials und so finden Textdokumente, Präsentationsfolien und Audio-/Videomaterial Berücksichtigung. Der Begriff Item betrachtet ein Lernmaterial dabei als zusammenhängendes und vollständiges Material. In Abschnitt 4.4 wird

sich noch eine Darstellung von inhaltlichen Teilbereichen in Lernmaterialien ergeben, die für eine generelle Darstellung allerdings noch unberücksichtigt bleibt.

LAOs User: Für einen User im System existieren die zwei Rollen *Lehrende(r)*⁶⁶ und *Lernende(r)*⁶⁷. Der Lehrende kann als Anbieter von Items gesehen werden und übernimmt eine eher passive Rolle im System. Der Lernende erhält durch die Möglichkeit der Nutzung von Lernmaterialien und der Durchführung von Transactions eine weitgehend aktive Rolle. Hierdurch bedingt findet vor allem die Rolle des Lerners nachfolgend Betrachtung.

LAOs Transaction: Das System verwendet explizite und implizite Informationen in Form von Tags. Explizite Tags werden dabei aktiv von einem User (Lernenden) für ein Item getätigt. Die Erhebung von impliziten Tags bedarf grundlegend keiner aktiven Beteiligung eines Users. Sicherlich ist für implizite Informationen, die über ein Nutzungsverhalten abgeleitet werden, ebenfalls eine aktive Beteiligung eines Users notwendig, allerdings ist mit der Aussage vielmehr die aktive Durchführung einer Bewertung in Form eines Tags gemeint, worüber sich ein klarer Unterschied zwischen expliziten und impliziten Tags ergibt.

User und Items werden übergeordnet innerhalb eines Kurses gekapselt. Ein Lehrender ist dabei der Besitzer des Kurses, der seine Materialien im Materialpool des jeweiligen Kurses anbietet. Die Usergemeinde der Lernenden erhält dementsprechend Zugriff auf den Materialpool des jeweiligen Kurses. Jeder Kurs enthält zusätzlich einen Aufgabenpool, der die Herleitung eines Kontextbezugs und die Einschätzung von Usern ermöglicht. Bevor eine weitgehend technische Betrachtungsweise eingenommen wird, illustriert Abbildung 23 zunächst die generelle Implementation des Systems.

⁶⁶ Nachfolgend nur noch als Lehrender bezeichnet.

⁶⁷ Nachfolgend nur noch als Lernender bezeichnet.

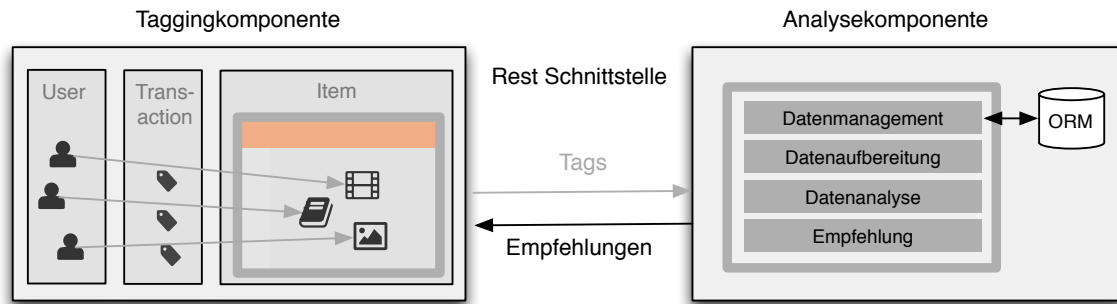


Abbildung 23 System Implementation

Die Implementation sieht eine bewusste Entkopplung der beiden Teilkomponenten vor. Die funktional voneinander getrennten Komponenten ermöglichen eine unabhängige Entwicklung, was neben einer verbesserten Struktur auch mehr Freiheiten bei der Auswahl eingesetzter Technologien ermöglicht. So ist die Taggingkomponente (in Abbildung 23 links) in Form einer Webanwendung implementiert und ermöglicht eine plattformunabhängige Nutzung des Systems. Die Serverkomponente ist als Java-basierte Webanwendung entwickelt, die sowohl das Datenmanagement, als auch die Implementation der Empfehlungsdienstalgorithmik vorsieht (vgl. Abschnitt 4.4). Der Austausch zwischen den Teilkomponenten erfolgt dabei über eine REST-Schnittstelle die Serverseitig implementiert ist. Die implementierte Taggingkomponente ist für ein Textdokument in Abbildung 24 bzw. für ein Videodokument in Abbildung 25 gezeigt. Die Darstellung von Präsentationsfolien bzw. Audiodokumenten ergibt sich analog zu Text- bzw. Videodokumenten. Unabhängig vom Dokumententyp ist der strukturelle Aufbau der Taggingkomponente gleich. Neben der eigentlichen Darstellung des Dokuments enthält die jeweilige Komponente eine Dokumentennavigation sowie eine (Tag)-Werkzeugleiste, innerhalb derer alle im System möglich zu setzenden Tags als Auswahl zur Verfügung stehen. Neben der Werkzeugleiste kann ebenfalls ein Kontextmenu über einen Rechtsklick auf das jeweilige Dokument geöffnet werden (vgl. auch Abbildung 27 auf Seite 116). Des Weiteren existiert eine Tagleiste, die bereits durchgeführte Taggingaktivitäten anzeigt (in Abbildung 24 rechts vom Dokument, in Abbildung 25 unten vom Dokument). Über die Tagleiste erhalten User die Möglichkeit bereits durchgeführte (Tag-)Aktivitäten anzeigen zu lassen und bei Bedarf zu bearbeiten. Bereits gesetzte Tags sind je nach Werkzeug farblich unterschiedlich markiert. Eine dauerhaft sichtbare Informationshilfe gibt Aufschluss über die Farbzueweisung sowie über die Funktionalität der verschiedenen Werkzeuge. Des Weiteren enthält das Informationsfenster einen Link zu einem weiterführenden Hilfeangebot. Der obere Bereich in beiden abgebildeten

Arbeitsbereichen dient der Navigation im System und ermöglicht dem User, zwischen verschiedenen Ebenen (Kursauswahl, Kursübersicht, Aufgabenpool, Materialpool, Arbeitsbereich) des Systems zu wechseln (vgl. hierzu auch Anhang A). Die Entwicklung der Taggingkomponente folgt einem modularen Aufbau, um eine möglichst hohe Flexibilität und Wiederverwertbarkeit im Entwicklungsprozess zu ermöglichen. Lediglich die Darstellung der beiden Arbeitsbereiche und der Zugriff auf die (Text, Video)-Dokumente über die Dokumentennavigation sind angepasst. Die Funktionalität der Werkzeugleisten, des Kontextmenüs, der Tagleisten und sämtliche Funktionalität zur Darstellung von Tags sowie zur Interaktion mit dem Server sind in beiden Arbeitsbereichen identisch. Auf eine detaillierte Beschreibung der technischen Umsetzung von Server- und Taggingkomponente wird an dieser Stelle verzichtet. Anhang A enthält allerdings eine Entwicklungsdokumentation, die detailliert auf die technische Umsetzung beider Komponenten und ebenso detailliert auf die genutzten Technologien eingeht. Anhang A stellt des Weiteren eine Übersicht über alle im System verfügbaren Anwendungsfunktionen und Oberflächen bereit, um einen ganzheitlichen Überblick über das System geben zu können.



Abbildung 24 Taggingkomponente Textdokument

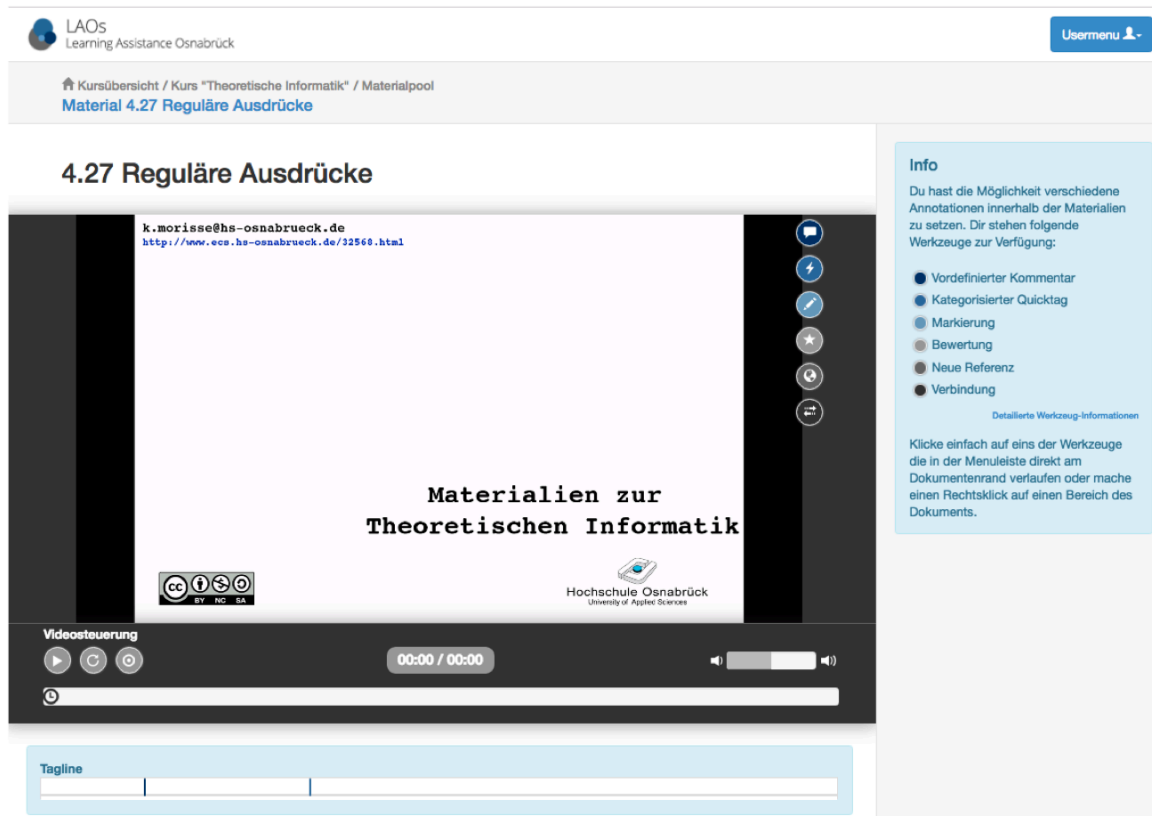


Abbildung 25 Taggingkomponente Videodokument

Das System unterscheidet zwischen expliziten und impliziten Tags. Explizite Tags kommen einer Art Bewertung gleich, durch die User Aussagen zu inhaltlichen Bereichen innerhalb von Lernmaterialien tätigen können. Hierüber soll eine inhaltliche Klassifizierung durch die User ermöglicht werden (vgl. Entwurfsziele K1, K3, L1, L2). In sonst üblichen Social Tagging Systems sind Tags in textueller⁶⁸ Form komplett frei wählbar. Verschiedene Gründe sprechen gegen ein etwaiges Vorgehen. Zum einen sind frei wählbare Tags häufig schwer zu analysieren und zu interpretieren. Dies gilt nicht zwangsläufig für maschinelle Systeme und so nutzen User unterschiedliche Begrifflichkeiten, um einen Sachverhalt zu beschreiben, was selbst für andere User schwer zu interpretieren sein kann [51, 91]. Des Weiteren zeigt die Literatur, dass User eher zur aktiven Mitarbeit angeregt werden, sofern eine vordefinierte Auswahl an Tags existiert [50]. Dies deckt sich mit den Ausführungen aus Abschnitt 3.5.3, wo bereits verschiedene Probleme in diesem Zusammenhang angeführt wurden. Aus diesem Grund ist innerhalb von

⁶⁸ Streng genommen wären natürlich auch numerische oder symbolische Zeichenfolgen möglich.

LAOs eine vordefinierte Menge von Tags vorgesehen. Tabelle 4 präsentiert die verfügbaren Tags geordnet nach expliziter und impliziter Ausprägung.

Tag	Explizit / Implizit	Beschreibung
Vordefinierter Freitext	Explizit	Vordefinierter Textkommentar (9x positiv, 9x negativ)
Kategorisierter Quicktag	Explizit	„Daumen hoch/runter“ Bewertung in den Kategorien Wichtigkeit, Nützlichkeit, Verständnis und Schwierigkeit
Bewertung	Explizit	Generelle Bewertung von 1 bis 5 Sterne
Neues Material	Explizit	Hinzufügen eines neuen Materials z.B. aus dem Internet
Material Verknüpfung	Explizit	Verknüpfung zwischen Materialbereichen für bestehende Materialien
Marker	Explizit	Marker zur Markierung von inhaltlich zusammenhängenden Bereichen
Use Flag	Implizit	Information über die Nutzung eines Materials
Verweildauer	Implizit	Verweildauer innerhalb eines bestimmten Bereiches eines Materials
Seitenwechsel	Implizit	Wechsel zwischen zwei Seiten in einem Textdokument
Zeitsprung	Implizit	Zeitsprung in einem zeitbasierten Material (Audio, Video)
Materialwechsel	Implizit	Wechsel zwischen zwei Materialien

Tabelle 4 Übersicht Tags im System

Durch die verschiedenen expliziten Tag-Typen soll der User die Möglichkeit erhalten, seiner Aussage in unterschiedlicher Form Ausdruck zu verleihen. Dies gilt in erster Linie für die Typen „vordefinierter Freitext“, „kategorisierter Quicktag“ und „Bewertung“, die auf unterschiedlichen Ebenen zur Nutzung anregen sollen. So ermöglicht ein vordefinierter Freitext die Erstellung einer sehr zielgenauen Aussage, wofür neun vordefinierte Statements in jeweils positiver oder negativer Ausprägung zur Verfügung stehen. Eine Aufstellung der Statements kann aus Gründen der Übersicht dem Anhang B entnommen werden. Die Statements sind über das *Learning Object Review Instrument (LORI)* abgeleitet worden, bei dem es sich um ein validiertes Evaluationsschema für die Entwicklung von Learning Objects handelt, die zum Zweck der

Qualitätssicherung in verschiedenen Kategorien zu bewerten sind [5, 89, 164]. Ein Überblick über das Evaluationsschema ist in Abbildung 26.

Item	Brief Description
Content quality	Veracity, accuracy, balanced presentation of ideas, and appropriate level of detail
Learning goal alignment	Alignment among learning goals, activities, assessments, and learner characteristics
Feedback and adaptation	Adaptive content or feedback driven by differential learner input or learner modeling
Motivation	Ability to motivate and interest an identified population of learners
Presentation design	Design of visual and auditory information for enhanced learning and efficient mental processing
Interaction usability	Ease of navigation, predictability of the user interface, and the quality of the interface help features
Accessibility	Design of controls and presentation formats to accommodate disabled and mobile learners
Reusability	Ability to use in varying learning contexts and with learners from different backgrounds
Standards compliance	Adherence to international standards and specifications

Abbildung 26 Übersicht Bewertungskriterien des Evaluationsschemas LORI (entnommen aus [89])

Ganz offensichtlich sind nicht alle Kriterien des LORIs zu berücksichtigen, da die Bewertung und Entwicklung von Learning Objects nicht mit der Bewertung von Lerninhalten durch Lernende gleichzusetzen ist. Allerdings bietet das Evaluationsschema vor allem in den Kategorien *Content Quality*, *Learning Goal Alignment*, *Feedback and Adaption*, *Motivation* und *Presentation Design* wertvolle Ideen zur generellen Bewertung von Lerninhalten, die für die vorliegende Arbeit adaptiert wurden. Um innerhalb des Analyse- bzw. Empfehlungsprozesses aggregierte inhaltliche Aussage über das System treffen zu können, ist jedes Statement einer der Kategorien *Wichtigkeit*, *Nützlichkeit*, *Verständnis* oder *Schwierigkeit* zugeordnet. Die Kategorien ergeben sich ebenfalls durch das LORI-Evaluationsschema. Die Zuordnung zwischen Statement und Kategorie ist ebenfalls dem Anhang B zu entnehmen. Der Marker Tag entspricht grundlegend der Ausführung eines vordefinierten Freitextes, allerdings besteht in diesem Zusammenhang die Möglichkeit, ganze Bereiche in einem statischen oder multimedialen Lernmaterial mit einem vordefinierten Freitext zu verbinden.

Mit dem Quicktag kann eine Aussage für die vier Kategorien *Wichtigkeit*, *Nützlichkeit*, *Verständnis* oder *Schwierigkeit* getroffen werden, was einer generellen Bewertung in den durch das LORI abgeleiteten Kategorien entspricht. Der Quicktag ermöglicht eine weniger konkrete, dafür aber schnelle und zielgerichtete Aussage durch die Bewertung der vier abgeleiteten

Kategorien (vgl. Abbildung 27). Der Bewertungs-Tag dient zur Bewertung von Inhalten, für die ein User keine konkrete Aussage treffen kann oder möchte. Mit dem Tag-Typ „neues Material“ kann ein Inhalt im Lernmaterial mit einem alternativen Material z.B. aus dem Internet verknüpft werden (vgl. Entwurfsziele K2, L3). Eine Materialverknüpfung findet dann Verwendung, wenn ein User zwei Punkte im gleichen oder in unterschiedlichen bestehenden Materialien miteinander verbinden möchte (vgl. ebenfalls Entwurfsziele K2, L3). Die Verbindung ist dabei unabhängig von den Materialtypen möglich. So kann z.B. eine Seite im Skript mit einem Bereich in einem Video miteinander verknüpft werden. In einem zeitbasierten Dokument - z.B. einem Video - besteht des Weiteren die Möglichkeit, einen expliziten Tag für eine frei wählbare (zeitliche) Sequenz zu setzen. Abbildung 27 zeigt beispielhaft die Nutzung eines Quicktags. Die Darstellung übriger Tag-Typen ist aus Gründen der Übersichtlichkeit in Anhang A präsentiert.

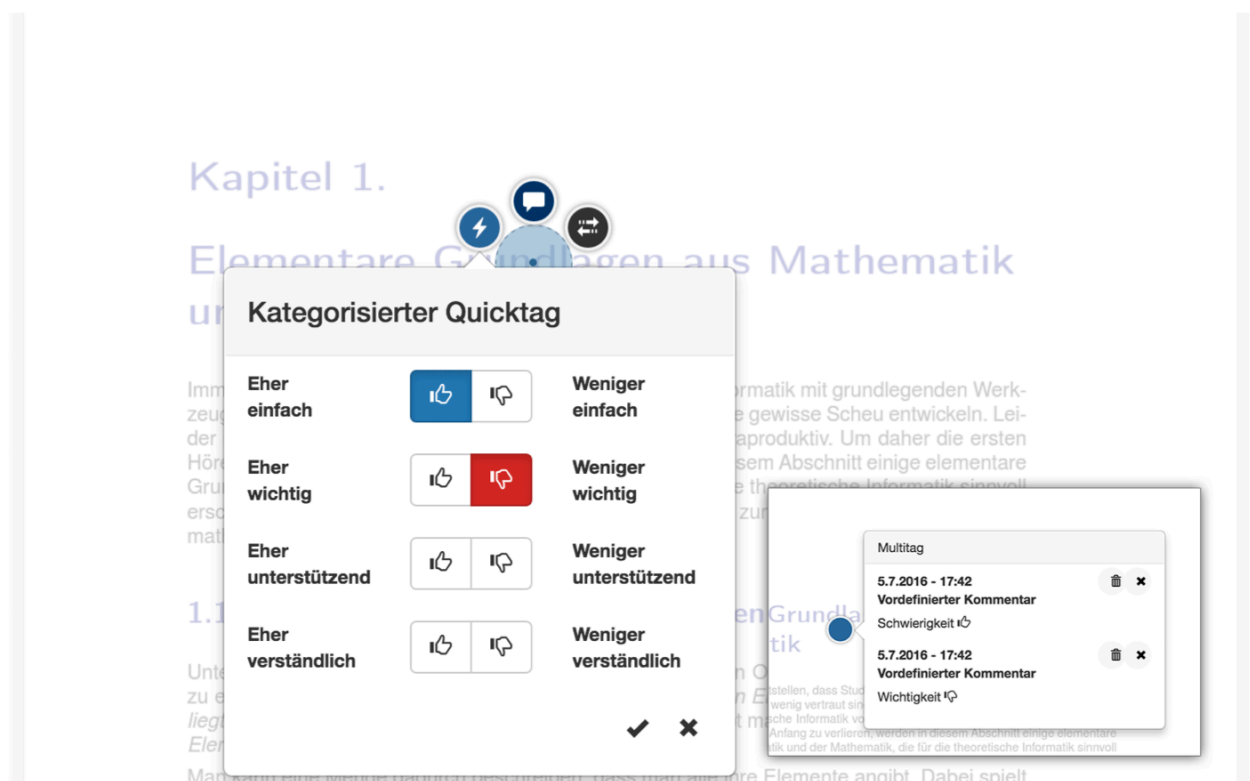


Abbildung 27 Quicktag im System LAOs

Implizite Tags werden wie bereits beschrieben durch das Nutzungsverhalten von Usern abgeleitet (vgl. Entwurfsziel L4). Grundlegend beinhalten die impliziten Tags alle nötigen Informationen, um das Arbeiten mit Materialien für jeden einzelnen der Lerner zu rekonstruieren. Vom Start/Ende der Nutzung eines Materials, über die Navigation innerhalb des

Materials inklusive entsprechender Verweilzeiten ergibt sich ein vollständiger Ablauf. Der Ablauf dient dem Lehrenden eine Einsicht in das Nutzungsverhalten von Lernenden zu erhalten (vgl. hierzu Abschnitt 4.4.4.2). Des Weiteren dienen implizite Tags dazu eine Relevanz für inhaltliche Bereiche zu bestimmen, was als eine Art Qualitätsmaß zu betrachten ist (vgl. Entwurfsziel K1, sowie für weitere Ausführungen Abschnitt 4.4 und Kapitel 8). Auch die Bestimmung von inhaltlichen Verbindungen ist mit impliziten Tags angedacht (vgl. hierzu Kapitel 8).

Aus einer formelleren Perspektive hält jeder explizite Tag einen vordefinierten reellen Wert zwischen +1 und -1 vor, der nachfolgend als *Tagscore* bezeichnet wird. Die Wertigkeit variiert dabei je nach Aussage und Aussagekraft und beschränkt so das Maß an Einfluss für die Klassifizierung von inhaltlichen Bereichen. Ein Wert von +1 deutet auf eine vollständig positive Aussage hin; ein Wert von -1 beschreibt äquivalent eine vollständig negative Aussage. Die Aussagekraft des jeweiligen Tags beeinflusst dabei die Ausprägung des Wertes. So enthalten Tags aus der Gruppe der vordefinierten Freitexte eine deutlich positivere bzw. negativere Aussagen, wodurch sich eine ausgeprägte Wertigkeit ergibt. Als Beispiel erhält ein Tag mit „*Das Material ist an dieser Stelle besonders wichtige für die aktuelle Aufgabe*“ bedingt durch eine klar positive Aussage eine Wertigkeit von +0,9. Ein positiver Quicktag in der Kategorie Wichtigkeit erhält hingegen eine Wertigkeit +0,7, da zwar eine klar positive Aussage erkennbar ist, sich die Ausprägung der Aussage zum Freitext Tag allerdings deutlich unterscheidet. Ein Rating Tag nimmt hingegen maximal eine Wertigkeit zwischen +0,5 (fünf Sterne) und -0,5 (einen Stern) an, da eine klare Aussagekraft nicht erkennbar ist und der Tag so weniger Einfluss auf die Klassifizierung von inhaltlichen Bereichen hat. Die Wertigkeiten sind vollständig vordefiniert und sind in ebenfalls in Abhängigkeit des LORI erhoben worden. Eine vollständige Auflistung aller Wertigkeiten ist Anhang B zu entnehmen. Auch auf Seiten der User bedarf es innerhalb des Systems einer Reglementierung des Einflusses ihrer gesetzten Tags. Hierfür erhalten User im System einen *Userscore*, der eine Wertigkeit zwischen +1 und 0 annimmt. Ein User mit einem höheren Userscore erhält somit höheren Einfluss bei der Klassifizierung von Inhalten. Die Regulierung des Userscores kann z.B. in Abhängigkeit zum Wissensstand oder zum Abschneiden in Tests erfolgen (vgl. z.B. [14]). Auch die Aktivität im System, so z.B. die Anzahl durchgeführter Tags, kann für die Regulierung herangezogen werden. Die Annahme zur

Gewichtung von Tags und Usern besteht darin, dass wichtige Tags von wichtigen Usern einen höheren Einfluss bei der Klassifizierung von inhaltlichen Bereichen erhalten sollen. Dies deckt sich grundlegend mit der Idee aus den Arbeiten zum PageRank- [18] und FolkRank-Algorithmus [65]. Die Gewichtung des Einflusses zeigt in beiden Arbeiten eine verbesserte Bestimmung von bedeutsamen Elementen (Internetseiten, Bookmarks). Dies wird in gleicher Form auch für die Bestimmung relevanter inhaltlicher Bereiche angenommen. Die Idee für die vorliegende Arbeit ist in Abbildung 28 illustriert und zeigt vier unterschiedliche User, die jeweils unterschiedliche Tags in zwei unterschiedlichen Bereichen eines Materials verwendet haben. Die Größe der Userkreise spiegelt dabei die Höhe des Userscores bzw. die Größe der Tagkreise die Ausprägungen der jeweiligen Tagscores wider. Gemäß der Abhängigkeit des User- bzw. Tagscores steigt die Bedeutung eines inhaltlichen Bereiches (r_1 , r_2).

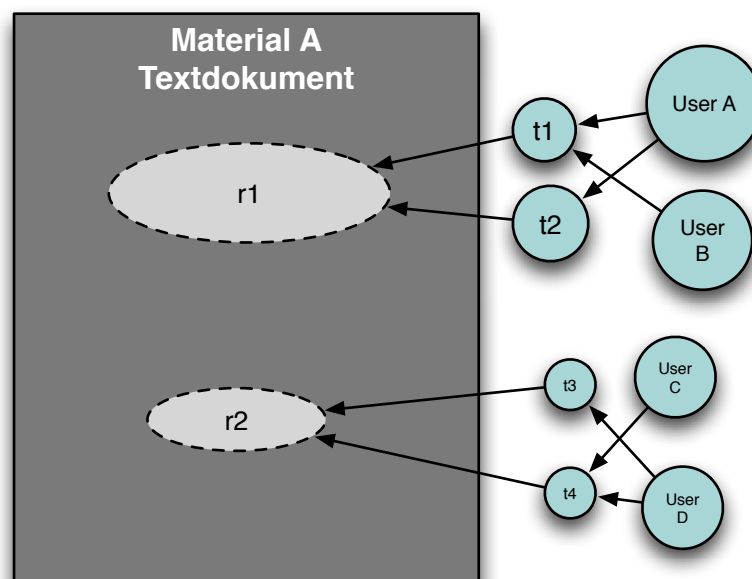


Abbildung 28 Abhängigkeit Relation zwischen User, Tag und Resource

Abbildung 28 illustriert bereits grundlegend die Idee zur Extraktion von relevanten inhaltlichen Bereichen innerhalb eines Materials. Die Idee besteht darin, ein Material (Item) je nach Anhäufung und Bedeutung von Tags in Teilbereiche zu gliedern. Genauer gesagt wird eine Menge von Tags unter bestimmten Bedingungen zu einem Teilbereich zusammengefasst bzw. geclustert. Die sich hieraus ergebenden Teilbereiche werden nachfolgend gemäß der Vorgabe aus Social-Tagging-Systems als Resource bezeichnet. Eine Resource ist somit ein relevanter inhaltlicher Bereich innerhalb eines Lernmaterials. Zum Auffinden dieser Ressourcen bedarf es

für jeden Tag eine Information über die Positionierung innerhalb des Materials. Im System erhält jeder Tag eine *Multimediakoordinate*, die die relative Position des jeweiligen Tags beschreibt. Die Koordinate variiert dabei je nach Typ des Lernmaterials. Für Textdokumente oder Präsentationsfolien wird die Tagposition auf X- und Y-Koordinaten einer Seite oder Folie abgebildet. In zeitbasierten Medien wie Audio- und Videodokumenten ist die Koordinate durch einen Zeitstempel repräsentiert. Die Positionierung von Tags ist in Abbildung 29 für ein Textdokument auf der linken Seite und für ein Video auf der rechten Seite – t entspricht dabei der Zeit – illustriert.

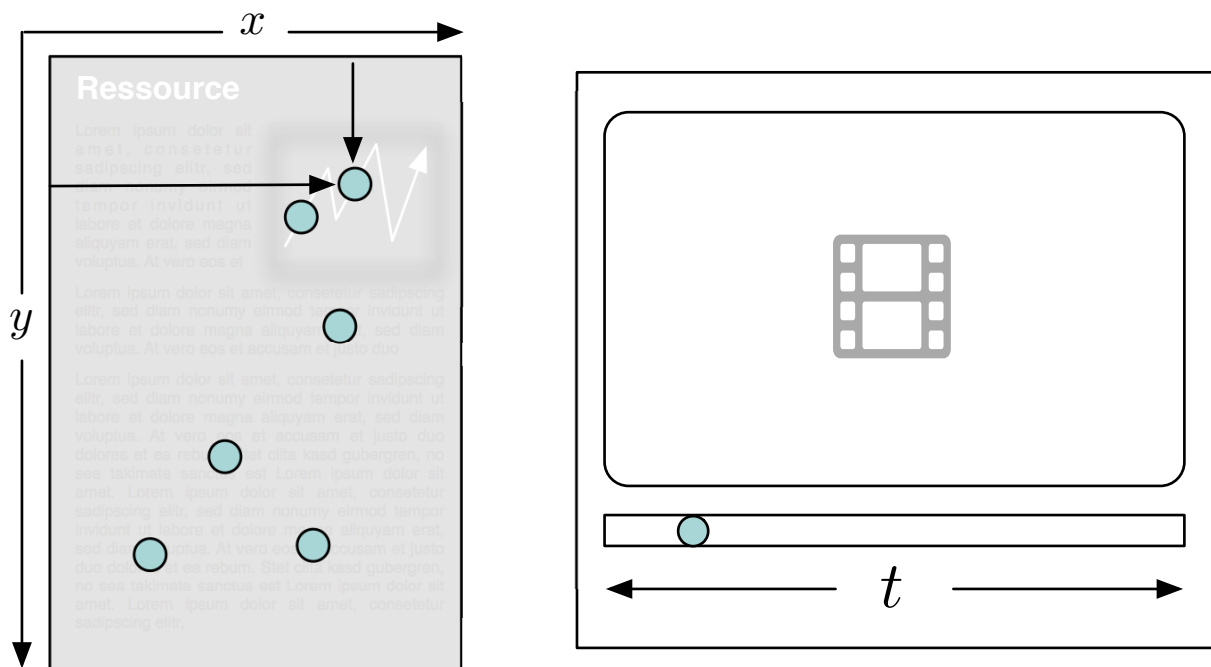


Abbildung 29 Multimediakoordinaten für Text und Video

Auf Grundlage der bis hierher vermittelten Beschreibung des Systems kann im nachfolgenden Abschnitt 4.4 zu einer formalen Darstellung des Empfehlungsprozesses übergegangen werden.

4.4 Empfehlungsdienst Algorithmik

Der nachfolgende Abschnitt beschreibt die Empfehlungsdienstkomponente innerhalb des Systems LAOs. Gemäß Abbildung 23 auf Seite 111 ist diese auch als Analysekompente zu bezeichnen. Die Bezeichnung als Empfehlungsdienstkomponente ist an dieser Stelle ggf. etwas irreführend, da ein ganzheitlicher Empfehlungsdienst grundlegend aus der Kombination

zwischen Userinterface (Taggingkomponente) und Analysekomponente besteht. Da sich die Algorithmik zur Herleitung von Empfehlungen allerdings innerhalb der Analysekomponente befindet, ist die Bezeichnung der serverseitigen Implementation als Empfehlungsdienstkomponente durchaus gerechtfertigt. Die nachfolgende Erläuterung erfolgt auf Grundlage der vermittelten Systembestandteile in Abschnitt 4.3 und gliedert sich gemäß Abbildung 13 auf Seite 37 nach Schritten eines Empfehlungsprozesses in die Unterpunkte Datenaufbereitung, Datenanalyse und Empfehlung. Die Datenaufbereitung beinhaltet dabei die Extraktion oder Clusterung von relevanten inhaltlichen Teilbereichen, die innerhalb der Datenanalyse bewertet werden. Auf Grundlage der Datenanalyse ist dann eine Empfehlung der Bereiche für Lehrende und Lernende möglich. Der nachfolgende Teilabschnitt 4.4.1 definiert zunächst allerdings eine formale Systemdefinition, auf der die Schritte der Datenaufbereitung, Datenanalyse und Empfehlung aufbauen.

4.4.1 Systemdefinition

Dem System LAOs liegt eine Folksonomy (vgl. Abschnitt 3.3.2) als Datenstruktur zugrunde. Wie in Abschnitt 4.3 beschrieben erfolgt die Ermittlung von Ressourcen allerdings erst mittels Extraktion von inhaltlich relevanten Bereichen innerhalb der Datenaufbereitung, die schlussendlich die hier benannten Ressourcen darstellen. Somit erfolgt die Systemdefinition zunächst mit der übergeordneten Menge an Items (Lernmaterialien):

Seien U , T und I endliche Mengen deren Elemente als User (u), Tags (t) und Items (i) bezeichnet werden. Sei $Y_{\Delta} \subseteq U \times T \times I$ eine Relation zwischen den Mengen definiert durch

$$Y_{\Delta} = \{(u, t, i) | u \in U, t \in T, i \in I, \text{User } u \text{ hat Tag } t \text{ Item } i \text{ bis zum Zeitpunkt } \Delta \text{ zugeordnet}\}. \quad (22)$$

Der Einfachheit halber sei angenommen, dass Y_{Δ} immer im aktuellen Systemzustand Betrachtung findet und so die vereinfachte Schreibweise Y genutzt wird. Zur Beschreibung des Userscores eines beliebigen Users u wird die Funktion

$$S_U: U \rightarrow [0,1] \quad (23)$$

verwendet. Des Weiteren wird zur Bestimmung des Tagscores eines beliebigen Tags t die Funktion

$$S_T: T \rightarrow [-1,1] \quad (24)$$

verwendet. Zur Bestimmung der Multimediakoordinate eines beliebigen Tags t , welche die relative Position innerhalb eines Items darstellt, betrachtet man die Funktion

$$K_T: T \rightarrow [0,100]. \quad (25)$$

Für die Extraktion von inhaltlich relevanten Bereichen innerhalb von Items wird fortan der Begriff Resource genutzt. So sei R eine endliche Menge deren Elemente als Ressourcen (r) bezeichnet werden. Für jedes beliebige Element $r \in R$ betrachtet man zwei Koordinatenfunktionen, die mit

$$O_R: R \rightarrow [0,100] \quad (26)$$

eine obere Grenze einer Resource bzw. mit

$$U_R: R \rightarrow [0,100] \quad (27)$$

eine untere Grenze einer Resource bestimmt, wobei gilt $O_R(r) < U_R(r)$. Für ein besseres Verständnis illustriert nachfolgende Abbildung 30 die Systemdefinition.

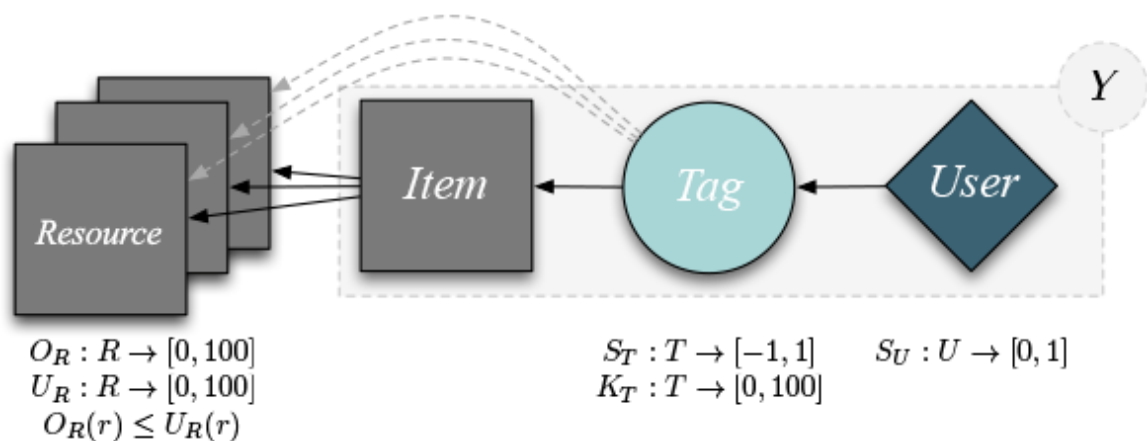


Abbildung 30 Systemdefinition

4.4.2 Datenaufbereitung

Die Datenaufbereitung oder vielmehr die Extraktion von Ressourcen erfolgt für LAOs durch die Clusterung von Datenbeständen bzw. Tags. Genauer gesagt handelt es sich um die Extraktion von inhaltlich relevanten Teilbereichen (Ressourcen) innerhalb von Lernmaterialien (Items), was grundlegend der Clusterung von Tags innerhalb eines Items gleichkommt. Hierzu ist ein Clusteringverfahren implementiert, welches für die Datenbasis der Folksonomy adaptiert wurde. Das implementierte Verfahren stellt ein hierarchisches Verfahren dar, was einer agglomerativen (Bottom-Up) Strategie folgt, so dass Tags zunächst als einzelne Cluster zu sehen sind und sukzessive gemäß einer Fusionierungsvorschrift zu größeren Clustern zusammengefasst werden. Die nachfolgende formale Darstellung des Clusteringverfahrens ist in Abbildung 31 für ein leichteres Verständnis schrittweise illustriert. Auf die Verfahrensschritte der Abbildung wird innerhalb der formalen Darstellung verwiesen.

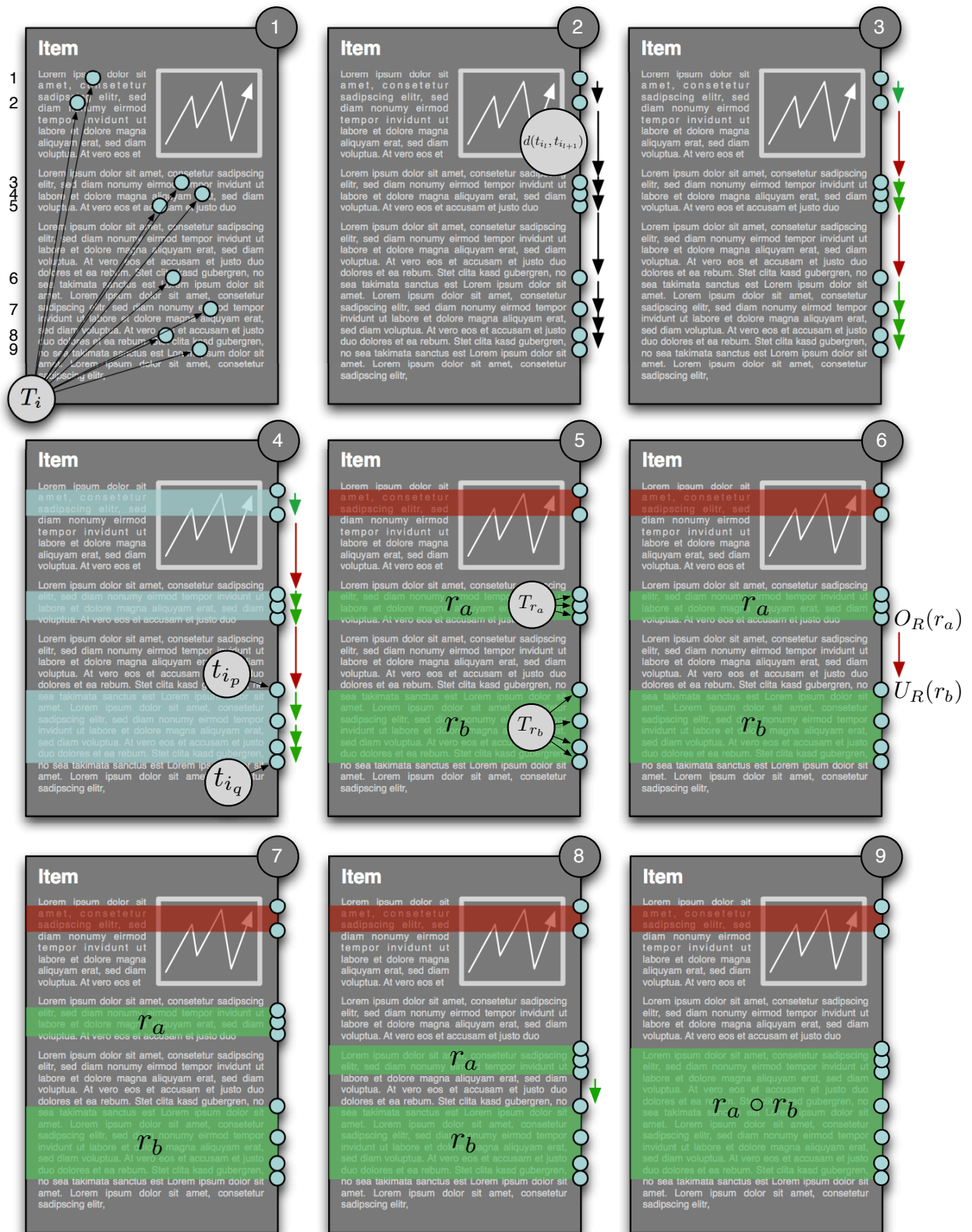


Abbildung 31 Ablauf Clustering Verfahren

Zur Extraktion der Menge R innerhalb eines frei gewählten Items $i \in I$ sei T_i eine Sequenz⁶⁹, die alle Tags die einem Item i zugeordnet wurden beinhaltet und die definiert ist durch (vgl. Abbildung 31 Schritt 1):

$$T_i = (t_{i_1}, \dots, t_{i_{n_i}}), \text{ mit} \quad (28)$$

$$\exists u \in U \text{ und } (u, t_{i_k}, i) \in Y \text{ f\u00fcr jedes } i_k.$$

Dies bedeutet, dass mindestens ein User u einen Tag t_{i_k} Item i zugeordnet hat. Dabei sei n_i die Anzahl der Elemente der Sequenz T_i . F\u00fcr das System LAOs wird innerhalb eines Textdokuments eine Textseite bzw. in Pr\u00e4sentationsfolien eine Folie als Item angesehen, um die Errechnung von Ressourcen auf einem abgegrenzten Bereich durchf\u00fchren zu k\u00f6nnen. Audio-/Videomaterial wird hingegen in vollem Umfang als Item betrachtet. F\u00fcr die Sequenz T_i wird eine nach Tagkoordinaten aufsteigend sortierte Ordnung vorausgesetzt, so dass gilt $K_T(t_{i_1}) \leq K_T(t_{i_2}) \leq \dots \leq K_T(t_{i_{n_i}})$. Sei $d(x, y)$ die Distanzfunktion f\u00fcr zwei beliebige Tags $x, y \in T$ zur Bestimmung des Abstandes zwischen beiden Tags definiert durch (vgl. Abbildung 31 Schritt 2)

$$d(x, y) = \|K_T(x) - K_T(y)\|^2 = \sqrt{K_T(x)^2 - K_T(y)^2}. \quad (29)$$

Sei $\overline{D_i}$ das arithmetische Mittel aller Distanzen aneinander grenzender Tags der Sequenz T_i definiert durch

$$\overline{D_i} = \frac{\sum_{l=1}^{n_i-1} d(t_{i_l}, t_{i_{l+1}})}{n_i}. \quad (30)$$

Die Clusterung von Ressourcen f\u00fcr ein Item i erfolgt \u00fcber die sukzessive Clusterung von benachbarten Tags der Sequenz T_i , wobei $t_{i_l}, t_{i_{l+1}}$ als benachbart gelten, sofern die Fusionierungsvorschrift

$$d(t_{i_l}, t_{i_{l+1}}) \leq \overline{D_i} \quad (31)$$

⁶⁹ Eine Sequenz ergibt sich aus der festen Anzahl an Tags innerhalb des Systems, weshalb ein Tag durch einen User mehrfach einem Item zugeordnet und demnach mehrfach in der Sequenz auftreten k\u00f6nnte.

gültig ist (vgl. Abbildung 31 Schritt 3). Die Anzahl n_i der Tags zu einem Item i stellt eine Obergrenze für die Anzahl der Ressourcen für dieses Item dar, d.h. es gilt stets $|R| \leq n_i$ für jede Ressourcenmenge R . Durch die Clusterung von Tags ergibt sich für jede Resource $r \in R$ eine Teilsequenz T_r von Tags der Sequenz T_i definiert durch (vgl. Abbildung 31 Schritt 4 und 5)

$$T_r = (t_{i_p}, \dots, t_{i_q}), \text{ mit } 1 \leq p, q < n_i \text{ und } p < q. \quad (32)$$

Eine Resource wird somit aus solchen benachbarten Tags extrahiert, deren Distanz durchweg kleiner oder gleich der mittleren Distanz einer Sequenz an Tags eines Item ist, z.B.

$$(d(t_{i_p}, t_{i_{p+1}}) \leq \overline{D_i}) \wedge \dots \wedge (d(t_{i_{q-1}}, t_{i_q}) \leq \overline{D_i}). \quad (33)$$

Des Weiteren muss für jede Resource $r \in R$ die Bedingung

$$(q - p) + 1 \geq \frac{n_i}{|R|} \quad (34)$$

gelten, um ein Minimum an Tags für jede Resource sicherzustellen. Dies entspricht der *Relevanz*⁷⁰ die in Abschnitt 4.3 benannt wurde. Die Ober- bzw. Untergrenze einer Resource ergibt sich dabei aus der Koordinate des ersten bzw. letzten Elements der Teilsequenz, d.h. mit $O_R(r) = K_T(t_{i_p}), U_R(r) = K_T(t_{i_q})$ (vgl. Abbildung 31 Schritt 6). Um die Extraktion von inhaltlich relevanten Bereichen bzw. Ressourcen zu verfeinern, können Ressourcen ihrerseits zusammengeführt werden, was wiederum der Zusammenführung zweier Teilsequenzen an Tags gleichkommt. Hierfür findet die Distanz zwischen zwei Ressourcen Betrachtung, die auf einen genügend kleinen Abstand zwischen den Ressourcen zu prüfen ist, wofür das arithmetische Mittel der Distanzen zwischen allen Ressourcen definiert wird durch

$$\overline{D_R} = \frac{\sum_{m=1}^{|R|-1} (U_R(r_{m+1}) - O_R(r_m))}{|R|}. \quad (35)$$

⁷⁰ Auch die Beurteilung der Relevanz über die implizite Information der Verweildauer z.B. über den Abgleich der durchschnittliche Verweildauer auf allen Ressourcen mit $\emptyset \text{Verweildauer}(R) < \text{Verweildauer}(r)$ ist vorgesehen. Eine schlüssige Annahme für eine etwaige Bedingung ist zum jetzigen Zeitpunkt aufgrund fehlender empirischer Daten oder Hinweisen aus der Literatur nicht möglich. Es bedarf zur Implementierung einer Beurteilung auf dem realen Verhalten Lernender, weshalb die Verwendung eines zweiten Relevanzfaktors in Kapitel 8 diskutiert wird.

Die Zusammenführung zweier Ressourcen erfolgt über die Zusammenführung der beiden Teilsequenzen an Tags. Hierfür sei ein Konkatenationsoperator \circ für zwei beliebige Ressourcen $r_a, r_b \in R$ und deren jeweilige Teilsequenz an Tags $T_{r_a} = (t_{j_1}, \dots, t_{j_m})$ bzw. $T_{r_b} = (t_{k_1}, \dots, t_{k_n})$ definiert durch (vgl. Abbildung 31 Schritt 9)

$$r_a \circ r_b = (t_{j_1}, \dots, t_{j_m}, t_{k_1}, \dots, t_{k_n}). \quad (36)$$

Die Zusammenführung zweier Ressourcen kann erfolgen, sofern diese unter der Bedingung $U_R(r_b) - O_R(r_a) \leq \overline{D_R}$ (vgl. Abbildung 31 Schritt 7 und 8) als benachbart gelten.

4.4.3 Datenanalyse

Innerhalb der Datenanalyse erfolgt eine Bewertung der extrahierten Ressourcen der Menge R innerhalb eines Items $i \in I$. Zur Bewertung findet eine Matrixdarstellung Verwendung, um die Relation zwischen einem Element $r \in R$ und den Mengen U und T zu beschreiben. So sei r eine beliebige Resource der Menge R und A_r die zugehörige Relationsmatrix definiert durch

$$A_r = (a_{r,m,n}) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,|U|} \\ \vdots & \ddots & \vdots \\ a_{|T|,1} & \cdots & a_{|T|,|U|} \end{pmatrix}, \text{ mit} \quad (37)$$

$$m, n \in \mathbb{N}, 1 \leq m \leq |T|, 1 \leq n \leq |U|, \text{ wobei}$$

$$a_{r,m,n} = |\{(u_n, t_m, i) \in Y \mid \exists i \in I \text{ und } r \text{ ist Resource in } i\}|$$

die Anzahl an Tags t_m , die durch einen User u_n einer Resource r zugeordnet wurden. Für Ressourcen $r \in R$ wird nun eine Gewichtungsfunktion

$$S_R: R \rightarrow [-1,1] \quad (38)$$

definiert durch

$$S_R(r) = \frac{\sum_{i=1}^{|T|} \sum_{j=1}^{|U|} (S_T(t_i) * S_U(u_j)) * a_{i,j}}{\sum_{i=1}^{|T|} \sum_{j=1}^{|U|} S_U(u_j) * a_{i,j}}. \quad (39)$$

$S_R(r)$ wird als Ressourcenscore bezeichnet und fasst die Gesamtheit der Bewertungen durch Tags aller User für diese Resource zusammen. Die Errechnung des Ressourcenscores bildet die

Einflussnahme von User- bzw. Tagscore zur Bestimmung einer Relevanz mit ab. Die Funktion S_R berücksichtigt somit die bestehenden User-/Tag-Wertigkeiten und ermittelt je nach Einflussnahme einen ausgeprägten Ressourcenscore, der als Grundlage für die nachfolgende Herleitung von Empfehlungen dient. Der Tagscore S_T für einen beliebigen Tag $t \in T$ ergibt sich aus einer vordefinierten Konfiguration gemäß Anhang B. Ein schlüssiges Modell zur Bestimmung des Userscores kann an dieser Stelle aufgrund fehlender empirischer Daten oder schlüssiger Annahmen aus der Literatur nicht präsentiert werden. Der Userscore wird deshalb für einen beliebigen User $u \in U$ mit $S_U(u) = 1$ festgelegt, so dass eine Einflussnahme durch den User zunächst keine Berücksichtigung findet. Auf Grundlage der Evaluation in Kapitel 5 und 6 kann allerdings eine Annahme in Kapitel 8 hergeleitet und ein Modell zur Diskussion gestellt werden.

4.4.4 Empfehlung

Die Herleitung von Empfehlungen ergibt für die Usergruppen *Lernende* und *Lehrende*, weshalb der vorliegende Abschnitt die Herleitung getrennt voneinander betrachtet. Es handelt sich bei Empfehlungen für Lehrenden vielmehr um die Aggregation von Lernaktivitäten der Gruppe der Lernenden. Der auf der Datenanalyse aufbauende Schritt zur Herleitung von Empfehlungen erfolgt deshalb in erster Linie für Lernende in Abschnitt 4.4.4.1.

4.4.4.1 Lernende

Um eine bessere Vermittlung für die Herleitung von Empfehlungen innerhalb des Systems zu ermöglichen, ist eine Systemempfehlung in Abbildung 32 gezeigt. Wie in der Abbildung zu sehen ist, erfolgt der Zugriff über einen gesonderten Empfehlungstag innerhalb des jeweiligen Lernmaterials (in der Abbildung rot unterlegt). Beim Aufruf des Tags wird der inhaltliche Bereich, auf den sich die Empfehlung bezieht, rot gefärbt.

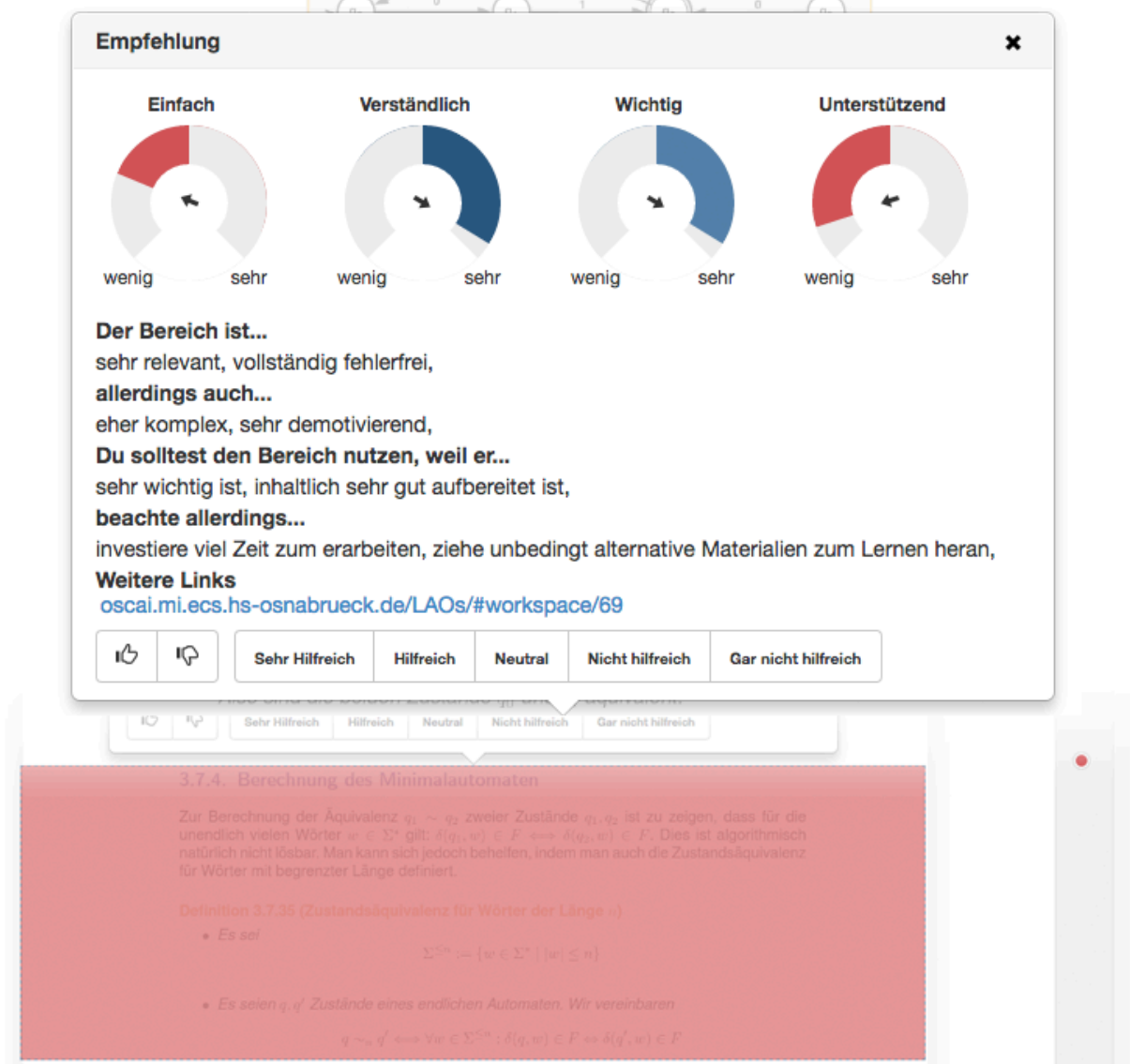


Abbildung 32 Empfehlung in LAOs

Wie Abbildung 33 zeigt, ergeben sich für eine Empfehlung fünf strukturelle Elemente, wobei der Bereich der Bewertung als Evaluationsinstrument für Lernende dient und keiner Herleitung durch das System bedarf (vgl. auch Abschnitt 5.3). Durch die *grafische Aussage* wird den Lernenden eine aggregierte Einschätzung in den Kategorien Schwierigkeit, Verständnis, Wichtigkeit und Unterstützung dargelegt. Die folgende *textuelle Aussage* verspricht die grafische Darstellung, um diese für Lernende besser interpretierbar zu machen. In Abhängigkeit der grafischen bzw. textuellen Aussagen präsentiert das System eine *Handlungsempfehlung* zum Lernen mit dem jeweiligen inhaltlichen Bereich. Die drei Elemente der grafischen und textuellen Aussage sowie der Handlungsempfehlung können als die Hilfestellung innerhalb der

Empfehlung gesehen werden. Innerhalb des Elements *alternative Lernmaterialien* erhalten Lernende Verknüpfungen zu weiterführenden Materialien, die entweder Verknüpfungen innerhalb des Systems oder externe Verknüpfungen z.B. aus dem Internet darstellen.



Abbildung 33 Struktureller Aufbau einer Empfehlung

Die Herleitung von Empfehlungen erfolgt innerhalb des Systems in fünf Schritten, die sich grundlegend am strukturellen Aufbau einer Empfehlung in Abbildung 33 ableiten lassen. Vor der Herleitung erfolgt zunächst eine Prüfung für eine Resource auf Grundlage der Funktion S_R , ob die Herleitung einer Empfehlung für die Resource sinnvoll erscheint oder nicht. Sofern das System die Herleitung als sinnvoll erachtet, erfolgt die Ermittlung der grafischen Aussagen auf Grundlage einer Neuberechnung des Ressourcenscores hinsichtlich der Kategorien Einfachheit, Verständnis, Wichtigkeit und Unterstützung. Hierauf aufbauend leitet das System textuelle Aussagen ab. Anschließend ist die Herleitung einer Handlungsempfehlung möglich. Abschließend verknüpft das System die Empfehlung mit alternativen Lernmaterialien, die durch die Lernenden mit in das System aufgenommen wurden. Der detailliert Ablauf der Herleitung ist nachfolgend beschrieben und zum besseren Verständnis als Pseudocode in Quellcode 1 abgebildet.

Eingabe:

r als beliebige Resource $\in R$

T_r als Tagsequenz von r

$S_R(r)$ als übergeordneter Ressourcenscore von r

Kategorien Einfachheit, Verständnis, Wichtigkeit, Unterstützung und alternatives Material

$SE_R(r)$, $SV_R(r)$, $SW_R(r)$, $SU_R(r)$ als kategorieabhängige Ressourcenscores von r

Ausgabe:

Empfehlung bestehend aus Aussagen, Handlungsempfehlungen (inkl. Ausprägung) und alternativem Material für r

if $S_R(r) > 0,1$ **or** $S_R(r) < -0,1$ **then**

for all T_r t **do**

if t ist aus Kategorie Einfachheit **and** $S_T(t)$ entspricht Vorzeichen von $SE_R(r)$ **then**

if Aussage, Handlungsempfehlung von t in Empfehlung enthalten **then**

increment Hitcount von Aussage, Handlungsempfehlung

else

add Aussage, Handlungsempfehlung zu Empfehlung

and set Ausprägungen von Aussage, Handlungsempfehlung gemäß Wertigkeit von $SE_R(r)$

endif

elseif t ist aus Kategorie Verständnis **and** $S_T(t)$ entspricht Vorzeichen von $SV_R(r)$ **then**

if Aussage, Handlungsempfehlung von t in Empfehlung enthalten **then**

increment Hitcount von Aussage, Handlungsempfehlung

else

add Aussage, Handlungsempfehlung zu Empfehlung

and set Ausprägungen von Aussage, Handlungsempfehlung gemäß Wertigkeit von $SV_R(r)$

endif

elseif t ist aus Kategorie Wichtigkeit **and** $S_T(t)$ entspricht Vorzeichen von $SW_R(r)$ **then**

if Aussage, Handlungsempfehlung von t in Empfehlung enthalten **then**

increment Hitcount von Aussage, Handlungsempfehlung

else

add Aussage, Handlungsempfehlung zu Empfehlung

and set Ausprägungen von Aussage, Handlungsempfehlung gemäß Wertigkeit von $SW_R(r)$

endif

elseif t ist aus Kategorie Unterstützung **and** $S_T(t)$ entspricht Vorzeichen von $SU_R(r)$ **then**

if Aussage, Handlungsempfehlung von t in Empfehlung enthalten **then**

increment Hitcount von Aussage, Handlungsempfehlung

else

add Aussage, Handlungsempfehlung zu Empfehlung

and set Ausprägungen von Aussage, Handlungsempfehlung gemäß Wertigkeit von $SU_R(r)$

endif

elseif t enthält ein alternatives Material **then**

if alternatives Material von t ist in Empfehlung enthalten **then**

increment Hitcount von alternativem Material

else

add alternatives Material zu Empfehlung

endif

endif

endfor

endif

Quellcode 1 Herleitung von Empfehlungen

Die Funktion $S_R(r)$ beschreibt für eine beliebige Resource $r \in R$ einen Ressourcenscore im Wertebereich $[-1,1]$ (vgl. Abschnitt 4.4.3). Die Ausprägung der Wertigkeit gibt Aufschluss über die Ausprägung der positiven bzw. negativen Bewertung durch die Usergruppe. Sofern die Bedingung $-0,1 < S_R(r)$ oder $S_R(r) > +0,1$ erfüllt ist, erwägt das System die Herleitung einer Empfehlung für die Resource r . Der Grund hierfür ist die Annahme, dass eine nicht klar zu deutende Ausprägung in positiver oder negativer Hinsicht zu einer ungenauen Herleitung von Hilfestellungen führen würde. Für die Ermittlung der grafischen Aussagen seien nachfolgend vier Funktionen betrachtet

$$SE_R: R \rightarrow [-1,1], \quad (40)$$

$$SV_R: R \rightarrow [-1,1], \quad (41)$$

$$SW_R: R \rightarrow [-1,1], \quad (42)$$

$$SU_R: R \rightarrow [-1,1], \quad (43)$$

die eine kategorieabhängige Ermittlung des Ressourcenscores vorsehen (SE_R für Einfachheit, SV_R für Verständnis, SW_R für Wichtigkeit, sowie SU_R für Unterstützung). Die Definition ergibt sich analog zur Definition des übergeordneten Ressourcenscores aus Abschnitt 4.4.3, wobei die Errechnung lediglich auf solchen Tags erfolgt, die der jeweiligen Kategorie zugeordnet sind (vgl. hierzu Abschnitt 4.3 bzw. Anhang B). Die grafische Darstellung innerhalb der Hilfestellung ergibt sich über die Darstellung des Ressourcenscores der jeweiligen Kategorie.

Jedes Tag im System ist mit einer vordefinierten Systemaussage bzw. mit einer vordefinierten Systemhandlungsempfehlung, die jeweils mit der Systemaussage zusammenhängt, verknüpft (vgl. Anhang B). Hilfestellungen beruhen grundlegend auf Aussagen bzw. Handlungsempfehlungen solcher Tags, die innerhalb einer jeweiligen Resource vorhanden sind. Die Herleitung textueller Aussagen kann als Verknüpfung von Systemaussagen gesehen werden. Jede vordefinierte Systemaussage kann dabei vier (2x für positive Tags, 2x für negative Tags) vordefinierten Ausprägungen⁷¹ folgen, die in Abhängigkeit des jeweiligen kategorieabhängigen

⁷¹ Eine Erweiterung bzw. Verringerung der Anzahl an Ausprägungen ist ebenfalls möglich.

Resourcenscores variieren. Als Beispiel könnte eine Systemaussage die Ausprägung schwer oder sehr schwer in Abhängigkeit des kategorieabhängigen Resourcenscores $SE_R(r)$ annehmen. Für die Auswahl der Ausprägung ergibt sich eine Fallunterscheidung in Abhängigkeit vordefinierter Grenzen⁷² (vgl. Gleichung 44 als Beispiel für die Kategorie Einfachheit). Die Fallunterscheidung ist äquivalent für die übrigen Kategorien Verständnis, Wichtigkeit und Unterstützung auf Grundlage des zugehörigen Resourcenscores durchzuführen.

$$Ausprägung = \begin{cases} Ausprägung A, falls & 1 \geq SE_R(r) > 0,4 \\ Ausprägung B, falls & 0,4 \geq SE_R(r) \geq 0,1 \\ Ausprägung C, falls & -0,1 \geq SE_R(r) \geq -0,4 \\ Ausprägung D, falls & -0,4 > SE_R(r) \geq -1 \end{cases} \quad (44)$$

Die Umsetzung ist an die Theorie der Fuzzy-Logik angelehnt, ohne allerdings eine sachgemäße Umsetzung der Theorie z.B. unter Berücksichtigung von Fuzzy-Sets zu befolgen. Vielmehr wurde sich der Idee bedient, einen numerischen Wert in den allgemeinen Sprachgebrauch zu übersetzen und einen Übergang zwischen den verschiedenen Ausprägungen zu schaffen. Des Weiteren wird das Aufkommen eines Tags innerhalb einer Resource gezählt, um eine Sortierung der Aussagen nach Häufigkeit und hierüber eine Darstellung des Meinungsbildes zu ermöglichen. Gemäß der Literatur ist die Anhäufung von Tags innerhalb einer Resource nachfolgend als Hit-Count bezeichnet [106]. Abbildung 34 stellt die Verkettung von Hit-Counts, Ausprägungen und Aussagen grafisch dar.

Hit-Count	Ausprägung (Textuelle Aussage)	Textuelle Aussage	Hit-Count	Ausprägung (Textuelle Aussage)	Textuelle Aussage	...
-----------	-----------------------------------	-------------------	-----------	-----------------------------------	-------------------	-----

Abbildung 34 Herleitung textuelle Aussagen

Die Herleitung von Handlungsempfehlungen erfolgt grundlegend analog zur Herleitung von textuellen Aussagen. Jede textuelle Aussage innerhalb der Tagdefinition sieht eine entsprechende Handlungsempfehlung vor, die in Abhängigkeit des jeweils ermittelten kategorischen Resourcenscores vier Ausprägungen folgen kann (vgl. ebenfalls Anhang B). Es entsteht wiederum eine Verkettung - in diesem Fall von Handlungsempfehlungen - die

⁷² Die Grenzen sind anpassbar und gilt es in Abhängigkeit der Evaluationsergebnisse in Kapitel 8 zu diskutieren.

ausgehend von den zuvor ermittelten Aussagen angeordnet sind. Wie bereits beschrieben, ist eine inhaltliche Abhängigkeit zwischen Aussage und Handlungsempfehlung gegeben, so dass Aussage und Handlungsempfehlung lediglich in Kombination auftreten können. Die Herleitung von Handlungsempfehlungen ist in Abbildung 35 illustriert.

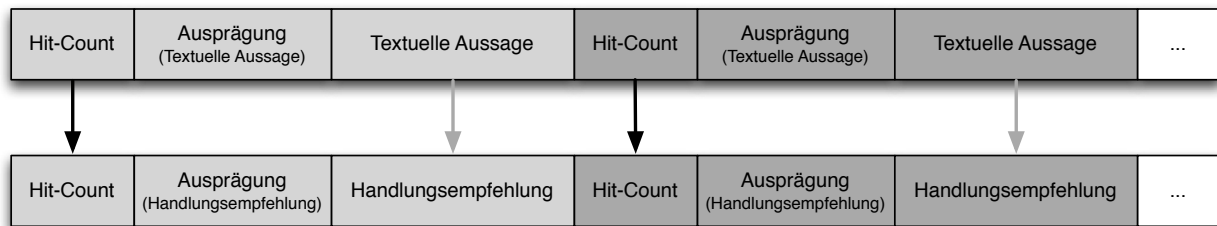


Abbildung 35 Herleitung Handlungsempfehlung

Wie in der Darstellung von Empfehlung in Abbildung 32 oder Abbildung 33 zu sehen ist, sind positive und negative Aussagen bzw. Handlungsempfehlungen getrennt voneinander dargestellt. Zur Herleitung werden allerdings nur solche Tags herangezogen, die mit ihrer Tagwertigkeit $S_T(t)$ einer positiven bzw. negativen Ausführung des jeweilig ermittelten kategorisierten Ressourcenscores entsprechen. Die Ableitung einer zugleich positiven und negativen Aussagen bzw. Handlungsempfehlung aus einer Kategorie kann somit nicht erfolgen.

Die Herleitung von alternativen Lernmaterialien ergibt sich aus den Materialverknüpfungen bzw. den Neuen Materialien, die mittels entsprechendem Tag innerhalb der Resource verlinkt wurden. Die Darstellung erfolgt dabei unabhängig von positiver oder negativer Ausprägung des Ressourcenscores und ebenso unter Vernachlässigung eines kategorisierten Ressourcenscores. Die Annahme besteht darin, dass die Darstellung von alternativen Lernmaterialien in jedem Fall von Nutzen ist, da hierbei vor allem der Prozess des Auffindens von alternativen Materialien vereinfacht werden soll. Vielmehr bedarf es der übergeordneten Reglementierung zur Anzeige von Empfehlungen. Dies ist in Kapitel 8 diskutiert. Die Abbildung eines alternativen Materials erfolgt über eine Beschreibung die durch die Lernenden bei der Verlinkung oder Verknüpfung mit angegeben wurden. Bei der Auswahl des alternativen Lernmaterials erfolgt eine direkte Weiterleitung an die geeignete Stelle des alternativen Materials.

4.4.4.2 Lehrende

Aus Sicht des Lehrenden stellt das System LAOs eine aggregierte Übersicht über die Aktivitäten Lernender zur Verfügung. Nachfolgende Funktionen sind innerhalb des Systems für Lehrende vorgesehen:

- Gesamtübersicht über die Nutzungsdauer für alle verfügbaren Materialien
- Gesamtübersicht über das Tagaufkommen für alle verfügbaren Materialien
- Detailansicht über die Nutzungsdauer innerhalb eines Materials (z.B. je Seite)
- Detailansicht über das Tagaufkommen innerhalb eines Materials (z.B. je Seite)
- Materialansicht mit allen gesetzten Tags und hergeleiteten Empfehlungen
- Nutzungsaktivitäten eines Users im System in Form von Nutzungszeiten und Tagaktivitäten aufgeschlüsselt nach Materialien

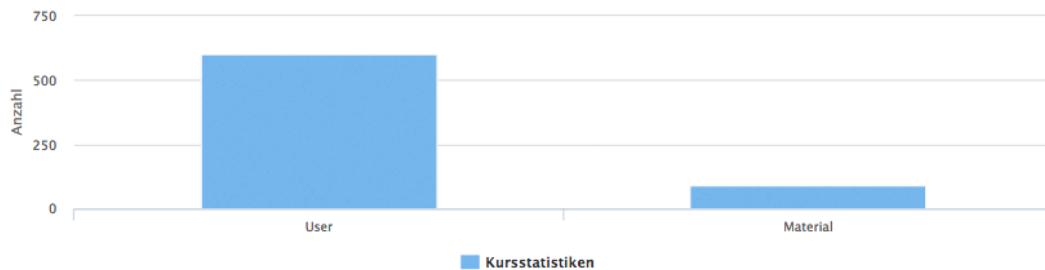
Eine Übersicht wird kursbezogen präsentiert, so dass sich die aggregierten Aktivitäten für Lernmaterialien auf den jeweiligen Materialpool des Kurses beziehen. Neben der Übersicht über Anzahl der Kursteilnehmern, sowie Anzahl der Kursmaterialien, erhält eine Lehrende/ein Lehrender Einsicht in die aggregierte Nutzungsdauer und in das aggregierte (explizite, implizite) Tagaufkommen je Lernmaterial. Ein Lehrender hat somit die Möglichkeit, Lernmaterialien mit hoher oder niedriger Nutzungsaktivität auszumachen. Für eine Übersicht der Analysetools für Lehrende vgl. Abbildung 36.

Admin Analyses Tools

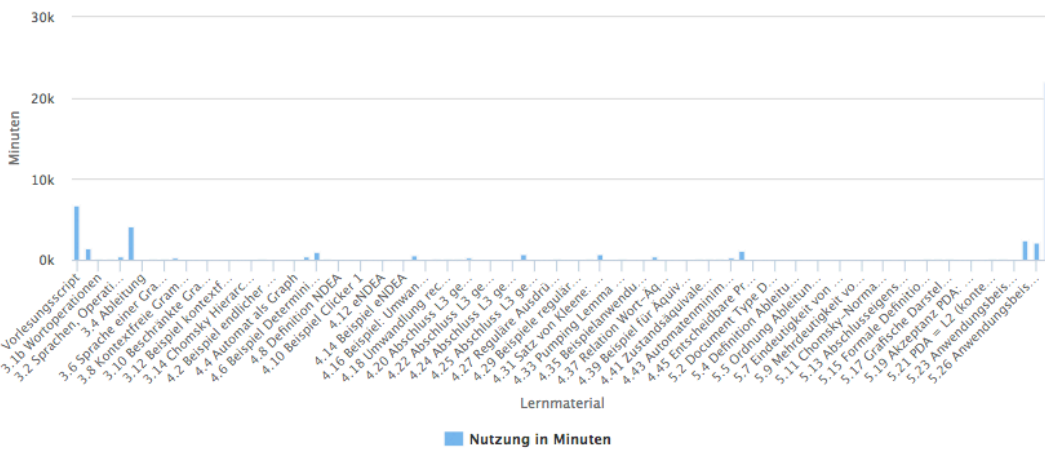
- Benutzeranalyse - - Materialanalyse -

Echtzeitdaten

Kursstatistiken



Materialnutzung



Tagaufkommen

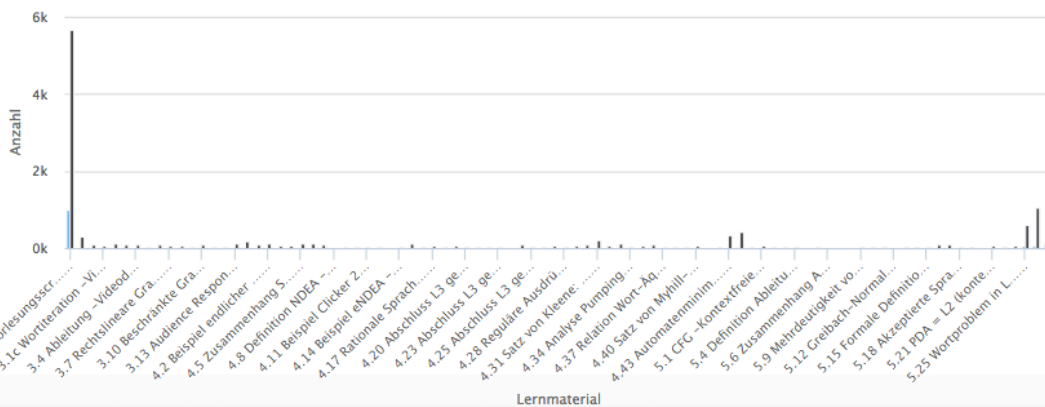


Abbildung 36 Übersicht Lehrender Analysetools⁷³

⁷³ Zur Darstellung aller Diagramme innerhalb des Systems wurde die Javascript Bibliothek Highcharts verwendet (<http://www.highcharts.com/>).

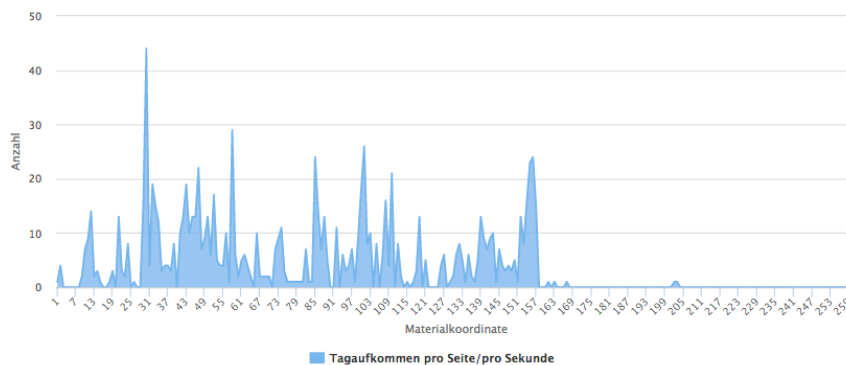
Des Weiteren erhalten Lehrende eine Detailübersicht für jedes Lernmaterial, welches die Nutzungsdauer sowie das Tagaufkommen je Seite (für statische Text-/Präsentationsdokumente) bzw. je Sekunde (für zeitbasierte Audio-/Videodokumente) darlegt. So ergibt sich für Lehrende die Möglichkeit, Bereiche in Lernmaterialien auszumachen, in denen die Usergruppe eine hohe oder niedrige Aktivität zeigt. Eine entsprechende Detailübersicht ist für ein Textdokument in Abbildung 37 gezeigt.

Admin Analyses Tools

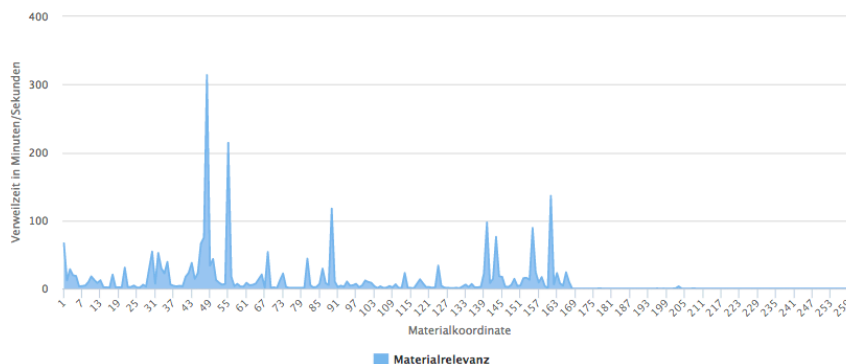
- Benutzeranalyse - *->Vorlesungsscript

Echtzeitdaten

Tagaufkommen in Minuten/Sekunden (Appearance)



Verweilzeit in Minuten/Sekunden (Relevanz)



Material Analytics

Strikte Relevanz

Start Recommender Process

Abbildung 37 Lehrender Detaillierte Materialanalyse

In Abbildung 37 ist des Weiteren die Möglichkeit gezeigt, über den Button *Material Analytics* (rot im Bild) das jeweilige Material zu öffnen und alle vorhandenen Tags der Usergruppe sowie alle vorhandenen vom System generierten Empfehlungen einzusehen (vgl. Abbildung 38). Ausschließlich die jeweilige Lehrperson ist berechtigt, den Empfehlungsprozess für ein jeweiliges Lernmaterial anzustoßen und entscheidet dementsprechend über die Freigabe von Empfehlungen für Lernende (vgl. Abbildung 37 blauer Button).

Beispiel 3.7.36 Es sei $A = (K, \Sigma, \delta, s_0, F)$ der Automat aus dem Beispiel 3.7.22, der in Abbildung 3.30 dargestellt ist.

\sim_0 Zwei Zustände p und q sind bezüglich \sim_0 äquivalent, wenn $\delta(p, \epsilon) \in F \iff \delta(q, \epsilon) \in F$. Im ersten Schritt werden daher die finalen und nicht-finalen Knoten unterschieden. Es ergeben sich die zwei Äquivalenzklassen $[q_1]_{\sim_0} = \{q_1, q_2\}$ und $[q_3]_{\sim_0} = \{q_3, q_4, q_5, q_6\}$.

\sim_1 Nun werden die Zustände ermittelt, die bezüglich Wortlänge 1 nicht unterscheidbar und somit äquivalent sind. Offenbar können nur solche Wörter äquivalent bezüglich \sim_1 sein, die bereits äquivalent bezüglich \sim_0 sind, da ja die \sim_1 -Äquivalenz für Wörter der Länge ≤ 1 gilt, also insbesondere auch Wörter der Länge 0 einschließt. Es müssen somit nur die Elemente einer jeden Äquivalenzklasse untereinander verglichen werden, ob sie bezüglich der neu dazugekommenen Wörter der Länge 1, also die Wörter a und b , äquivalent sind.

Zur systematischen Überprüfung werden die Übergänge in einer Tabelle (3.3) eingetragen. Ein Eintrag q in der Zeile x und Spalte p bedeutet: $\delta(p, x) = q$. Ein farbig markierter

	q_1	q_3	q_2	q_4	q_5	q_6
a	q_2	q_3	q_5	q_1	q_5	q_5
b	q_4	q_6	q_2	q_5	q_5	q_5

Tabelle 3.3.: Beispiel 3.7.36 - Zustandsäquivalenz \sim_1

Eintrag ist ein finaler Zustand. Die Äquivalenzklasse $[q_2]_{\sim_1}$ zerfällt in drei Äquivalenzklassen $[q_2]_{\sim_2} = \{q_2\}$, $[q_4]_{\sim_2} = \{q_4, q_6\}$ und $[q_5]_{\sim_2} = \{q_5\}$. Bei dieser Verfeinerung verbleiben diejenigen Zustände in einer Äquivalenzklasse, deren Spaltenstruktur identisch ist. So gehen beispielsweise die beiden Zustände q_1 und q_3 mit einem a jeweils in einen finalen Zustand und mit einem b in einen nicht-finalen Zustand. Das die erreichten finalen Zustände bei gelesenem a verschieden sind ist dabei irrelevant. Es ist lediglich wichtig, dass jeweils ein finaler Zustand erreicht wird. Da es drei unterschiedliche Spaltenstrukturen (a führt zu finalen Zustand, b führt zu finalen Zustand und weder a noch b führen zu finalen Zustand) gibt ergeben sich drei neue Äquivalenzklassen.

\sim_2 Nun werden die Zustände ermittelt, die bezüglich Wortlänge 2 nicht unterscheidbar und somit äquivalent bezüglich \sim_2 sind. Dazu tragen wir die Übergänge wieder in einer Ta-

Abbildung 38 Lehrender Detaillierte Metarialeinsicht

Ein letzter Punkt zur Analyse für Lehrende bietet das System in Form einer detaillierten Useranalyse, die für jeden User eines Kurses zur Verfügung steht (vgl. Abbildung 39). Innerhalb der Useranalyse zeigt das System einen zeitlichen Ablauf aller Aktivitäten eines Users im System untergliedert nach Lernmaterialien. Neben einem Start- und Endzeitpunkt hinsichtlich der Nutzung eines Materials ist auch das Hinzufügen von Tags inklusive deren Inhalten protokolliert. Ein Lehrender erhält somit die Möglichkeit, das Lernverhalten von Usern detailliert nachzuvollziehen.

Nutzerstatistik

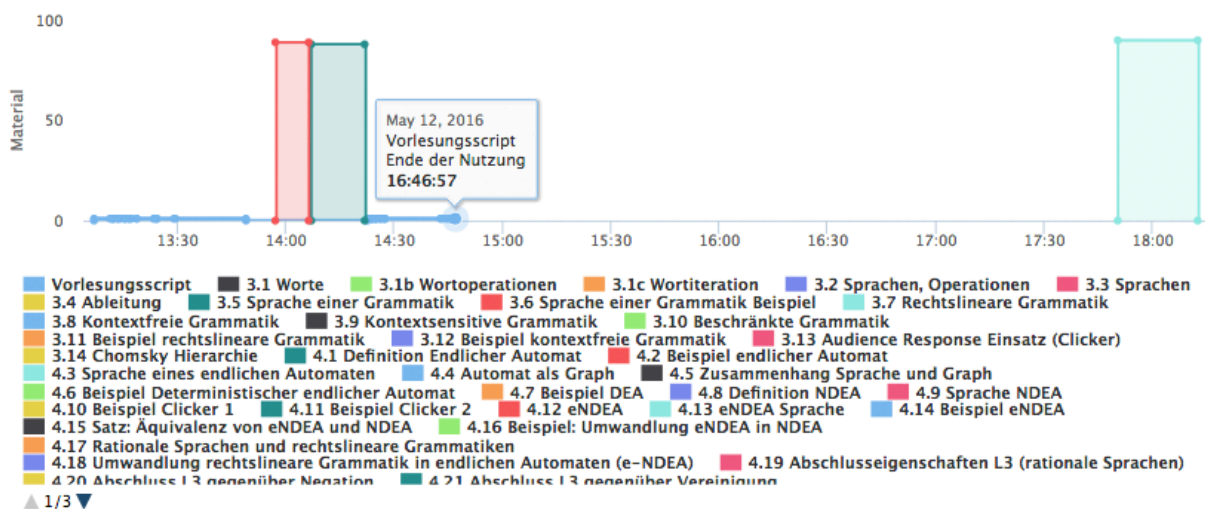


Abbildung 39 Lehrender Detaillierte Nutzeranalyse

5 Evaluation

Der entwickelte Empfehlungsdienst LAOs ist im Rahmen einer Evaluationsstudie im Zeitraum von drei Wochen von Studierenden evaluiert worden. Der Evaluationsstudie geht eine Vorstudie voraus, innerhalb derer der Empfehlungsdienst in einem verkürzten Evaluationsszenario von 90 Minuten von Studierenden getestet wurde. In diesem und den Folgekapiteln sind die Entwicklung und die Durchführung der Evaluation beschrieben. Abschnitt 3.2.5 hat bereits gängige Evaluationsmethoden für Empfehlungsdienste erörtert. Wie sich in Abschnitt 5.3 zeigt, stellt die vorliegende Arbeit eine Mischung aus Messung der Qualität (Nutzungswahrscheinlichkeit, Nützlichkeit) gemäß klassischer Evaluationsformen für Empfehlungsdienste und einem ergänzenden Fragebogen zur Prüfung der übrigen Merkmale in Abhängigkeit des TAM vor. Die nachfolgenden Abschnitte leiten ein geeignetes Evaluationsdesign auf den bereits erläuterten Methoden her. Abschnitt 5.1 legt zunächst die Problemstellung für die vorliegende Arbeit dar, über die Aspekte herausgestellt werden, die es innerhalb der Evaluation für das vorgestellte System zu untersuchen gilt. In diesem Zusammenhang wird eine erste Zielvorgabe ausgesprochen, die es in Abschnitt 5.2 durch die Formulierung geeigneter Hypothesen zu fixieren gilt. Der Abschnitt 5.3 präsentiert geeignete Methoden und Evaluationsszenarien, über die die Beantwortung der Hypothesen möglich ist. In diesem Zusammenhang werden insgesamt zwei Evaluationsszenarien beschrieben, für die in Kapitel 6 entsprechende Ergebnisse präsentiert werden können.

5.1 Problemstellung und Zielvorgabe

Für die Fixierung einer konkreten Zielvorgabe der Evaluation gilt es die Forschungsfragen der Arbeit zentral zu berücksichtigen. Die übergeordnete Forschungsfrage aus Abschnitt 1.3 erfragt

kurz zusammengefasst die Entwicklung eines Empfehlungsdienstes, welcher maschinell inhaltliche Hilfestellungen und Materialempfehlungen zur Unterstützung des Lernenden generiert. Abschnitt 1.3 stellt des Weiteren vier Teilfragen, die grundlegend eine sinnvolle Form von Hilfestellungen und Materialempfehlungen, sowie deren Güte bzw. die Güte der zugrundeliegenden Empfehlungen hinterfragen. Über eine Evaluation müssen sowohl die übergeordnete Forschungsfrage als auch die Teilforschungsfragen zu beantworten sein. Klar ersichtlich ist, dass sich alle Fragestellungen auf das entwickelte Verfahren und dessen Empfehlungen beziehen. Es bedarf somit einer erneuten Betrachtung des Verfahrens zur Darlegung geeigneter Zielvorgaben.

Der in Abschnitt 4 vorgestellte Empfehlungsdienst berücksichtigt bereits eine Vielzahl an Annahmen, die in der Literatur geprüft wurden. So ist die Herleitung von Empfehlungen auf Grundlage kollaborativ erhobener Daten einer Usergruppe eine etablierte Form. Die Annahme des „Word of Mouth“ Prinzips, also die Weiterempfehlung von Ressourcen oder Items auf Grundlage weiterer User, darf in diesem Zusammenhang als geprüft angenommen werden. Für das grundlegende Verfahren auf dem die algorithmische Entwicklung in Abschnitt 4.4 aufbaut, existieren ebenso quantitative Ergebnisse, die die positive Herleitung von Empfehlungen bestätigen (vgl. Abschnitt 3.3). Die vorliegende Arbeit präsentiert zwar ein adaptiertes, aber dennoch neu zu prüfendes Verfahren. Bedingt durch den mehrstufigen Prozess des vorgestellten Empfehlungsdienstes ergeben sich verschiedene Ansatzpunkte zur Erhebung von Messgrößen, die die (Teil-)Forschungsfragen beantworten. Gemäß Abschnitt 4.4 untergliedert sich der Prozess in die Schritte Extraktion von inhaltlichen Bereichen, Bewertung der inhaltlichen Bereiche und Herleitung von Empfehlungen (vgl. Abbildung 40). Abschnitt 4.4 legt ebenso dar, dass eine Abhängigkeit zwischen den Prozessschritten vorliegt. In besonderer Form ist die Bewertung inhaltlicher Bereiche, sowie die Herleitung von Empfehlungen vom ersten Prozessschritt, also der Extraktion inhaltlicher Bereiche, abhängig. Die Bewertung der inhaltlichen Bereiche dient grundlegend der Herleitung von Empfehlungen, weshalb beide Schritte nachfolgend als die Herleitung von Empfehlungen betrachtet werden. Die Annahme besteht darin, dass eine ungenügende Extraktion von inhaltlichen Bereichen ebenso in einer ungenügenden Herleitung von Empfehlungen münden würde. Wie in Abbildung 40 dargelegt ergeben sich somit zwei zu prüfende Stellen innerhalb des Prozesses.

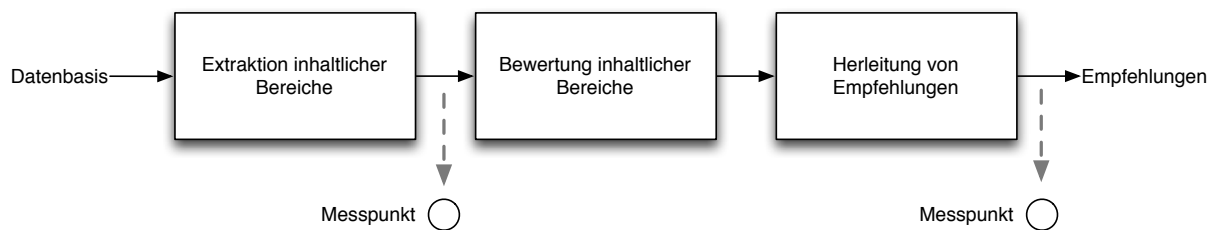


Abbildung 40 Empfehlungsdienstprozess LAOs

Die Extraktion inhaltlicher Bereiche wird innerhalb eines bestimmten Lernkontextes durchgeführt. Zur Bestimmung einer Güte bedarf es somit zunächst einer Beurteilung, wie relevant sich die extrahierten Bereiche innerhalb eines bestimmten Lernkontextes darstellen. Dies kann in Form der *Nutzungswahrscheinlichkeit* gemessen werden. Als zweiter Ansatzpunkt kann aufbauend auf den extrahierten inhaltlichen Bereichen eine Güte für die hergeleiteten Empfehlungen ermittelt werden. Es bedarf in diesem Zusammenhang der Messung der Güte hergeleiteter Empfehlung in Form der *Nützlichkeit* für Lernende. Wenn sowohl eine genügende Nutzungswahrscheinlichkeit als auch eine genügende Nützlichkeit für extrahierte inhaltliche Bereiche und deren Empfehlungen gegeben ist, kann von einer generellen Unterstützung durch das System ausgegangen werden.

Ein weiterer Ansatzpunkt, den es innerhalb der Evaluation zu prüfen gilt, ist wann das System eine genügende Nutzungswahrscheinlichkeit bzw. Nützlichkeit für Empfehlungen und somit eine Unterstützung für Lernende erreicht. Die Extraktion inhaltlicher Bereiche erfolgt auf der Clusterung von Tags. Es kann deshalb davon ausgegangen werden, dass eine unterschiedliche Menge an Tags (bzw. eine unterschiedlich große Datenbasis) zu unterschiedlichen Ergebnissen hinsichtlich der Extraktion inhaltlicher Bereiche führt, die sowohl Einfluss auf die Nutzungswahrscheinlichkeit als auch Einfluss auf die Nützlichkeit für Empfehlungen nach sich ziehen. Innerhalb der Evaluation ist es deshalb von Bedeutung zu prüfen, ob ein Zusammenhang zwischen der Größe der Datenbasis und der Qualität von Empfehlungen existiert (vgl. Abbildung 41). In diesem Zusammenhang ist die Nennung eines Schwellwertes wünschenswert, ab welcher Anzahl an Tags das Verfahren eine genügende Nutzungswahrscheinlichkeit inhaltlicher Bereiche bzw. eine genügende Nützlichkeit für Empfehlungen gewährleistet. Des Weiteren wurde angenommen, dass die Präsentation von alternativen Lernmaterialien oder Lernmaterialverknüpfungen einen Mehrwert für Lernende erwirkt. Auch in diesem Zusammenhang ist es sinnvoll zu erörtern, ob ein Zusammenhang zwischen der Anzahl bzw.

dem Aufkommen von alternativen Lernmaterialien/Lernmaterialverknüpfungen und der bewerteten Nützlichkeit existiert.

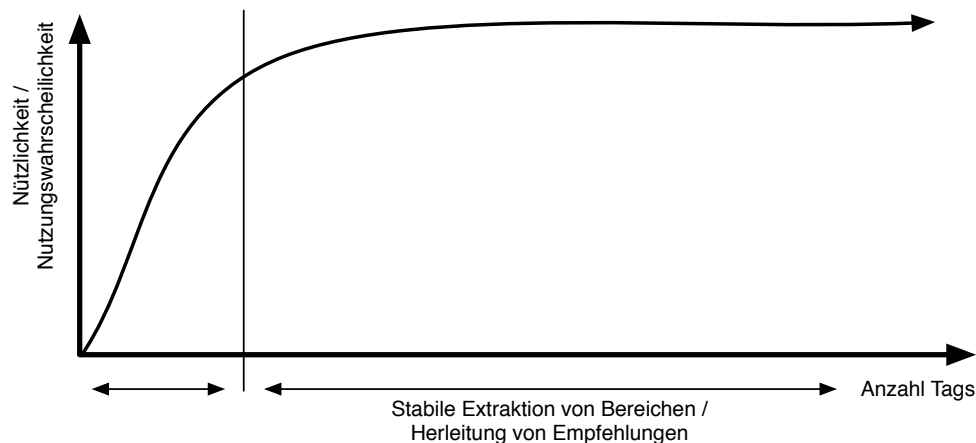


Abbildung 41 Zusammenhang Datenbasis und Qualität⁷⁴

Da es sich um eine grundlegende Neuentwicklung des Systems handelt, ist auch die Prüfung der Akzeptanz des Systems aus Sicht der Lernenden sinnvoll. Diese kann unter Berücksichtigung der Qualität oder Güte der Empfehlungen als auch der Prüfung des Aufwandes für Lernende z.B. mit Hinblick auf die Nutzerfreundlichkeit (Usability) des Systems bestimmt werden. Ein in der Literatur gängiges Verfahren zur Prüfung der Akzeptanz ist die Nutzung des Technology Acceptance Model (TAM) nach Davis. In diesem Zusammenhang ergibt sich die Möglichkeit sowohl die inhaltliche als auch die visuelle Form von Empfehlungen zu hinterfragen.

5.2 Hypothesen

Auf Grundlage der zuvor erhobenen Zielvorgaben sollen nachfolgend geeignete Hypothesen hergeleitet werden, die insbesondere die Forschungsfragen aus Abschnitt 1.3 beantworten. Die durchzuführende Evaluation ist eine Mischung aus Hypothesenprüfender Untersuchung und explorativer quantitativer Datenanalyse (vgl. [15] S. 490 bzw. S. 371). Wo die Zusammenhänge zwischen Datenmenge und Güte einer theoretisch fundierten (Veränderungs-)Hypothese ($H_3 - H_5$) entsprechen die es mittels entsprechender Signifikanztests zu prüfen gilt, ergeben

⁷⁴ Die Abbildung dient lediglich der Illustration der getätigten Annahme und trifft keine Aussage ob und in welchem Maß ein Anstieg der Nützlichkeit bzw. Nutzungswahrscheinlichkeit zu erwarten ist.

sich für die Frage nach der Güte des Systems Hypothesen ($H_1, H_2, H_6 - H_{10}$), deren theoretischer Hintergrund durch entsprechende Beantwortung erst gebildet wird. Die Hypothesen sind im Sinne der Übersichtlichkeit den zu Beginn der Arbeit gestellten Forschungsfragen untergeordnet, die es final mittels Hypothesenprüfung zu beantworten gilt.

Forschungsfrage 1: Welche Güte haben die durch das System zur Verfügung gestellten Empfehlungen?

Hypothese (H_1): Das System extrahiert innerhalb eines bestimmten Lernkontextes relevante inhaltliche Bereiche, für die eine positiv erkennbare Nutzungswahrscheinlichkeit seitens der Nutzer existiert.

Hypothese (H_2): Das System leitet innerhalb eines bestimmten Lernkontextes Empfehlungen her, für die eine positiv erkennbare Nützlichkeit seitens der Nutzer existiert.

Forschungsfrage 2: Welche Menge an Daten (Datenbasis) muss für die Herleitung positiv empfundener Empfehlungen zur Verfügung stehen?

Hypothese (H_3): Es existiert ein Zusammenhang zwischen der Anzahl an Tags, die für die Extraktion eines inhaltlichen Bereichs Verwendung finden, und der ermittelten Nutzungswahrscheinlichkeit des jeweiligen Bereiches.

Hypothese (H_4): Es existiert ein Zusammenhang zwischen der Anzahl an Tags, die für die Herleitung einer Empfehlung Verwendung finden, und der ermittelten Nützlichkeit der jeweiligen Empfehlung.

Hypothese (H_5): Es existiert ein Zusammenhang zwischen der Anzahl alternativer Lernmaterialien oder Materialverknüpfungen, die innerhalb einer Empfehlung präsentiert werden und der ermittelten Nützlichkeit der jeweiligen Empfehlung.

Forschungsfrage 3: Was umfasst eine unterstützende inhaltliche Hilfestellung innerhalb des Empfehlungsdienstes?

Hypothese (H_6): Für inhaltliche Hilfestellungen innerhalb hergeleiteter Empfehlungen existiert eine erkennbare positiv unterstützende inhaltliche sowie visuelle Darstellung aus Sicht der Nutzer.

Forschungsfrage 4: Was umfasst eine unterstützende inhaltliche Verbindung zwischen Lernmaterialien innerhalb des Empfehlungsdienstes?

Hypothese (H_7): Für eine inhaltliche Verbindung zwischen Lernmaterialien innerhalb hergeleiteter Empfehlungen existiert eine erkennbare positiv unterstützende inhaltliche sowie visuelle Darstellung aus Sicht der Nutzer.

Forschungsfrage 5: Kann ein Empfehlungsdienst entwickelt werden, der maschinell inhaltliche Hilfestellungen für bestehende Lernmaterialien als auch inhaltliche Verknüpfungen zwischen bestehenden und neuen, ungenutzten Lernmaterialien zur Unterstützung von Lernenden in Form von Empfehlungen generiert?

Hypothese (H_8): Das System leitet innerhalb eines bestimmten Lernkontextes Empfehlungen her, deren Verknüpfungen zwischen bestehenden und neuen, ungenutzten Lernmaterialien eine positiv erkennbare Unterstützung für die Nutzer darstellen.

Hypothese (H_9): Das System leitet innerhalb eines bestimmten Lernkontextes Empfehlungen her, deren Hilfestellungen für bestehende Lernmaterialien eine positiv erkennbare Unterstützung für die Nutzer darstellen.

Hypothese (H_{10}): Es existiert eine positiv erkennbare Akzeptanz zur Nutzung des Systems seitens der Nutzer.

5.3 Aufbau der Studie

Der nachfolgende Abschnitt beschreibt das für die Evaluation anzuwendende Forschungsdesign⁷⁵. Das Forschungsdesign konkretisiert die Methoden und Abläufe, mit deren Hilfe Fragestellungen empirisch untersucht werden können. Als Grundlage für die vorliegende Arbeit dienen die in den vorangegangenen Abschnitten diskutierte Problemstellung, sowie die gestellten Hypothesen. In Abschnitt 5.3.1 wird zunächst eine Methodik für die Studie hergeleitet, auf deren Grundlage die Abschnitte 5.3.2 und 5.3.3 zwei Evaluationsszenarien zur Durchführung beschreiben.

⁷⁵ Oder auch Untersuchungsdesign, Untersuchungsanordnung oder Versuchsplan

5.3.1 Methode

Wie Kapitel 4 beschreibt, ist der Empfehlungsdienst inklusive entsprechender Tagging Funktionalität und gemäß der Zielvorgaben vollständig implementiert. Eine zentrale Anforderung an das System ist die Möglichkeit bestehende Lernmaterialien Lehrender zu verwenden, in denen Lernende aktiv arbeiten können und Empfehlungen erhalten. Bedingt durch die neu konstruierte Umgebung, die der Vorgabe nachkommen, existiert hierfür keine geeignete Datenbasis, auf der ein offline Experiment durchgeführt werden könnte. Ein offline Experiment wäre zunächst zwar wünschenswert, um in erster Linie das algorithmische Verfahren in einem kostengünstigen Schritt zu evaluieren, allerdings ist die aktive Einbeziehung von Personengruppen zum Aufbau einer geeigneten Datenbasis unabdingbar. Dennoch soll auf ein schrittweises Vorgehen nicht verzichtet werden und so ist zunächst eine Prüfung des algorithmischen Verfahrens sinnvoll, bevor das System in einem ganzheitlichen Rahmen evaluiert wird. Hierfür bedarf es einer zweistufigen Evaluation, die aus einer Benutzerstudie unter Laborbedingungen und einem online Experiment besteht. Innerhalb der Benutzerstudie wird eine Datenerhebung durchgeführt, die schlussendlich für eine Beurteilung und Anpassung des algorithmischen Verfahrens durch einen Experten dient. Gemäß Abbildung 40 in Abschnitt 5.1 handelt es sich in diesem Zusammenhang zunächst um die Prüfung des ersten Messpunktes und somit um die Beurteilung der Extraktion inhaltlicher Bereiche. In diesem Zusammenhang können zudem erste Erkenntnisse für einen Zusammenhang zwischen einer ansteigenden Datenbasis und der Auswirkung auf die Extraktion inhaltlicher Bereiche ermittelt werden. Hierauf aufbauend kann dann ein online Experiment erfolgen, in dem das ganzheitliche Verfahren hinsichtlich der Nutzungswahrscheinlichkeit, Nützlichkeit und Akzeptanz zu prüfen ist. In diesem Zusammenhang werden sowohl die extrahierten Bereiche als auch die hergeleiteten Empfehlungen evaluiert (vgl. Abbildung 40 Messpunkt 1 und Messpunkt 2). Die Prüfung eines Zusammenhangs zwischen ansteigender Datenbasis und ansteigender Qualität hinsichtlich der Nutzungswahrscheinlichkeit bzw. Nützlichkeit erfolgt dann auf Grundlage von Probandenaussagen. Eine zweistufige Evaluation in Form einer Benutzerstudie und einer nachgelagerten online Simulation ist deshalb sinnvoll, um ggf. Schwachstellen des algorithmischen Verfahrens hinsichtlich der Extraktion relevanter Bereiche auf Grundlage einer Expertenaussage in einem kostengünstigen Schritt aufzudecken, bevor eine finale Evaluation der hergeleiteten Empfehlungen in Form eines aufwändigeren online Experimentes erfolgt.

Es ergeben sich somit zwei getrennt voneinander durchzuführende Evaluationsschritte, deren Aufbau in Form eines Laborversuchs bzw. Benutzerstudie in Abschnitt 5.3.2 bzw. in Form eines Realversuchs bzw. online Simulation in Abschnitt 5.3.3 beschrieben ist. Abschnitt 5.3.2 präsentiert dabei ein eher konstruiertes Lernszenario mit frei gewählten Lernmaterialien über die Dauer einer schriftlichen Prüfungsleistung, um eine erste Datenbasis zu erheben und erste Erkenntnisse für das Verfahren ableiten zu können. Abschnitt 5.3.3 beschreibt dann die Nutzung in einem curricularen Kurs an einer Hochschule über die Dauer mehrerer Wochen.

5.3.2 Benutzerstudie unter Laborbedingungen

Im ersten Evaluationsschritt ist es von Bedeutung beurteilen zu können wie gut das entwickelte Verfahren inhaltlich zusammenhängende bzw. relevante Bereiche extrahiert. Wie bereits dargelegt, kann dies in Form einer Benutzerstudie unter Laborbedingungen erfolgen. Verschiedene Aspekte gilt es in diesem Zusammenhang zu berücksichtigen. Zum einen bedarf es einer genügenden Datenbasis, die es durch reale Probanden zu erheben gilt. Um eine genügend große Datenbasis hinsichtlich gesetzter Tags erheben zu können, sind Mindestanforderungen im Umgang mit dem System zu stellen. Gleichzeitig bedarf es allerdings möglichst realer Rahmenbedingungen, so dass Aussagen, die innerhalb des ersten Evaluationsschrittes getätigt werden, ebenso Gültigkeit für den Folgeschritt, der unter realen Bedingungen stattfindet, besitzen. Die Benutzerstudie ist als eine Art Entwicklertest zu sehen, der erste Aussagen über die Qualität des Verfahrens trifft.

So sieht der erste Evaluationsschritt die Bearbeitung eines Aufgabenblattes mit vorgegebenen Lernmaterialien vor, welches von Studierenden im universitären Umfeld bearbeitet wird. Um an ein reales Szenario anknüpfen zu können, ist die Bearbeitung eines Themenfeldes vorgesehen, welches aus dem Studienverlauf der Probandengruppe entstammt. Für den Bereich der Informatik bietet sich in diesem Zusammenhang das Themenfeld der „Algorithmen und Datenstrukturen“ an, das als grundständiges Fach in jedem Studium mit Informatikschwerpunkt zu sehen ist. Hierdurch ergibt sich der Vorteil, dass eine Vielzahl an Studierenden angesprochen werden kann.

Das Aufgabenblatt umfasst insgesamt 20 Aufgaben zu fünf verschiedenen Themenbereichen (Grundbegriffe Algorithmen, Insertionsort Algorithmus, Mergesort Algorithmus, Quicksort Algorithmus, Dijkstra Algorithmus). Eine Kopie des Aufgabenblattes kann dem Anhang der Arbeit entnommen werden (vgl. Anhang C). Das Schwierigkeitsniveau der Aufgaben ist moderat und sowohl Studierende mit als auch Studierende ohne Vorkenntnisse sollten die Aufgaben unter zu Hilfenahme passender Lernmaterialien lösen können. Neben der Beantwortung von reinem Faktenwissen sind algorithmische Verfahren – vorrangig Sortierverfahren – anzuwenden. Die zur Verfügung stehenden Lernmaterialien entstammen dem Internet und sind frei verfügbar. Des Weiteren ist das Niveau der Materialien für eine Lehrveranstaltung im Hochschulkontext angemessen. Insgesamt stehen den Probanden ein Vorlesungsskript, ein Buch und fünf Lernvideos zur Verfügung (vgl. Tabelle 5). Für jeden abgefragten Themenbereich gibt es ein entsprechendes Kapitel im Vorlesungsskript, sowie jeweils ein entsprechendes Lernvideo. Teilinformationen sind in den Videos zu finden, obgleich die Aufgaben ausschließlich mit dem Vorlesungsskript lösbar sind. Die Videos dienen somit der Ergänzung des Skriptes und zur Veranschaulichung der algorithmischen Probleme. Für die Bearbeitung der Aufgaben bedarf es lediglich der Nutzung weniger Seiten und Teilbereiche des Vorlesungsskriptes, so dass eine tiefgehende Recherche nicht notwendig ist. Als Beispiel ist von den Probanden die Sortierung einer Zahlenfolge unter Verwendung des Mergesort Algorithmus durchzuführen. Im Vorlesungsskript gibt es hierzu einen Abschnitt, der die Durchführung in formeller und beispielhafter Form über drei Seiten erläutert. Als Ergänzung steht ein Video zum Mergesort Algorithmus zur Verfügung, welches eine beispielhafte Durchführung des Algorithmus erläutert. Über das Inhaltsverzeichnis des Vorlesungsskriptes sind die entsprechenden Kapitel direkt auffindbar. Ebenso beinhalten die Namen der Videos einen klaren inhaltlichen Bezug. Als Alternative ergeben sich ebenso Hilfestellungen innerhalb des bereitgestellten Buches, welches als umfassendes Nachschlagewerk zu beurteilen ist. Die Aufgaben sind bewusst so gestellt, dass für die Bearbeitung der Aufgaben von der Nutzung einiger weniger inhaltlicher Teilbereiche innerhalb der Materialien ausgegangen werden kann. So soll in einem ersten Schritt der Evaluation eine Bewertung der extrahierten inhaltlichen Bereiche zur sinnhaften Nutzung im Kontext der gestellten Aufgaben durchgeführt werden. Anders gesprochen wird von einer Klassifizierung in relevanter bzw. nicht relevanter inhaltlicher Bereiche ausgegangen, die zur

Lösung der Aufgaben sinnvoll sind und die durch das algorithmische Verfahren bestätigt werden sollen.

Lernmaterial	Größe
Vorlesungsskript	160 Seiten
Lehrbuch Algorithmen und Datenstrukturen	384 Seiten
Lernvideo Funktionsweise Mergesort Algorithmus	04:55 Minuten
Lernvideo Funktionsweise Quicksort Algorithmus	04:30 Minuten
Lernvideo Funktionsweise Quicksort Algorithmus (In-Place)	07:35 Minuten
Lernvideo Funktionsweise Dijkstra Algorithmus	24:48 Minuten
Lernvideo Funktionsweise Dijkstra Algorithmus APSP	21:46 Minuten

Tabelle 5 Lernmaterialien 1. Evaluationsschritt

Die Probanden erhalten vor der Bearbeitung des Aufgabenblattes eine kurze Einführung in das System LAOs, die sowohl die Navigation im System als auch innerhalb der Lernmaterialien umfasst, und des Weiteren eine Übersicht über die zur Verfügung stehenden Taggingwerkzeuge beinhaltet. Des Weiteren sind die Taggingwerkzeuge, sowie die wichtigsten Funktionsweisen des Systems im Anhang des Aufgabenblattes detailliert beschrieben (vgl. Anhang C). Da die Lernmaterialien grundlegend neu für die Probanden sind, bedarf es vor der Bearbeitung des Aufgabenblattes ebenfalls einer kurzen Einführung in die zur Verfügung stehenden Materialien.

Nach der Einführung beginnen die Probanden mit der Bearbeitung der Aufgaben. Für die Bearbeitung der Aufgaben erhalten die Probanden 90 Minuten Zeit, wobei die Bearbeitung eigenständig durchzuführen ist. Die Probanden schreiben ihre Lösung auf das Aufgabenblatt, welches nach Beendigung abgegeben wird. Für den Zugang zum System erhält jeder Proband eine Kennung, die auf dem jeweiligen Aufgabenblatt vermerkt ist. Es ist somit eine klare Zuordnung zwischen Kennung innerhalb des Systems und abgegebenem Aufgabenblatt möglich. Personenbezogene Daten werden in Form von Geschlecht, Alter, Semester und Studienfach erhoben. Es erfolgt somit eine anonyme Teilnahme durch die Probanden. Neben der Bearbeitung des Aufgabenblattes sind die Probanden angehalten genutzte Stellen mit freigewählten Tags zu annotieren. Um eine genügende Datenbasis generieren zu können, erhalten die Probanden die Vorgabe mindestens 20 Tags in beliebigen Lernmaterialien zu setzen. Des Weiteren können die Probanden das Internet zu Recherchezwecken verwenden, sind allerdings dazu angehalten

alternativ genutzte Materialien oder Internetseiten unter Verwendung des entsprechenden Tags mit in das System einzubringen und zu verlinken.

Das Evaluationssetting wurde positiv mit zwei Informatikstudierenden vor dem Evaluationsdurchlauf getestet, um zu prüfen ob der zeitliche Rahmen als auch der Umfang und die Menge an Lernmaterialien zur Bearbeitung der Aufgaben genügen.

Auf Grundlage der Datenerhebung durch die Probanden extrahiert das System inhaltliche Teilbereiche innerhalb der Lernmaterialien, die als kontextrelevante Bereiche zur Lösung der Aufgaben anzusehen sind. Dabei nimmt das System intern eine Klassifizierung zwischen relevanten und nicht relevanten Bereichen vor. Einer fachkundigen Person⁷⁶ werden sowohl die Aufgaben als auch die extrahierten Bereiche – *ohne* sichtbare Klassifikation – präsentiert, die ihrerseits eine eigene Klassifikation durchführt. Auf dieser Grundlage kann eine Klassifikationstabelle gemäß Tabelle 2 (Seite 59) erstellt werden, die eine Beurteilung der Klassifikation über die Errechnung von Precision und Recall erlaubt. In Abhängigkeit der Klassifikation kann des Weiteren eine Beurteilung über den Zusammenhang zwischen Anzahl an Tags und der Anzahl richtig klassifizierter Bereiche erfolgen, um die entsprechende Hypothese aus Abschnitt 5.2 bestätigen bzw. ablehnen zu können. Grundlegend wird geprüft, wie häufig inhaltliche Bereiche bei einer bestimmten Anzahl an Tags als richtig bzw. falsch klassifiziert wurden. Das Vorgehen ist in Abbildung 45 auf Seite 155 zum besseren Verständnis illustriert. Diese Betrachtungsweise soll abschließend die Nennung eines Schwellwertes ermöglichen, ab wann das Verfahren zuverlässig inhaltliche Bereiche richtig klassifiziert. Einen etwaigen Schwellwert gilt es in Abhängigkeit des zweiten Evaluationssettings (vgl. Abschnitt 5.3.3) zu bestätigen.

5.3.3 Online Simulation in realem Kursumfeld

Im Vergleich zum Evaluationssetting in Abschnitt 5.3.2 ergeben sich, bedingt durch die unterschiedliche Zielsetzung, für die online Simulation grundlegende Änderungen. Innerhalb der

⁷⁶ In der Benutzerstudie ist mit fachkundiger Person der Ersteller des Aufgabenblattes sowie der Entwickler des Verfahrens gemeint, da es sich in einem ersten Schritt – wie erläutert – der Beurteilung des Verfahrens und somit einer Art Entwicklertest bedarf.

online Simulation soll eine Bewertung sowohl für die extrahierten inhaltlichen Bereiche als auch für die Empfehlungen durch die User des Systems durchgeführt werden. Hierfür bedarf es der längerfristigen Arbeit mit dem System und entsprechender Empfehlungen. Es bietet sich deshalb die Durchführung des zweiten Evaluationsschrittes innerhalb eines realen, curricularen Kurses an. In diesem Zusammenhang besteht die Möglichkeit zur Nutzung des Systems im Kurs *Theoretische Informatik* an der Hochschule Osnabrück im Sommersemester 2016. Der Kurs ist eine Pflichtveranstaltung für die Informatikstudiengänge Medien- und Technische Informatik mit einem zeitlichen Umfang von 4 Semesterwochenstunden⁷⁷. Studierende erhalten für den Abschluss des Kurses 5 ECTS Punkte.

Auch für den Realversuch bedarf es der Durchführung innerhalb eines bestimmten Lernkontextes, in dem die Probanden arbeiten. Für den Kurs *Theoretische Informatik* sind regelmäßige Übungsblätter vorgesehen, die Studierenden zur Erlernung eines bestimmten Themenfeldes bereitgestellt werden. Darüber hinaus erhalten die Studierenden gegen Ende des Semesters ein umfangreicheres Übungsblatt, welches verschiedene Themen abdeckt und zur Vorbereitung auf die abschließende Prüfung dient. Das Übungsblatt zur Prüfungsvorbereitung scheint ein sinnvoller Kontext für die Durchführung zu sein, da durch die Bearbeitung verschiedener Themenbereiche ebenso verschiedene Lernmaterialien bzw. Teile von Lernmaterialien zu nutzen sind. So werden Empfehlungen in unterschiedlichen Stellen der Lernmaterialien erwartet, was die Aussagekraft der Ergebnisse steigern würde.

Die online Simulation sieht somit die Bearbeitung eines Aufgabenblattes im Kurs *Theoretische Informatik* vor, welches von realen Kursteilnehmern zu bearbeiten ist. Das Aufgabenblatt umfasst insgesamt fünf Aufgaben, die grundlegend jeweils einem Themengebiet der theoretischen Informatik zugeordnet werden können. Anders als im Laborversuch aus Abschnitt 5.3.2 lassen sich die Aufgaben weniger differenziert betrachten und so bauen die Themenbereiche in der online Simulation aufeinander auf. In Anhang D ist eine Kopie des Übungsblattes hinterlegt. Innerhalb der Übungen bedarf es nicht der Beantwortung von reinem Faktenwissen, sondern vielmehr der Anwendung von Methodenwissen und Beweistechniken.

⁷⁷ entspricht 2x Mal pro Woche 90 Minuten Präsenzzeit

Das Niveau entspricht dabei dem einer abschließenden Klausur. Zur Bearbeitung des Übungsblattes stehen den Probanden die im Verlauf des Semesters bereits bekannten und genutzten Lernmaterialien zur Verfügung. Neben einem Vorlesungsskript ist eine Vielzahl an themenspezifischen Lernvideos bereitgestellt, die jeweils einem Themenbereich zugeordnet werden können. Die Videos dienen dabei als Ergänzung zum Skript und sollen die Inhaltsvermittlung in Audio-/Visueller Form unterstützen. Insgesamt stehen 88 Videos zur Verfügung. Anders als in der Benutzerstudie in Abschnitt 5.3.2 wurde zudem das Aufgabenblatt selbst mit im Materialpool zur Verfügung gestellt. Eine Zusammenfassung der zur Verfügung stehenden Materialien ist in Tabelle 6 gelistet. Bedingt durch die Vielzahl an Videos ist eine ganzheitliche Auflistung zu Gunsten der Übersicht Anhang E zu entnehmen.

Lernmaterial	Größe
Vorlesungsskript Theoretische Informatik	261 Seiten
Übungsblatt	2 Seiten
Lernvideos	Durchschnittliche Länge ~6 Minuten

Tabelle 6 Lernmaterialien 2. Evaluationsschritt

Die Einführung in das System erfolgt in Form eines Videos, welches die Probanden vor Beginn der Evaluation online abrufen können. In diesem Video werden alle Aspekte und Funktionen des Systems detailliert erläutert. Neben der Erläuterung von Taggingwerkzeugen und Navigationselementen enthält das Video auch eine Beschreibung zum Abrufen und Bewerten von Empfehlungen. Als weitere Unterstützung erhalten die Probanden eine textuelle Beschreibung, in der die Taggingwerkzeuge sowie die wichtigsten Funktionalitäten im System zusammengefasst sind.

Für die Evaluation erhalten die Probanden zwei separat voneinander durchzuführende Arbeitsaufträge, die in einen Bearbeitungszeitraum und einen Bewertungszeitraum unterteilt werden können (vgl. Abbildung 42). Die Probanden werden eine Woche vor Start der Evaluation innerhalb eines regulären Veranstaltungstermins über ihre durchzuführenden Aufgaben informiert. Des Weiteren erhalten alle Teilnehmer des Kurses eine separate E-Mail über das Lernmanagementsystem mit gleichen Instruktionen sowie individuelle Zugangsdaten zum System als auch die Information über den Zugang zu den vorab erläuterten Hilfsmaterialien.

Dies gewährleistet, dass alle Teilnehmer des Kurses informiert sind. Zum Start des Bewertungszeitraums erhalten die Probanden zudem eine erneute E-Mail, die in erster Linie der Erinnerung dient. Des Weiteren wird innerhalb der E-Mail noch einmal das Bewertungsverfahren erläutert, um den Probanden einen möglichst klaren und transparenten Arbeitsauftrag darzulegen. Die Befragung zur Akzeptanz des Systems erfolgt nach Beendigung des Bewertungszeitraums und wird ebenfalls via E-Mail angekündigt. Der Fragebogen wird den Probanden online zur Verfügung gestellt.

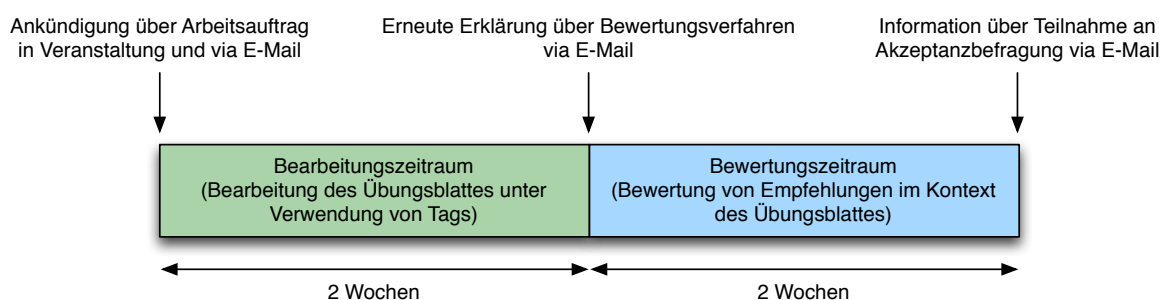


Abbildung 42 Evaluationsablauf online Simulation

Im Bearbeitungszeitraum erhalten die Probanden die Aufgabe das Übungsblatt unter Berücksichtigung der Lernmaterialien im System zu bearbeiten. Um eine Abhängigkeit zum System zu schaffen, ist das Übungsblatt ausschließlich im LAOs System abrufbar. Die Probanden werden dazu angehalten ausschließlich online mit den Lernmaterialien zu arbeiten und diese mit entsprechenden Tags während des Lernens zu versehen. Die Nutzung des Internets ist unter der Bedingung erwünscht, dass die Probanden alternativ genutzte Lernmaterialien unter Verwendung des entsprechenden Tags mit in das System aufnehmen. Hierfür erhalten die Probanden zwei Wochen Zeit. Grundsätzlich wäre für die Bearbeitung eines entsprechenden Übungsblattes ca. eine Woche vorgesehen. Damit die Probanden allerdings genügend Zeit für die Einarbeitung in das System aufbringen können, ist der Zeitraum entsprechend angepasst.

Auf Grundlage der durch die Probanden gesetzten Tags generiert das System Empfehlungen, die innerhalb des Bewertungszeitraumes durch Probanden zu bewerten sind. Das System nimmt hier zunächst eine Extraktion von inhaltlichen Bereichen vor, die eine Klassifizierung in relevante bzw. nicht relevante Bereiche erhalten. Diese Klassifizierung ist durch die Probanden nicht ersichtlich. Für die extrahierten inhaltlichen Bereiche leitet das System des Weiteren Empfehlungen ab. Die Lernmaterialien, in denen Empfehlungen zu finden sind, erkennen die

Probanden an einer entsprechenden Markierung im System (vgl. Abbildung 43)⁷⁸. Zum einen erhalten die Probanden die Möglichkeit für inhaltliche Bereiche eine Klassifizierung in relevant bzw. nicht relevant durchzuführen, was einer Bewertung hinsichtlich der Nutzungswahrscheinlichkeit entspricht. Hierfür stellt das System einen entsprechenden Bewertungsbutton zur Verfügung (vgl. Abbildung 44 unten links). Des Weiteren erhalten die Probanden die Aufgabe Empfehlungen gemäß ihrer Nützlichkeit zu bewerten. Hierfür steht eine Bewertung in Form von Sternen (1 Stern entspricht „gar nicht nützlich“; 5 Sterne entspricht „sehr nützlich“) zur Verfügung (vgl. ebenfalls Abbildung 44 unten rechts). Der Zeitraum für die Bewertung durch die Probanden beträgt ebenfalls zwei Wochen.

Materialpool

90 Materialien

Textdokument Videodokument Audiodokument Slides Link

Textdokument







 Vorlesungsscript   Vorlesungsscript Theoretische Informatik
 Übungsklausur   Übungsklausur TI 2016

Abbildung 43 Hinweis auf Empfehlungen im System LAOs

⁷⁸ Das Augensymbol verweist auf eine häufige Nutzung durch die Probandengruppe. Das Informationssymbol verweist auf Empfehlungen innerhalb des Materials.

Empfehlung ✕

Einfach **Verständlich** **Wichtig** **Unterstützend**

wenig sehr wenig sehr wenig sehr wenig sehr

Der Bereich ist...
 wenig schwierig, anschaulich, motivierend, sehr eingängig, sehr präzise, relevant, hilfreich,
Du solltest den Bereich nutzen, weil er...
 wenig schwierig ist, zum lernen anregt, inhaltlich sehr gut aufbereitet ist, wichtig ist,

Weitere Links
www.informatikseite.de/theorie/node61.php
 CNF
 kontextfreie Grammatik
 CYK Idee
 CYK Beispiel
de.wikipedia.org/wiki/Chomsky-Normalform
www.informatikseite.de/theorie/node62.php

Abbildung 44 Bewertungsschema für Empfehlungen

Im Anschluss an den Bewertungszeitraum erfolgt eine online Befragung zur Akzeptanz des Systems gemäß Technology Acceptance Model (TAM) in Form eines Fragebogens. Abschnitt 3.2.5 beschreibt das TAM bereits ausführlich. Kurz zusammengefasst erläutert das Modell im Kontext von Empfehlungsdiensten den Zusammenhang zwischen den Einflussfaktoren Qualität der Empfehlungen⁷⁹ und den Aufwand für den Erhalt von Empfehlungen⁸⁰, die maßgeblichen Einfluss auf die Nutzung eines Systems nach sich zieht. Somit beinhaltet der Fragebogen Aussagen zu den Kategorien Qualität und Aufwand, die durch die Probanden auf einer fünfstufigen Lickertskala mit den Werten „trifft gar nicht zu“, „trifft nicht zu“, „neutral“, „trifft zu“ und „trifft vollständig zu“ zu bewerten sind. Den Kategorien Qualität und Aufwand schließen sich Fragen zur generellen Akzeptanz an, die auf der selben Skala zu bewerten sind. Der Fragebogen entspricht einer Adaption eines validierten Fragebogens aus [68, 77] und ist in

⁷⁹ Gemäß Definition in Abschnitt 3.2.5 auch als Perceived Usefulness bezeichnet.

⁸⁰ Gemäß Definition in Abschnitt 3.2.5 auch als Perceived Ease of Use bezeichnet.

Tabelle 7 abgebildet⁸¹. Für die Beantwortung des Fragebogens erhalten die Probanden eine Woche Zeit.

Qualität	
1	Die Empfehlungen waren für die Suche nach Informationen wertvoll
2	Ich hätte die Bereiche der Empfehlungen ebenfalls markiert
3	Die Empfehlungen würden meine Produktivität erhöhen
4	Die Empfehlungen würden es mir erleichtern Inhalte zu erlernen
5	Mir würde es gefallen, dass das System meine Stärken und Schwächen erkennt
6	Ich habe durch das System Inhalte entdeckt, die ich vorher nicht kannte
7	Ich habe durch das System neue Materialien entdeckt, die ich vorher nicht kannte
Aufwand	
8	Ich habe mich im System schnell zurecht gefunden
9	Mir fiel es leicht durch das Textdokument zu navigieren
10	Mir fiel es leicht durch das Videodokument zu navigieren
11	Mir fiel es leicht Werkzeuge im System zu nutzen und Annotationen zu setzen
12	Die Auswahl der Werkzeuge war übersichtlich
13	Die Auswahl der Werkzeuge war verständlich
14	Die Menge an Werkzeugen war ausreichend
15	Die Empfehlungen waren leicht zugänglich
16	Die Darstellung der Empfehlungen war verständlich
17	Das System würde mich beim Lernen einschränken
18	Ich würde durch das System länger für die Bearbeitung von Aufgaben benötigen
Akzeptanz	
19	Mir würde es helfen zu wissen welche Informationen in Lernmaterialien wichtig, unterstützend, schwierig oder verständlich sind
20	Mir würde es helfen zu wissen wie andere Studierende Lerninhalte einschätzen
21	Ich würde das System in der regulären Lehre nutzen wollen
22	Mir würde es helfen alternative Lerninhalte vorgeschlagen zu bekommen

Tabelle 7 Fragebogen Akzeptanz gemäß TAM

⁸¹ Negativ gestellt Fragen sind in Tabelle 7 rot hinterlegt

Die Auswertung der durch die Probanden abgegebenen Bewertungen erfolgt nach Abschluss des Bewertungszeitraums. In diesem Zusammenhang erfolgt zunächst ein Abgleich der durch die Probanden bzw. durch das System durchgeführte Klassifikation der inhaltlichen Bereiche hinsichtlich ihrer Relevanz. Dies erfolgt wie in Abschnitt 3.2.5 bzw. Abschnitt 5.3.2 beschrieben, über eine Klassifikationstabelle, über die die Metriken Precision und Recall abgeleitet werden können. Gemäß der Ausführung in Abschnitt 3.2.5 ergibt sich hierüber die Nutzungswahrscheinlichkeit für die extrahierten inhaltlichen Bereiche und dementsprechend der abgeleiteten Empfehlungen. Der Vorgang ist in Abbildung 45 noch einmal grafisch verdeutlicht.

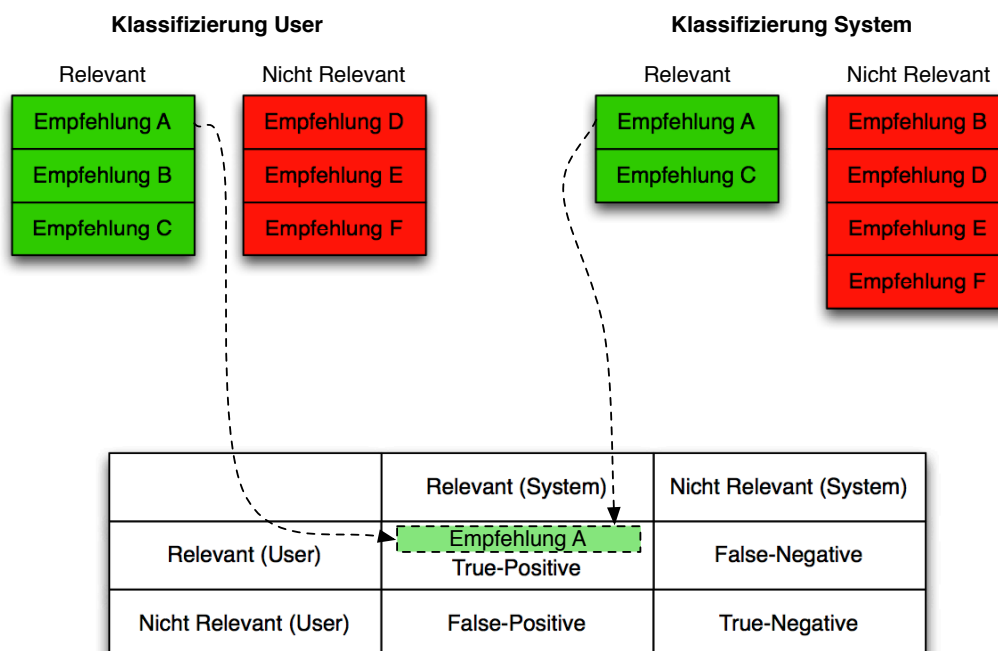


Abbildung 45 Klassifikationsabgleich

Die Beurteilung der Nützlichkeit für Empfehlungen erfolgt über alle durch die Probanden bewerteten inhaltlichen Bereiche. Für die Beurteilung ist eine Betrachtung der Bewertungen in Abhängigkeit der getätigten Klassifizierung durchzuführen, um Zusammenhänge zwischen relevanten bzw. nicht relevanten Bereichen und getätigter Bewertung ausmachen zu können. Die Auswertung des abschließenden Fragebogens kann separat erfolgen. In diesem Zusammenhang ist eine einfache Aggregation der getätigten Aussagen gemäß der Bewertungsskala durchzuführen.

Analog zum Evaluationssetting aus Abschnitt 5.3.2 erfolgt über die Auswertung des Klassifikationsprozesses eine Beurteilung über den Zusammenhang zwischen der Anzahl an Tags und der Anzahl richtig klassifizierter inhaltlicher Bereiche. Wiederum bedarf es der Prüfung, in wieweit eine steigende Anzahl an Tags zu einer besseren Klassifizierung von Ressourcen führt. Ebenso wie in Abschnitt 5.3.2 ist die Nennung eines Schwellwertes das Ziel, der Aussage darüber trifft, ab welcher Anzahl an Tags das Verfahren eine zuverlässige Klassifikation durchführt. Darüber hinaus bedarf es innerhalb der online Simulation der Prüfung zweier weiterer Faktoren, um die Hypothesen aus Abschnitt 5.2 genügend prüfen zu können. Zum einen erfolgt die Prüfung, ob ein Zusammenhang zwischen der bewerteten Nützlichkeit für Empfehlungen und der Anzahl an Tags, auf derer die Herleitung der Empfehlung beruht existiert. In diesem Zusammenhang ist es ebenfalls das Ziel, einen Schwellwert herzuleiten, ab welcher Anzahl an Tags mit einer genügenden Nützlichkeit für User gerechnet werden kann. Des Weiteren bedarf es der Prüfung, ob ein Zusammenhang zwischen einer steigenden Anzahl an alternativen oder verlinkten Lernmaterialien und der bewerteten Nützlichkeit durch die Probanden existiert. In diesem Zusammenhang ist weniger die Nennung eines Schwellwertes das Ziel, als die Herleitung einer allgemeinen Aussage ob ein Zusammenhang existiert oder nicht.

6 Präsentation der Ergebnisse

Das nachfolgende Kapitel präsentiert die erhobenen Ergebnisse der zwei Evaluationsszenarien, die in Abschnitt 5.3 beschrieben wurden. Es erfolgt wiederum eine zweistufige Unterteilung für Ergebnisse der Benutzerstudie (vgl. Abschnitt 6.1) bzw. Ergebnisse aus der online Simulation (vgl. Abschnitt 6.2). Das nachfolgende Kapitel dient zunächst ausschließlich der Präsentation der Ergebnisse ohne dabei auf eine Wertung der Ergebnisse einzugehen. Aus Gründen der Übersicht erfolgt eine ausführliche Interpretation der Ergebnisse im anschließenden Kapitel 7.

6.1 Benutzerstudie unter Laborbedingungen

An der Benutzerstudie nahmen insgesamt 33 Teilnehmer teil, die mit einer durchschnittlichen Dauer von ~67 Minuten pro Proband mit den zur Verfügung gestellten Materialien gearbeitet haben. Die Datenerhebung ergab 7207 implizite Tags, sowie 858 explizite Tags, auf deren Grundlage die Extraktion von insgesamt 53 inhaltlichen Bereichen (Ressourcen) angestoßen wurde. Eine detaillierte Aufstellung aller erhobenen Daten aufgeschlüsselt nach Lernmaterialien ist Tabelle 8 zu entnehmen. Die 53 Ressourcen teilen sich dabei in 43 positive (relevante) und 10 negative (nicht relevante)⁸² Bereiche auf. Insgesamt wurden 45 nicht relevante Bereiche extrahiert (Drops), wobei zehn der negativen Bereiche mit in den Evaluationsprozess aufgenommen wurden. Es sind lediglich solche nicht relevanten Bereiche enthalten, denen als Datenbasis mindestens zwei Tags zu Grunde liegen, da von einer sinnvollen Aussage sonst nicht auszugehen ist. Ein Großteil der extrahierten Ressourcen befindet sich aufgrund der entsprechend hohen Anzahl an expliziten Tags innerhalb des Skriptes. Die bereitgestellten Videos weisen

⁸² In Tabelle 8 sind nicht relevante Bereiche als *Drops (gezählt)* gelistet.

untereinander ein durchweg ähnliches Aufkommen an Tags und Ressourcen auf. Ebenso ist für das Buch eine genügende Aktivität durch die Probanden zur Extraktion inhaltlicher Bereiche festzustellen. Lediglich das Videos Dijkstra APSP fand durch die Probanden eine geringe Berücksichtigung und enthält keine extrahierten Ressourcen.

	Nutzungs- dauer	Implizite Tags	Explizite Tags	Ressourcen	Drops	Drops (gezählt)
Skript	1667	4724	722	26	45	9
Buch	312	1529	38	10	5	1
Video Mergesort	88	212	28	2	7	0
Video Quicksort	84	282	32	3	6	0
Video Quicksort In-Place	31	224	13	1	1	0
Video Dijkstra	42	193	24	1	5	0
Video Dijkstra APSP	9	43	1	0	0	0
Summe	2233	7207	858	43	69	10
Durschnitt pro User	67,67	218,39	26			

Tabelle 8 Auswertung Gesamtübersicht Benutzerstudie

Wie in Abschnitt 5.3.2 dargelegt erfolgt die Bewertung der Klassifikation in Form einer Expertenanalyse. In diesem Zusammenhang muss noch einmal erwähnt werden, dass die Benutzerstudie zunächst einem ersten Test – also einer Art Entwicklertest – des Verfahrens dient, weshalb eine vielmehr qualitative Betrachtung erfolgt. Eine quantitative Analyse ist gemäß Abschnitt 5.3.3 durch die beschriebene online Simulation abgedeckt. Die Bewertung ergibt sich aus der Klassifikationstabelle in Tabelle 9. Die zugrundeliegenden Daten für die Ermittlung der Klassifikationstabelle können Anhang G entnommen werden. Die Klassifikationstabelle zeigt die Klassifizierung von 53 extrahierten inhaltlichen Bereichen mit entsprechender Zuordnung durch das System bzw. durch den Experten (User). Es kann mit 41 (~77%) richtig klassifizierter Bereiche⁸³ eine überwiegend positive Klassifizierung festgestellt werden.

⁸³ Hierzu gehören Bereiche die richtig als relevant (TP) bzw. richtig als nicht relevant (TN) klassifiziert wurden.

	Relevant (System)	Nicht Relevant (System)	
Genutzt (User)	33	2	35
Nicht Genutzt (User)	10	8	18
	43	10	N=53

Tabelle 9 Klassifikationstabelle Benutzerstudie

Zur Errechnung der statistischen Signifikanz bedarf es der Nutzung des exakten Tests von Fisher (vgl. Abschnitt 3.2.5). Gemäß Gleichung (15) auf Seite 60 ergibt sich eine Wahrscheinlichkeit von $p = 0,0014$. Unter Berücksichtigung eines Fehlers von 1% bzw. $\alpha = 0,01$ ist die Bedingung $p < \alpha$ erfüllt, so dass die Nullhypothese abgelehnt werden kann.

In Tabelle 10 ist die Klassifikation aus Tabelle 9 nach Lernmaterialien aufgeschlüsselt. Bedingt durch den größten Anteil an Ressourcen innerhalb des Skriptes zeigt Tabelle 10 für dieses Lernmaterial die größte Aktivität. Über die Klassifikationstabelle in Tabelle 9 erfolgt die Berechnung von Precision und Recall zur Bewertung der Nutzungswahrscheinlichkeit in den Gleichungen (45) und (46):

$$Precision = \frac{33}{33 + 10} = \frac{33}{43} = 0.76744186 \quad (45)$$

$$Recall = \frac{33}{33 + 2} = \frac{33}{35} = 0.9428571429 \quad (46)$$

Das aggregierte F_1 erfolgt in Gleichung (47).

$$F_1 = \frac{2 * 0,76744186 * 0,9428571429}{(0,76744186 + 0,9428571429)} = \frac{1,447176079}{1,710299003} = 0,8461538459 \quad (47)$$

Skript	Relevant (System)	N. Relevant (System)	Summe
Relevant (User)	20	1	21
N. Relevant (User)	6	8	14
Summe	26	9	35
Buch			
Relevant (User)	7	1	8
N. Relevant (User)	3	0	3
Summe	10	1	11
Mergesort (V)			
Relevant (User)	2	0	0
N. Relevant (User)	0	0	0
Summe	2	0	2
Quicksort (V)			
Relevant (User)	2	0	2
N. Relevant (User)	1	0	1
Summe	3	0	3
Quicksort-IP (V)			
Relevant (User)	1	0	0
N. Relevant (User)	0	0	0
Summe	1	1	1
Dijkstra (V)			
Relevant (User)	1	0	1
N. Relevant (User)	0	0	0
Summe	1	0	1
APSP (V)			
Relevant (User)	0	0	0
N. Relevant (User)	0	0	0
Summe	0	0	0

Tabelle 10 Klassifikationstabelle nach Material Benutzerstudie

Abbildung 46 zeigt die Klassifikation von TP bzw. FP in Abhängigkeit der Anzahl an Tags, die für eine Resource geclustert wurden. Abbildung 47 stellt gleiches für die Klassifikation von TN bzw. FN dar. Beide Abbildungen sind durch die Daten aus Anhang G abgeleitet worden. In Abbildung 46 zeigt sich unterhalb von 6-7 Tags ein häufiger Wechsel zwischen richtig positiv

und falsch positiv klassifizierten inhaltlichen Bereichen. Ab der Anzahl von 7 Tags ist innerhalb der Benutzerstudie eine durchweg richtige Einordnung von relevanten Bereichen festzustellen.

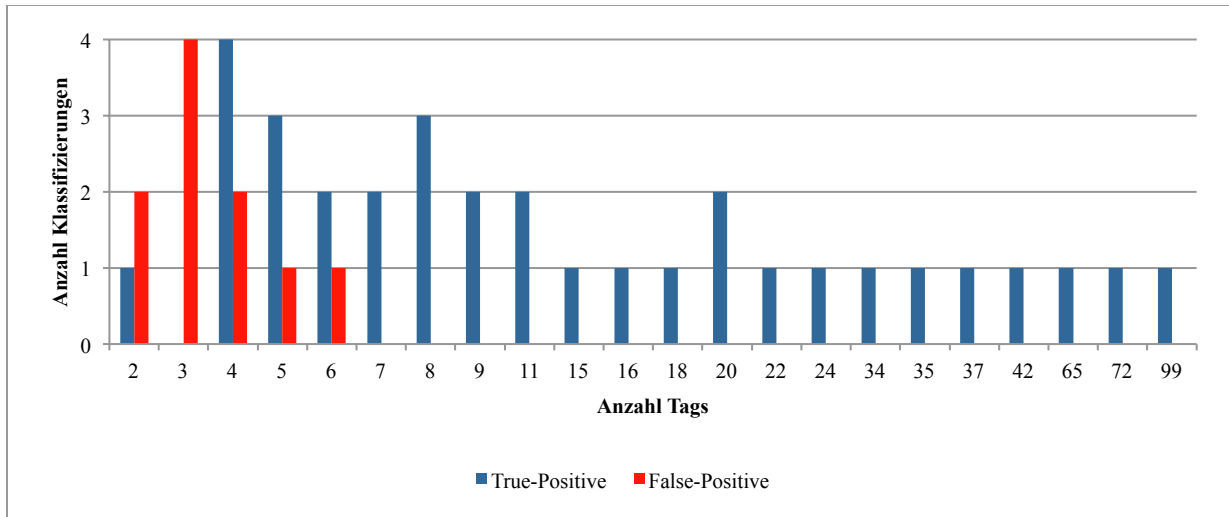


Abbildung 46 Klassifikation (TP, FP) in Abhängigkeit Anzahl Tags Benutzerstudie

Eine deutliche Tendenz für richtig falsch klassifizierte Bereiche kann in Abbildung 47 nicht gesehen werden. Dies liegt zum einen an der ohnehin hohen Genauigkeit bei der Einordnung falscher Bereiche, als auch an der recht geringen Anzahl an falsch klassifizierten Bereichen, die innerhalb der Klassifikation Berücksichtigung fanden. Weitere Feststellungen bedürfen der Betrachtung der Ergebnisse aus dem zweiten Evaluationsszenario.



Abbildung 47 Klassifikation (TN, FN) in Abhängigkeit Anzahl Tags Benutzerstudie

6.2 Online Simulation in realem Kursumfeld

Für die online Simulation wurden insgesamt 90 Teilnehmer eingeladen, wovon 56 tatsächlich im System angemeldet waren und eine aktive Teilnahme aufwiesen. Nachfolgend ist deshalb von 56 Teilnehmern auszugehen. Die 56 Teilnehmer können dabei in 33 aktive Teilnehmer und 23 Teilnehmer mit geringfügiger Aktivität untergliedert werden. Ein Teilnehmer gilt als aktiv, sofern er oder sie mehr als zehn explizite Tags im System gesetzt hat. Grundlegend ist die Untergliederung auch anhand der Nutzungsdauer denkbar, allerdings spiegelt die Nutzungsdauer nicht zwangsläufig ein aktives Vorgehen im System wieder. So ist eine Klassifizierung nach Tags als sinnvoller einzuschätzen. Die Nutzungsdauer im System für alle Teilnehmer betrug insgesamt 49215 Minuten (~8202 Stunden). Der Empfehlungsprozess – inklusive der Extraktion inhaltlich relevanter Bereiche – wurde auf Basis von 9772 impliziten, sowie 1369 expliziten Tags durchgeführt. Das Verfahren hat auf dieser Grundlage 122 relevante und 154 nicht relevante Bereiche (Drops) inklusive entsprechender Empfehlung generiert. Von den nicht relevanten Bereichen wurden aktiv elf in den Klassifikationsschritt der Evaluation mit aufgenommen⁸⁴ (vgl. in diesem Zusammenhang die Auswahlkriterien zur Aufnahme nicht relevanter Bereiche in Abschnitt 6.1). Für die 133 Bereiche wurde in 256 Fällen eine Klassifikation durchgeführt und es wurden insgesamt 324 Bewertungen hinsichtlich der Nützlichkeit durch die Probanden abgegeben. Ein Großteil der Ressourcen wurde innerhalb des Skriptes hergeleitet (vgl. Tabelle 11), weshalb sich eine Vielzahl an Klassifizierungen bzw. Bewertungen für das Skript ergibt. Die Vielzahl an Ressourcen spiegelt sich für das Skript auch in der Anzahl an expliziten Bewertungen und der Nutzungsdauer wider. Lediglich das Übungsblatt weist eine insgesamt höhere Nutzungsdauer als das Skript auf. Tabelle 11 ist eine Gesamtübersicht der Auswertung für die online Simulation nach Lernmaterialien zu entnehmen. Aus Gründen der Übersicht zeigt die Tabelle allerdings nur solche Lernmaterialien, innerhalb derer tatsächlich Empfehlungen enthalten sind. Eine Übersicht für alle Lernmaterialien kann Anhang H entnommen werden.

⁸⁴ In Tabelle 11 als *Drops (gezählt)* gelistet

	Nutzungsdauer (Minuten)	Implizite Tags	Implizite Tags (Datenanalyse)	Explizite Tags	Ressourcen	Bewertungen	Drops	Drops (gezählt)
Skript	6749	5590	5221	988	105	198	78	11
Worte	1478	290	217	5	2	12	1	0
Fehlende Zeilen (vgl. Anhang H für vollständige Auflistung)								
Automatenminimierung	314	340	255	10	1	3	2	0
Minimierung endlicher Automaten - Beispiel	1213	432	382	34	2	8	7	0
Fehlende Zeilen (vgl. Anhang H für vollständige Auflistung)								
Idee CYK-Algorithmus	2454	605	428	61	3	27	14	0
Anwendungsbeispiel CYK-Algorithmus	2155	1066	682	67	4	22	12	0
Übungsklausur	21995	852	769	123	5	54	7	0
Summe (gesamt)	49215	13484	9772	1369	122	324	154	11
Durchschnitt pro User (gesamt)	878,84	240,79	174,5	24,45		12,96		

Tabelle 11 Auswertung Gesamtübersicht Online Simulation

Die Klassifikation für 256 inhaltliche Bereiche ist gemäß Abschnitt 3.2.5 bzw. Abschnitt 5.3.3 in Tabelle 12 gezeigt. Die Ermittlung der Klassifikationstabelle ergibt sich aus den Daten in Anhang I. Mit insgesamt 220 (~86%) korrekt (positiv, negativ) klassifizierten Bereichen zeigt sich innerhalb der online Simulation ein positiver Trend, der mit den Ergebnissen aus dem ersten Evaluationsszenario grundlegend übereinstimmt.

	Relevant (System)	Nicht Relevant (System)	
Genutzt (User)	204	2	206
Nicht Genutzt (User)	34	16	50
	238	18	N=256

Tabelle 12 Klassifikationstabelle Online Simulation

Zur Errechnung der statistischen Signifikanz bedarf es der Nutzung des exakten Tests von Fisher (vgl. Abschnitt 3.2.5). Gleichung (15) auf Seite 60 ergibt sich eine Wahrscheinlichkeit von $p < 0,0001$. Unter Berücksichtigung eines Fehlers von 1% bzw. $\alpha = 0,01$ ist die Bedingung $p < \alpha$ erfüllt, so dass die Nullhypothese abgelehnt werden kann.

Eine nach Lernmaterialien unterteilte Klassifikation ist Tabelle 13 zu entnehmen. Die hohe Nutzung des Skriptes durch die Probanden spiegelt sich ebenfalls im Kontext der Klassifizierung wider. Ein bedeutend hoher Teil an Bereichen wurde innerhalb des Skriptes durch die Probanden klassifiziert. Auch innerhalb des Übungsblattes und in den CYK Videos hat eine aktive

Klassifizierung stattgefunden. Für die Videos Worte und Automatenminimierung ergibt sich lediglich eine geringe Klassifizierungsaktivität.

Skript	Relevant (System)	N. Relevant (System)	Summe
Relevant (User)	124	2	126
N. Relevant (User)	23	16	39
Summe	147	18	165
Worte			
Relevant (User)	7	0	7
N. Relevant (User)	2	0	2
Summe	9	0	9
Automatenminimierung			
Relevant (User)	3	0	3
N. Relevant (User)	0	0	0
Summe	3	0	3
Automatenminimierung BSP			
Relevant (User)	6	0	6
N. Relevant (User)	0	0	0
Summe	6	0	6
CYK Idee			
Relevant (User)	19	0	19
N. Relevant (User)	4	0	4
Summe	23	0	23
CYK Beispiel			
Relevant (User)	14	0	14
N. Relevant (User)	3	0	3
Summe	17	0	17
Übungsblatt			
Relevant (User)	31	0	31
N. Relevant (User)	2	0	2
Summe	33	0	33

Tabelle 13 Klassifikationstabelle nach Material Online Studie

Über Tabelle 12 ergibt sich die Berechnung von Precision und Recall in den Gleichungen (48) und (49).

$$Precision = \frac{204}{204 + 34} = \frac{204}{238} = 0,857142857 \quad (48)$$

$$Recall = \frac{204}{204 + 2} = \frac{204}{206} = 0,990291262 \quad (49)$$

Das aggregierte F_1 -Maß ergibt sich aus Gleichung (50).

$$F_1 = \frac{2 * 0,857142857 * 0,990291262}{(0,857142857 + 0,990291262)} = \frac{1,697642163}{1,847434119} = 0,918918919 \quad (50)$$

Abbildung 48 zeigt die Anzahl richtig bzw. falsch durch das System als relevant klassifizierten inhaltlichen Bereichen in Abhängigkeit der Anzahl an Tags, die der Extraktion des jeweiligen Bereichs zugrunde liegen. Abbildung 49 stellt den Zusammenhang analog für als nicht relevant klassifizierte Bereiche dar. Die Ermittlung erfolgt auf Grundlage der Auswertung in Anhang I. In Abbildung 48 ist anders als im ersten Evaluationsdurchlauf keine klare Grenze zu erkennen, ab der eine durchweg richtige Klassifikation durch das System stattfindet. Auffällig ist allerdings die Abnahme an falsch positiv klassifizierten Bereichen ebenfalls ab einer Grenze von ~7 Tags. Deutlicher auszumachen ist allerdings die starke Zunahme von richtig positiv klassifizierten Bereichen.

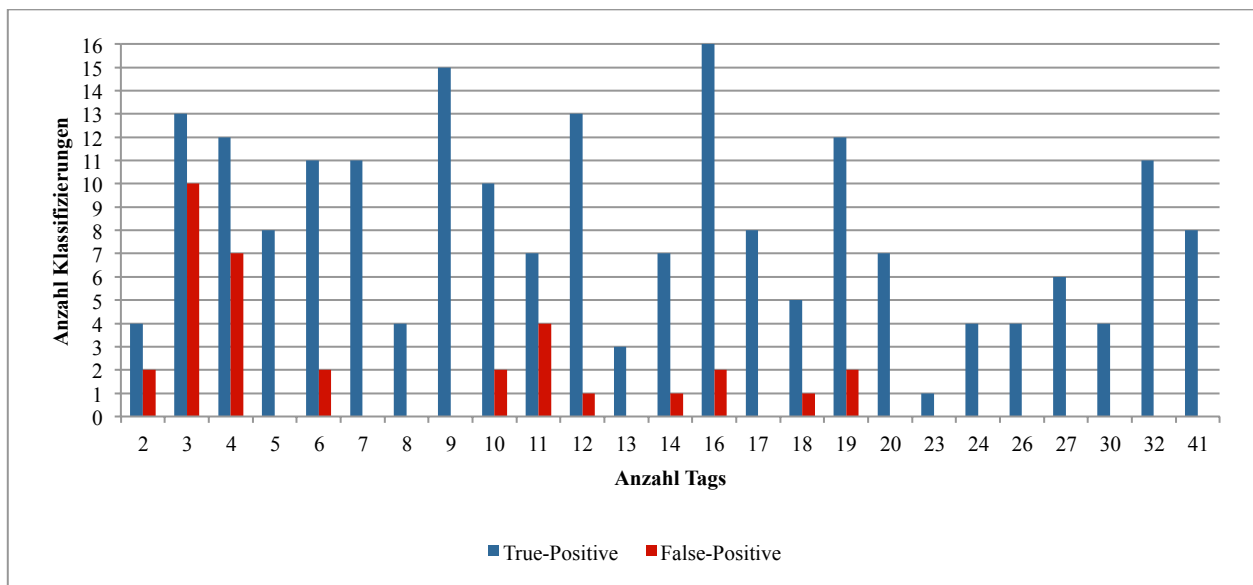


Abbildung 48 Klassifikation (TP, FP) in Abhängigkeit Anzahl Tags online Simulation

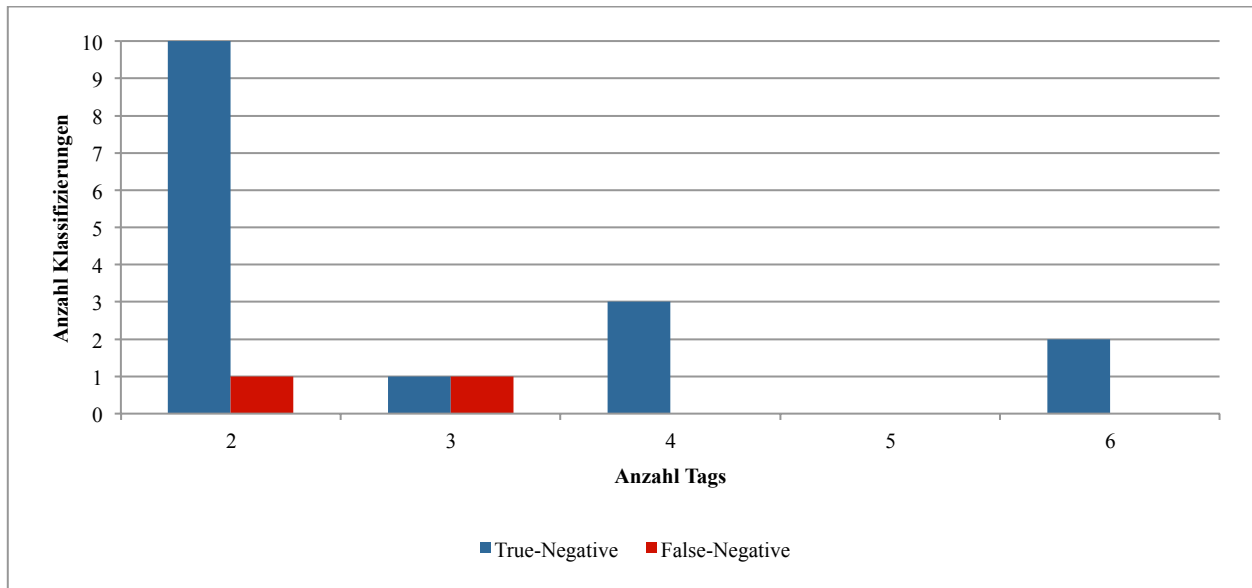


Abbildung 49 Klassifikation (TN, FN) in Abhängigkeit Anzahl Tags online Simulation

Eine Gesamtübersicht für die Bewertung der Nützlichkeit der Empfehlungen ist in Tabelle 14 gezeigt und als aggregiertes Ergebnis in Form des Mittelwertes in Abbildung 50 illustriert. In der Spalte „Ang.“ in Tabelle 14 sind solche Ereignisse angegeben, in denen eine Empfehlung lediglich angesehen, aber nicht bewertet wurde. Ähnlich wie für den Schritt der Klassifizierung, ergibt sich vor allem für das Skript, aber auch für das Übungsblatt und die CYK Videos, ein aktives Bewertungsverhalten durch die Probanden.

Bewertung Gesamt

	5 Sterne	4 Sterne	3 Sterne	2 Sterne	1 Sterne	Ang.	Summe
Skript	51	28	38	8	32	41	198
Worte	3	0	3	0	3	3	12
Automatenminimierung	3	0	0	0	0	0	3
Automatenminimierung (Beispiel)	3	0	2	0	0	3	8
CYK Idee	9	6	4	1	4	3	27
CYK Beispiel	4	5	5	1	1	6	22
Übungsblatt	19	16	5	1	1	12	54
Summe	92	55	57	11	41	68	324
Summe %	28,40	16,98	17,60	3,40	12,65	20,99	

Tabelle 14 Bewertung Gesamtübersicht Online Simulation

Abbildung 50 zeigt das aggregierte Mittel der Bewertungen ausgehend von den Ergebnissen in Tabelle 14. Ein überwiegender Teil der Empfehlungen wurde mit 5 Sternen bewertet. Zu etwa gleichen Teilen sind Empfehlungen mit 4- oder 3-Sternen bewertet. 2- und 1-Sterne Bewertungen sind zusammengefasst weniger häufig vergeben worden als die übrigen Bewertungen. Das Diagramm zeigt auch, dass die Probanden Empfehlungen häufig angesehen, aber nicht zwangshäufig bewertet haben.

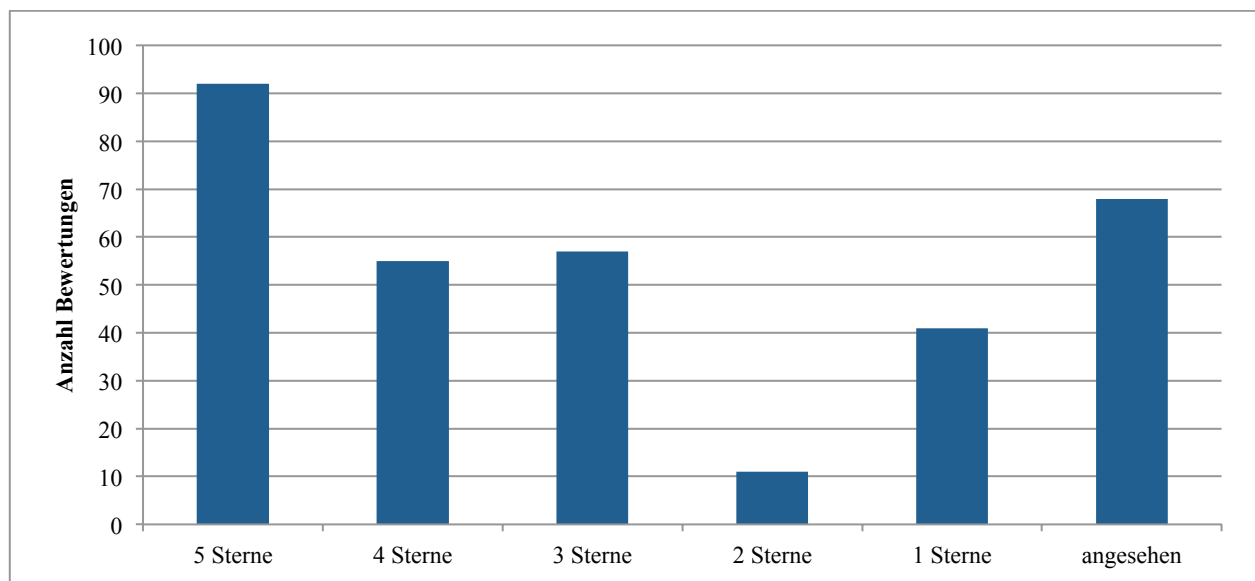


Abbildung 50 Auswertung Bewertungen Aggregiert

Für eine bessere Interpretation enthalten die nachfolgenden Tabelle 15 bis Tabelle 19 eine Aufschlüsselung der Bewertungen hinsichtlich der Nützlichkeit gemäß der jeweils durchgeführten Klassifizierung des inhaltlichen Bereichs. So ergeben sich insgesamt fünf unterschiedliche Betrachtungsweisen für jede Möglichkeit der Klassifikationstabelle (True-Positive, True-Negative, False-Positive, False-Negative), sowie für solche Empfehlungen, die durch den Probanden nicht klassifiziert wurden. Für eine übersichtlichere Betrachtung sind die Ergebnisse aus den nachfolgenden Tabellen noch einmal in Abbildung 51 illustriert.

True Positives

	5 Sterne	4 Sterne	3 Sterne	2 Sterne	1 Sterne	Ang.	Summe
Skript	50	28	31	0	9	6	124
Worte	3	0	3	0	1	0	7
Automatenminimierung	3	0	0	0	0	0	3
Automatenminimierung (Beispiel)	3	0	2	0	0	0	5
CYK Idee	9	6	2	0	2	1	20
CYK Beispiel	4	5	5	0	0	0	14
Übungsblatt	14	12	4	1	0	0	31
Summe	86	51	47	1	12	7	204
Summe %	42,16	25	23,04	0,50	5,88	3,43	

Tabelle 15 Auswertung Bewertungen in Abhängigkeit True-Positives online Simulation

True Negatives

	5 Sterne	4 Sterne	3 Sterne	2 Sterne	1 Sterne	Ang.	Summe
Skript	0	0	0	5	10	1	16
Worte	0	0	0	0	0	0	0
Automatenminimierung	0	0	0	0	0	0	0
Automatenminimierung (Beispiel)	0	0	0	0	0	0	0
CYK Idee	0	0	0	0	0	0	0
CYK Beispiel	0	0	0	0	0	0	0
Übungsblatt	0	0	0	0	0	0	0
Summe	0	0	0	5	10	1	16
Summe %	0	0	0	31,25	62,5	6,25	

Tabelle 16 Auswertung Bewertungen in Abhängigkeit True-Negatives online Simulation

False Positive

	5 Sterne	4 Sterne	3 Sterne	2 Sterne	1 Sterne	Ang.	Summe
Skript	0	0	5	3	13	2	23
Worte	0	0	0	0	2	0	2
Automatenminimierung	0	0	0	0	0	0	0
Automatenminimierung (Beispiel)	0	0	0	0	0	0	0
CYK Idee	0	0	1	1	2	0	4
CYK Beispiel	0	0	0	1	1	1	3
Übungsblatt	0	0	1	0	1	0	2
Summe	0	0	7	5	19	3	34
Summe %	0	0	20,59	14,71	55,88	8,82	

Tabelle 17 Auswertung Bewertungen in Abhängigkeit False-Positives online Simulation

False Negative

	5 Sterne	4 Sterne	3 Sterne	2 Sterne	1 Sterne	Ang.	Summe
Skript	1	0	1	0	0	0	2
Worte	0	0	0	0	0	0	0
Automatenminimierung	0	0	0	0	0	0	0
Automatenminimierung (Beispiel)	0	0	0	0	0	0	0
CYK Idee	0	0	0	0	0	0	0
CYK Beispiel	0	0	0	0	0	0	0
Übungsblatt	0	0	0	0	0	0	0
Summe	1	0	1	0	0	0	2
Summe %	50	0	50	0	0	0	

Tabelle 18 Auswertung Bewertungen in Abhängigkeit False-Negatives online Simulation

Ohne Klassifikation

	5 Sterne	4 Sterne	3 Sterne	2 Sterne	1 Sterne	Ang.	Summe
Skript	0	0	1	0	0	32	33
Worte	0	0	0	0	0	3	3
Automatenminimierung	0	0	0	0	0	0	0
Automatenminimierung (Beispiel)	0	0	0	0	0	2	2
CYK Idee	0	0	1	0	0	3	4
CYK Beispiel	0	0	0	0	0	5	5
Übungsblatt	5	4	0	0	0	12	21
Summe	5	4	2	0	0	57	68
Summe %	7,35	5,89	2,94	0	0	83,82	

Tabelle 19 Auswertung Bewertungen in Abhängigkeit keiner Klassifikation online Simulation

Abbildung 51 zeigt die Anzahl von Bewertungen je Kategorie in Abhängigkeit der jeweiligen Klassifikation. Die Form der Klassifikation ist dabei farblich unterschieden. Für 5-, 4-, sowie 3-Sterne Bewertungen ergeben sich mehrheitlich TP-klassifizierte Bereiche. 2-Sterne Bewertungen enthalten zu gleichen Teilen TN und FP Klassifikationen. 1-Sterne Bewertungen ergeben sich aus einer Hälfte TN und jeweils zu einem Viertel aus TN, sowie TP Klassifikationen. Nicht bewertete Empfehlungen sind zu einem überwiegenden Anteil nicht klassifiziert worden. Wie aus den vorherigen Tabellen, sowie Abbildung 51 zu entnehmen ist, ergibt sich für korrekt positiv klassifizierte inhaltliche Bereiche inklusive ihrer Empfehlung in 90% der Fälle eine Probandenbewertung von 3 Sternen oder besser. Für richtig falsch klassifizierte Bereiche ergibt sich in ~94% der Fälle eine Probandenbewertung von 2-Sternen oder 1-Stern. Für falsch positiv klassifizierte Bereiche inklusive entsprechender Empfehlung ergibt sich in ~91% der Fälle eine Bewertung von 3-Sternen oder schlechter. Für nicht klassifizierte Bereiche ergibt sich in vier von fünf Fällen keine Bewertung.

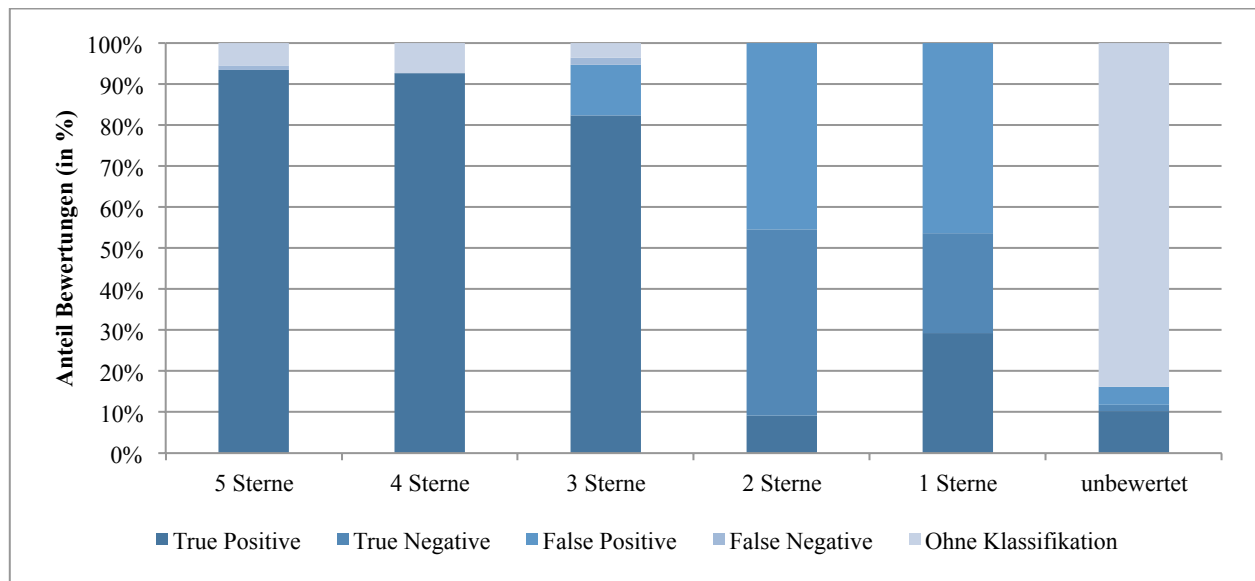


Abbildung 51 Aggregierte Auswertung der Bewertungen in Abhängigkeit der Klassifikation (online Simulation)

Abbildung 52 zeigt die aggregierte Bewertung der Probanden in Abhängigkeit der Anzahl an Tags, auf deren Grundlage ein inhaltlicher Bereich extrahiert bzw. eine Empfehlung abgeleitet wurde. Das Diagramm zeigt lediglich die durch das System als relevant klassifizierten Bereiche, da eine geringe Bewertung für nicht relevante Bereiche ohnehin erwartet wird und sich die Darstellung als weniger aussagekräftig darstellen würde. Die Trendlinie im Diagramm zeigt eine positive Entwicklung der Bewertungen mit ansteigender Anzahl an Tags. Das Punktdiagramm in Abbildung 53 zeigt den Trend aus einer weiteren, anschaulicheren Perspektive, wo die mittlere Bewertung für jede Resource in Abhängigkeit der Anzahl an Tags aufgetragen ist. In Abbildung 53 ist eine positive Trendwende ab ~5 Tags auszumachen ist.

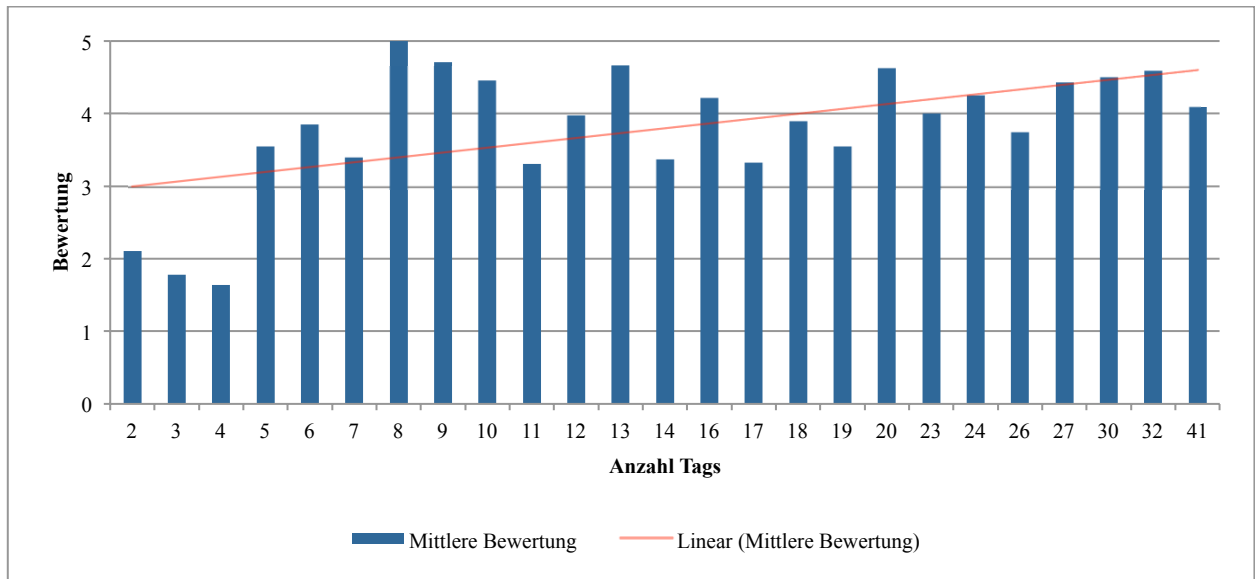


Abbildung 52 Bewertungen in Abhängigkeit Anzahl an Tags online Simulation

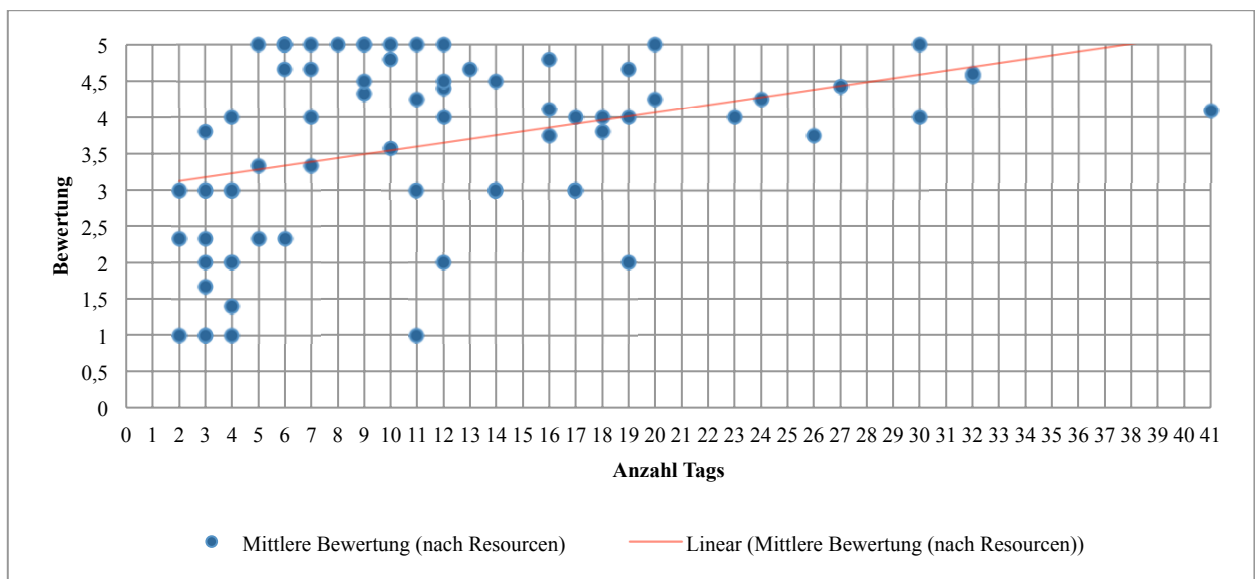


Abbildung 53 Bewertungen je Resource in Abhängigkeit Anzahl an Tags online Simulation

In Abbildung 54 sind die Probandenbewertungen für Empfehlungen in Abhängigkeit von alternativ zur Verfügung gestellten Lernmaterialien bzw. Materialverknüpfungen dargestellt. Für eine sinnvolle Darstellung zeigt das Diagramm wiederum nur solche Bewertungen an, die für Bereiche mit einer durch das System als relevant klassifizierten Zuordnung existieren. Für Empfehlungen mit alternativen Materialien oder Materialverknüpfungen ergibt sich so eine durchschnittliche Bewertung von ~4,5 Sternen, wohingegen sich für Empfehlungen ohne etwaigen alternativen Verweis eine durchschnittliche Bewertung von ~3 Sternen ergibt.

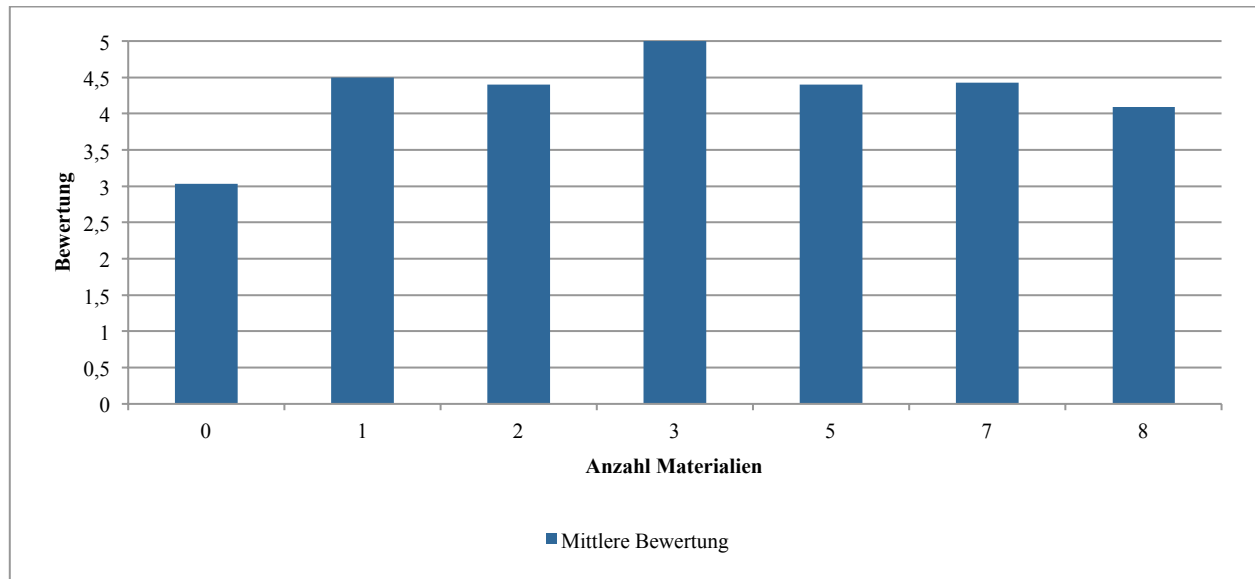


Abbildung 54 Bewertung in Abhängigkeit Anzahl alternativer Lernmaterialien online Simulation

Die aggregierte Auswertung zur Akzeptanzbefragung ist in Abbildung 55 in Form des aggregierten Mittels präsentiert⁸⁵. Die Ergebnisse sind aus Gründen der Übersicht gemäß ihrer Fragen-ID präsentiert (vgl. Fragebogen in Tabelle 7 auf Seite 154). Die Darstellung basiert auf der Auswertung aus Anhang J. Die Auswertung der Bewertungsskala des Fragebogens von „trifft voll zu“ bis hin zu „trifft gar nicht zu“ ist zur besseren Analyse in eine Dezimalskala von 5 („trifft voll zu“) bis 1 („trifft gar nicht zu“) überführt worden. Positiv gestellte Fragen erhielten, abgesehen von den Fragen 1,3,4 und 14, eine Bewertung von mindestens 3,5 oder besser. Die Fragen 1,3,4 und 14 sind durchweg mit besser als 3 beantwortet worden. Eine Wertung über 4 ergibt sich für die Fragen 2,8 und 9. Für die negativ gestellten Fragen 17 und 18 ergibt sich eine mittlere Bewertung von jeweils unter 2,5.

⁸⁵ Negativ gestellte Fragen sind rot gekennzeichnet.

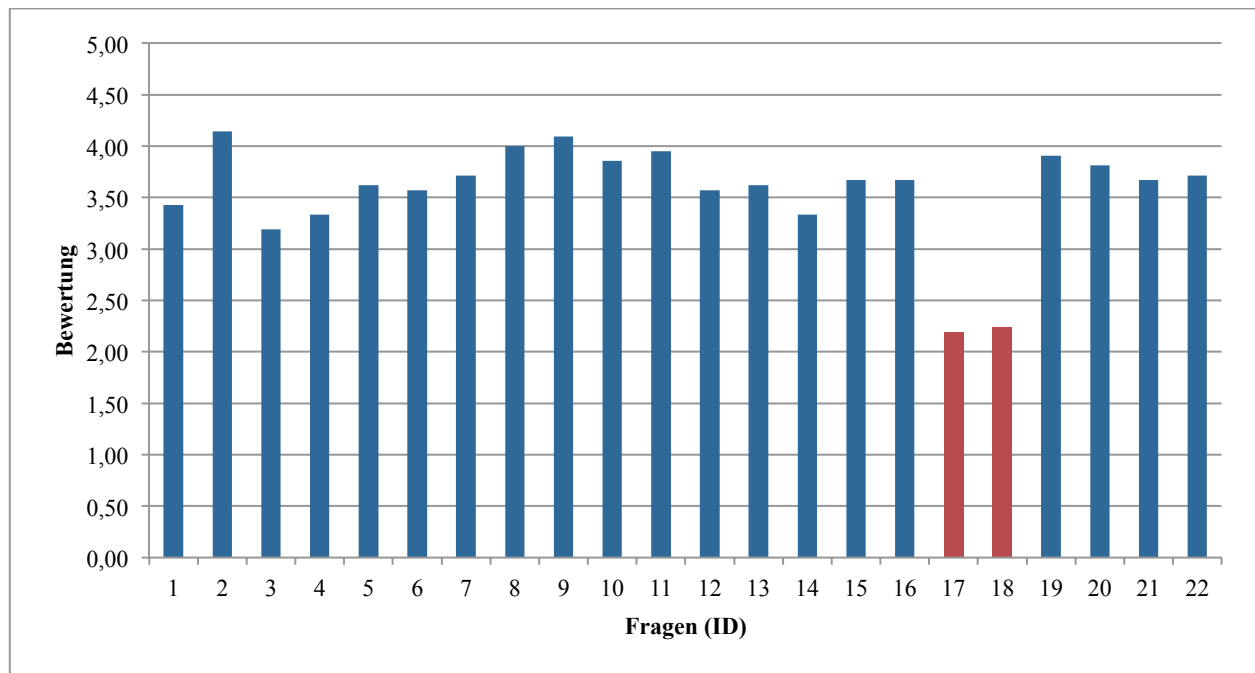


Abbildung 55 Auswertung Akzeptanzbefragung Aggregiert

Befragungen gemäß des Technology Acceptance Models lassen sich in die Kategorien Qualität, Aufwand, sowie Akzeptanz im Allgemeinen untergliedern (vgl. Abschnitt 3.2.5 bzw. 5.3.3). Für die Auswertung in der vorliegenden Arbeit sind die Fragen deshalb aufgeschlüsselt gemäß der zugrunde liegenden Kategorie in Abbildung 56, Abbildung 57 und Abbildung 58 präsentiert. Für jede Frage ist der prozentuale Anteil der jeweiligen Bewertungskategorie mit angegeben. Für alle positiv gestellten Fragen ergibt sich ein Anteil von ~85% für eine Bewertung von mindestens drei oder besser. Hierüber kann eine durchweg positive Tendenz für alle positiv gestellten Fragen festgestellt werden. Dieser Trend ist in gleicher Form auch der Abbildung 55 zu entnehmen. Für die beiden negativ gestellten Fragen (17, 18) ergibt sich ein Anteil von jeweils ~65% für eine Bewertung von zwei oder schlechter.

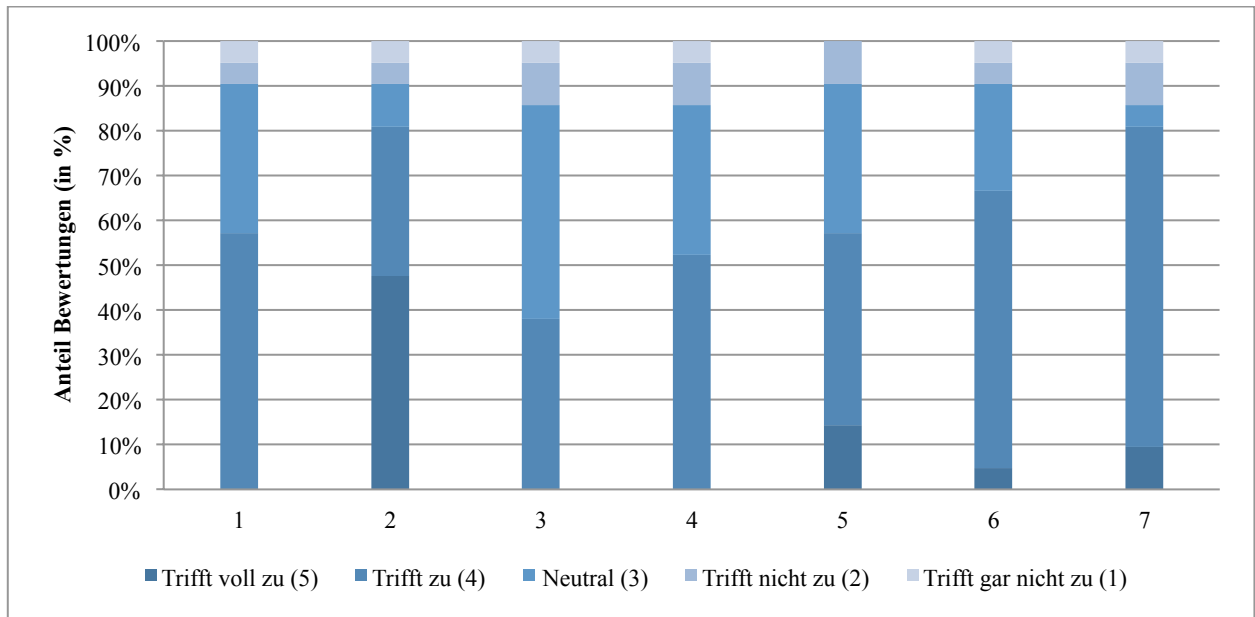


Abbildung 56 Auswertung Akzeptanzbefragung (Kategorie Qualität)

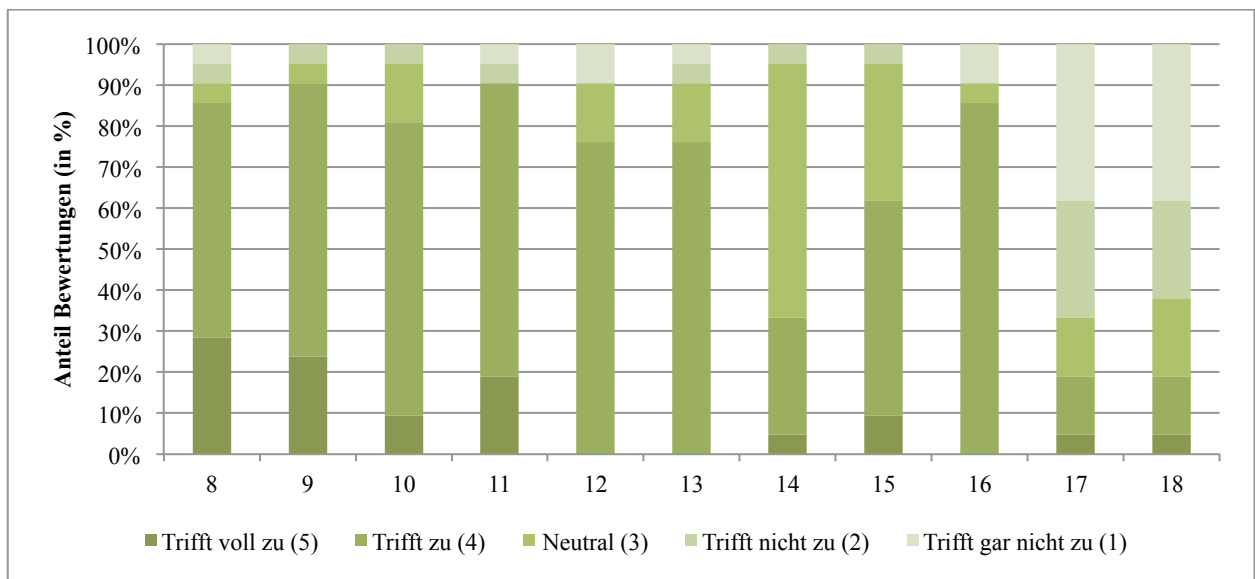


Abbildung 57 Auswertung Akzeptanzbefragung (Kategorie Aufwand)

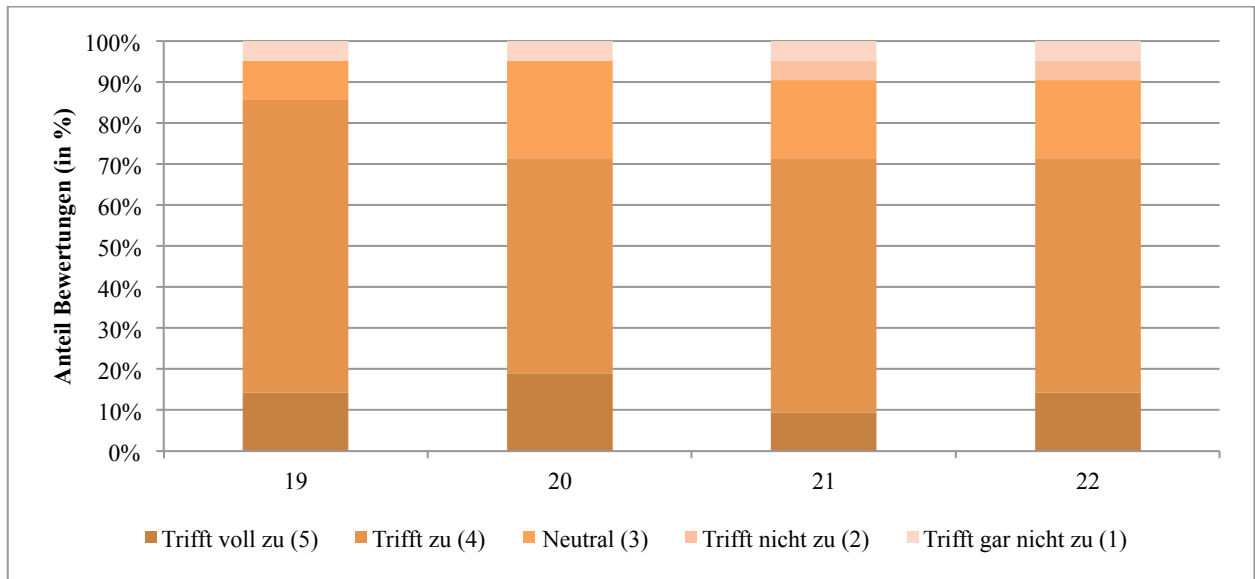


Abbildung 58 Auswertung Akzeptanzbefragung (Kategorie Akzeptanz allgemein)

7 Interpretation der Evaluationsergebnisse

Das nachfolgende Kapitel dient der Interpretation der Evaluationsergebnisse aus Kapitel 6 und zur entsprechenden Beantwortung der Forschungsfragen bzw. Hypothesen aus Abschnitt 5.2.

Zunächst seien die Erhebungen der Evaluation mit der Voruntersuchung aus Kapitel 2 in Zusammenhang gebracht. Die Evaluationsergebnisse beider Szenarien zeigen hinsichtlich des Nutzungsverhaltens durchweg ähnliche Ergebnisse wie die Voruntersuchung, was die Annahmen aus Kapitel 2 noch einmal stärkt. Das Skript steht für die Lernenden in beiden Evaluationsszenarien bei der Nutzung im Vordergrund. Dies scheint angesichts der ergänzenden Rolle der Videos in beiden Szenarien allerdings auch plausibel. Man sieht allerdings im zweiten Evaluationsszenario, dass Videos die eine Relevanz für das Übungsblatt besitzen, im Vergleich zu den übrigen Videos eine höhere Nutzung aufweisen. Dies bestätigt die angedachte Rolle der Videos. Interessant ist allerdings die Tatsache, dass im zweiten Evaluationsszenario die Nutzungsdauer des Übungsblattes dominiert. Das Ergebnis muss insofern relativiert werden, als dass das Übungsblatt einen zentralen Ausgangspunkt für das Lernen darstellt. Ohne die Nutzung des Übungsblattes kann demzufolge keine Aktivität von Lernenden ausgeübt werden. Die extrem hohe Nutzung des Übungsblattes ergibt sich aus zwei Annahmen. Zum einen kann davon ausgegangen, dass einige Lernende das Übungsblatt parallel zum Lernmaterial geöffnet hatten (z.B. über einen zweiten Tab im Browser). Somit ergibt sich eine grundlegend dauerhafte Aktivität für das Übungsblatt. Des Weiteren kann davon ausgegangen werden, dass einige Lernende lediglich das Übungsblatt aufgerufen haben, allerdings offline – also z.B. mit dem ausgedruckten Skript – gearbeitet haben. Interessant ist allerdings die Tatsache, dass das Übungsblatt ein hohes Aufkommen an expliziten Tags, sowie Materialverknüpfungen aufweist. Betrachtet man die hergeleiteten Empfehlungen des Systems, so ergibt sich für jede Übung eine

eigene Empfehlung, die jeweils Materialverbindungen aufweist. Das Vorgehen der Lernenden lässt die Vermutung zu, dass das Skript zwar als zentrales Lernmedium gesehen wird, allerdings der Wunsch nach inhaltlichen Verknüpfungen innerhalb des Übungsblattes existiert. Dies entspricht einem eher pragmatischen und lösungsorientierten Vorgehen, was in ähnlicher Weise bereits in Kapitel 2 angenommen wurde.

Gemäß des Evaluationsdesigns in Abschnitt 5.3 bedarf es der Untersuchung des Systems in Abhängigkeit des Empfehlungsprozesses. So sind zunächst die Ergebnisse im Zusammenhang der relevanten inhaltlichen Bereiche zu betrachten. Gemäß der Vorgabe ermitteln beide Evaluationsszenarien eine Güte hinsichtlich der Klassifizierung in relevante bzw. nicht relevante Bereiche, die durch das System ermittelt und durch die Probandengruppe(n) bewertet wurden. Ausgedrückt durch die ermittelten Maße Precision und Recall bzw. über das aggregierte F1-Maß zeigt sich zunächst eine positive Klassifizierungsquote in beiden Szenarien. Vor allem beide Recall Werte mit $\sim 0,9429$ und $\sim 0,9903$ zeigen ein sehr hohes Niveau für richtig negativ klassifizierte Bereiche. Allerdings muss die Aussagekraft an dieser Stelle ein Stück weit eingegrenzt werden. Das Verhältnis zu bewertender positiver und negativer Bereiche ist deutlich unausgewogen und die Anzahl an negativ klassifizierten Bereichen nimmt keinen Einfluss auf die Errechnung von Precision und Recall. Betrachtet man an dieser Stelle die True Negative Rate, die sich mit

$$\text{True Negative Rate} = \frac{TN}{TN + FP} \quad (51)$$

über die Anzahl der richtig negativ klassifizierten Bereiche und der Gesamtanzahl an negativ klassifizierten Bereichen ergibt, so kann mit 0,8 für das erste Szenario und $\sim 0,8888$ für das zweite Szenario ein immer noch positives, aber dennoch besser einzuordnendes Ergebnis präsentiert werden. Die erhobenen Precision Werte dienen an dieser Stelle als Argumentationsgrundlage. Das Klassifikationsniveau ist sowohl für Text-, als auch für Videodokumente beständig positiv, so dass sich keine Unterschiede zwischen den Materialformen ausmachen lassen. Im Zusammenhang mit der automatisierten Klassifikation inhaltlicher Bereiche ergibt sich die Hypothese H_1 aus Abschnitt 5.2 wie folgt:

Hypothese (H_1): Das System extrahiert innerhalb eines bestimmten Lernkontextes relevante inhaltliche Bereiche, für die eine positiv erkennbare Nutzungswahrscheinlichkeit seitens der Nutzer existiert.

Die erhobenen Ergebnisse zeigen, dass ein klar positiver Trend hinsichtlich der Nutzungswahrscheinlichkeit für die durch das System klassifizierten Bereiche vorliegt. Mit einer Genauigkeit von ~86% im zweiten Evaluationsszenario zeigt dies eine klar positive Nutzungswahrscheinlichkeit auch aus Sicht der Probanden. Betrachtet man zudem Frage 2 aus der Akzeptanzbefragung, die hinterfragt, in wie weit die Probanden die Bereiche der Empfehlungen ebenfalls markiert hätten, zeigt sich bei expliziter Nachfrage mit einer Bewertung von über 4 eine klar positive Beurteilung durch die Probanden. Gemäß der Begründung wird die Hypothese H_1 auf Grundlage der Evaluationsergebnisse angenommen. Im Zusammenhang mit der Nutzungswahrscheinlichkeit hinterfragt Hypothese H_3 in wie weit die Anzahl an Tags Einfluss auf die Nutzungswahrscheinlichkeit besitzt.

Hypothese (H_3): Es existiert ein Zusammenhang zwischen der Anzahl an Tags, die für die Extraktion eines inhaltlichen Bereichs Verwendung finden und der ermittelten Nutzungswahrscheinlichkeit des jeweiligen Bereiches.

Die Evaluationsergebnisse aus beiden Szenarien zeigen deutlich, dass eine klare Abhängigkeit zwischen der Datenmenge und der richtigen Klassifizierung von relevanten Bereichen existiert. Innerhalb der Benutzerstudie ist ab einer Anzahl von 7-8 Tags eine durchweg richtige Klassifizierung auszumachen. Die Ergebnisse der online Simulation zeigen zwar keine Eliminierung falsch klassifizierter Bereiche ab einem bestimmten Punkt, allerdings liegt die Anzahl falsch klassifizierter Bereiche ab einer Anzahl von 7 Tags bei ~1/3 der Summe. Klar erkennbar ist vor allem im zweiten Evaluationsszenario, dass die Anzahl der richtig positiv klassifizierten Bereiche ab einer Grenze von 7 Tags erheblich ansteigt. Für richtig falsch klassifizierte Bereiche ist ein klarer Trend so nicht zu erkennen, allerdings ist die Rate der falsch negativ klassifizierten Bereiche ohnehin sehr gering. Auf Grundlage der Argumentation kann auch die Hypothese H_3 angenommen werden, da eine deutliche Abhängigkeit zwischen Datenbasis und Klassifizierung ausgemacht werden konnte.

Gemäß Abschnitt 5.1 erfolgt auf Grundlage relevanter Bereiche die Herleitung von geeigneten Empfehlungen. Die Hypothese H_2 stellt in diesem Zusammenhang nachfolgende Annahme:

Hypothese (H_2): Das System leitet innerhalb eines bestimmten Lernkontextes Empfehlungen her, für die eine positiv erkennbare Nützlichkeit seitens der Nutzer existiert.

Die Argumentation erfolgt ausschließlich aus den Ergebnissen des zweiten Evaluationsszenarios, da hier eine Erhebung im Kontext der Nützlichkeit möglich war. Die Ergebnisse zeigen, dass die überwiegende Anzahl an Bewertungen positiv ausfällt. Es ist interessant zu sehen, dass Empfehlungen richtig klassifizierter Bereiche i.d.R. eine gute bis sehr gute Bewertung (4-5 Sterne) erhalten haben. Wenn negative Bewertungen für Empfehlungen relevanter Bereiche ermittelt wurden, dann fallen diese allerdings sehr schlecht (1 Stern) aus. Da dies allerdings nur in ~5% der Fälle zutrifft, bleibt ein positiver Trend bestehen. Empfehlungen auf Grundlage richtig falsch klassifizierter Bereiche sind durchweg negativ bewertet, was allerdings einem zu erwartenden Vorgehen nachkommt, da diese Bereiche in einem realen Szenario ohnehin keine Empfehlung erhalten würden. Kritisch sind allerdings solche Empfehlungen zu sehen, die auf Grundlage einer falsch positiven Klassifizierung hergeleitet sind. Solche Empfehlungen haben eine durchweg mittelmäßige bis sehr schlechte Bewertung (3-1 Sterne) erhalten. Zwar ist die Anzahl der falsch relevanten Bereiche mit ~13% gering, allerdings können solche Empfehlungen einen positiven Gesamteindruck der Users mindern. Die erhobenen Bewertungen durch die Probanden zeigen dennoch, dass eine klar positive Nützlichkeit für Empfehlungen innerhalb des Systems vorliegt, weshalb Hypothese H_2 angenommen werden kann. Die Annahme erfolgt unter Berücksichtigung eines gewissen Risikos, welches von einer fehlerhaften Klassifizierung bzw. Herleitung von Empfehlungen ausgeht, was innerhalb eines Empfehlungsdiensteszenarios üblich ist. Dieser Sachverhalt wurde bereits in Abschnitt 3.2.5 der Arbeit diskutiert und findet für die Ergebnisse der vorliegenden Evaluation in Kapitel 8 noch einmal eine ausführliche Betrachtung. Mit der Argumentation zur Annahme der Hypothesen H_1 und H_2 ist gleichzeitig die übergeordnete Forschungsfrage F_1 mitbeantwortet worden, die die Güte der Empfehlungen in LAOs hinterfragt. Die Güte, ausgedrückt durch Nutzungswahrscheinlichkeit und Nützlichkeit, ist in quantitativer Form im vorherigen Abschnitt dargelegt und mit Beantwortung der Hypothesen

klar interpretiert. Eine mögliche Steigerung der Güte wird innerhalb der Diskussion thematisiert (vgl. Kapitel 8).

Analog zur Ermittlung relevanter Bereiche besteht die Annahme, dass die Qualität von Empfehlungen in Abhängigkeit der Datenbasis (Anzahl an Tags bzw. alternative Materialverbindungen) variiert. Die Hypothesen H_4 und H_5 beinhalten diese Annahme:

Hypothese (H_4): Es existiert ein Zusammenhang zwischen der Anzahl an Tags die für die Herleitung einer Empfehlung Verwendung finden und der ermittelten Nützlichkeit der jeweiligen Empfehlung.

Hypothese (H_5): Es existiert ein Zusammenhang zwischen der Anzahl alternativer Lernmaterialien oder Materialverknüpfungen die innerhalb einer Empfehlung präsentiert werden und der ermittelten Nützlichkeit der jeweiligen Empfehlung.

Die Ergebnisse zeigen zunächst für die Anzahl an Tags, dass eine positive Tendenz innerhalb der Bewertungen erst ab fünf Tags oder mehr zu erkennen ist. Diese Tatsache ist insofern nicht verwunderlich, da eine Vielzahl an falsch positiv klassifizierten Bereichen eine Anzahl von sieben oder weniger Tags aufweisen. Somit gilt auch für den Zusammenhang zwischen Datenmenge und Nützlichkeit, dass ein fundierter Informationsgehalt in Form einer genügenden Menge an Tags zu einer höheren Qualität führt. Somit kann Hypothese H_4 angenommen werden. Und auch für den Zusammenhang zwischen Bewertungen und der Anzahl an alternativen Lernmaterialien innerhalb einer Empfehlung zeigt sich ein ähnlicher Trend. Anders als zuvor entscheidet allerdings nicht die Menge an alternativen Materialien über eine positive Bewertung, sondern vielmehr das Vorhandensein eines alternativen Materials. Es ist ersichtlich, dass die Empfehlungen mit mindestens einer Empfehlung durchweg positiv bewertet sind. Für alle übrigen Empfehlungen ist zwar ebenfalls ein positiver Trend erkennbar, allerdings variieren die Bewertungen stärker. Man muss die Ergebnisse insofern relativieren, als dass alle Empfehlungen, die auf einer falschen Klassifizierung beruhen keine Materialverbindung aufweisen und somit stärkeren Einfluss auf die übrigen Empfehlungen ohne alternatives Material einnehmen. Betrachtet man das Ergebnis ohne falsch klassifizierte Bereiche, so erhalten Empfehlungen dennoch häufiger eine nur mittelmäßige Bewertung (3 Sterne), als wenn die

Empfehlung eine Materialverbindung aufweist. Hypothese H_5 kann deshalb angenommen werden. Mit der Annahme der Hypothesen $H_3 - H_5$ kann die übergeordnete Forschungsfrage F_2 Beantwortung finden. F_2 hinterfragt die Größe einer Datenbasis zur Herleitung positiv empfundener Empfehlungen aus Sicht der Probanden. Dass das System positiv empfundene Empfehlungen hinsichtlich der Nutzungswahrscheinlichkeit sowie der Nützlichkeit herleitet konnte zuvor bereits dargelegt werden. Die Argumentation zur Beantwortung der Hypothesen $H_3 - H_5$ hat zudem gezeigt, dass eine Abhängigkeit zwischen der Datenbasis und dem Resultat des Empfehlungsprozesses besteht. Die Forschungsfrage F_2 ist in zweierlei Hinsicht zu beantworten. Zum einen ist die Anzahl an Tags, auf denen eine Herleitung einer Empfehlung beruht, von Bedeutung. In beiden Evaluationsszenarien konnte gezeigt werden, dass zwischen 7-8 Tags eine Veränderung in der Klassifikation von relevanten Bereichen stattfindet. Die Anzahl falsch positiver Bereiche geht ab dieser Grenze deutlich zurück. Des Weiteren konnte gezeigt werden, dass Empfehlungen, die auf mindestens fünf Tags beruhen, verbesserte Bewertungen erhalten haben. Es ist an dieser Stelle somit die Nennung eines Schwellwertbereichs möglich, der zwischen 5-8 Tags liegt. Die Fixierung eines Wertes ist zu einem jetzigen Zeitpunkt nicht möglich, da die Verschiebung der Grenze die Nutzungswahrscheinlichkeit oder die Nützlichkeit beeinflussen würden. Diesen Sachverhalt gilt es im nachfolgenden Kapitel 8 noch einmal gesondert zu diskutieren. Einen weiteren Faktor gilt es im Zusammenhang mit der Größe der Datenbasis und zur Beantwortung der Forschungsfrage F_2 zu berücksichtigen. Im Verlauf des Abschnittes wurde dargelegt, dass das Vorhandensein einer inhaltlichen Materialverbindung – unabhängig ob alternatives Material aus dem Internet oder als Verbindung zwischen bestehenden Materialien – Einfluss auf die Bewertung von Empfehlungen nach sich zieht. Die Anzahl der Materialverbindungen scheint dabei keinen besonderen Einfluss zu haben. Aus den erhobenen Daten konnte zudem nicht ausgemacht werden, ob es einen qualitativen Unterschied zwischen den Materialverbindungen gibt (z.B. ob ein Link aus dem Internet inadäquat zum Lernen war). Dennoch sind alternative Lernmaterialien in Empfehlungen aus Sicht der Probanden wünschenswert, so dass Empfehlungen diese möglichst beinhalten sollten. Es ist allerdings problematisch, dass die Herleitung von Materialverbindungen in Abhängigkeit der Systemuser geschieht. Zwar bleibt die Annahme bestehen, dass hilfreiche Lernmaterialien eines Users ggf. sinnvoll für andere Lernende sind, zumindest sagen die erhobenen Ergebnisse nichts Gegenteiliges aus, allerdings wäre eine Art „Backupstrategie“ wünschenswert, die ggf. ohne

explizite Rückmeldung der Usergruppe auskommt. Auch diesen Sachverhalt gilt es in Kapitel 8 zu diskutieren.

Die Hypothesen H_6 und H_7 treffen die Annahme, dass eine sinnvolle inhaltliche, sowie visuelle Darstellung für Hilfestellungen und Verbindungen zwischen Lernmaterialien in Empfehlungen existiert:

Hypothese (H_6): Für inhaltliche Hilfestellungen innerhalb hergeleiteter Empfehlungen existiert eine positiv erkennbare unterstützende inhaltliche, sowie visuelle Darstellung aus Sicht der Nutzer.

Hypothese (H_7): Für inhaltliche Verbindungen zwischen Lernmaterialien innerhalb hergeleiteter Empfehlungen existiert eine positiv erkennbare unterstützende inhaltliche, sowie visuelle Darstellung aus Sicht der Nutzer.

Bedingt durch die erhobene Nützlichkeit von Empfehlungen kann eine generelle Sinnhaftigkeit von Empfehlungen angenommen werden. Eine negative Darstellung von Hilfestellungen und Materialverbindungen hätte sonst in einer geringen Bewertung der Nützlichkeit durch die Probanden gemündet. Eine explizite Aussage ergibt sich des Weiteren durch die Aussagen 16, 19 und 20 der Akzeptanzbefragung. In Aussage 16 beurteilen die Probanden die allgemeine Darstellung der Empfehlungen, die mit einer durchschnittlichen Bewertung von $\sim 3,7$ positiv beantwortet wurde. Aussage 19 erhebt in wieweit sich eine hilfreiche Darstellung gemäß der Kategorien Wichtigkeit, Unterstützung, Schwierigkeit und Verständnis ergibt. Mit einem durchschnittlichen Wert von $\sim 3,90$ ist diese Frage ebenfalls positiv bewertet. In Aussage 20 erhebt die Befragung, in wieweit die Probanden die Einschätzung anderer Lernende zum Lerninhalt als sinnvoll erachten. Diese Einschätzung ist innerhalb von Empfehlungen sowohl durch die grafische als auch durch die textuelle Darstellung repräsentiert. Die positive Einschätzung durch die Probanden ist durch einen mittleren Wert von $\sim 3,81$ ebenfalls gegeben. Auf dieser Grundlage kann die Hypothese H_6 angenommen werden. Zur Prüfung der Hypothese H_7 bedarf es des Weiteren der Hinzunahme von Aussage 22 der Akzeptanzbefragung, die das Maß an Hilfe für Materialverbindungen hinterfragt. Die Auswertung ergab einen durchschnittlichen Wert von 3,71 für Aussage 22 und somit eine positive Einschätzung zur

Hilfestellung der alternativen Verbindungen. Zwar wurde keine explizite Aussage über die Darstellung der Materialverbindungen getätigt, allerdings kann über die Gesamtbewertungen der Empfehlungen und über die Akzeptanzerhebung von einer genügenden inhaltlichen und visuellen Darstellung von Materialverbindungen ausgegangen werden. Somit ergibt sich auch für H_7 eine Annahme der Hypothese. Die übergeordneten Forschungsfragen F_3 und F_4 der Hypothesen H_6 und H_7 liegt die Fragestellung zugrunde, was eine unterstützende Hilfestellung bzw. Materialverbindung innerhalb einer Empfehlung umfasst. Beide Fragestellungen mussten bereits innerhalb der Entwicklung des Systems thematisiert werden. Wie in Abschnitt 4.4.4 beschrieben, basiert die inhaltliche Darstellung von Empfehlungen grundlegend auf dem Learning Object Review Instrument (LORI). Die Annahme besteht darin, dass das System inhaltliche Lerneinheiten extrahiert, die einer Art Learning Object gleich kommen und dementsprechend über das LORI bewertbar sind. Von großer Bedeutung schien bei der Entwicklung, dass eine möglichst einfache Interpretation, aber gleichzeitig eine genügend detaillierte Interpretation der Empfehlungen möglich ist. Die Kombination aus einer einfachen grafischen Übersicht für die vier herausgestellten Kategorien, sowie die einfache, aber aussagekräftige Erläuterung in Textform, erschien als inhaltliche Hilfestellung angemessen. Die angeführte Argumentation kommt der Beantwortung von Forschungsfrage F_3 gleich. Die Darstellung von Materialverbindungen ist im Detailgrad der Konzeption weniger ausgeprägt. Neben einem direkten Einstiegspunkt via Link in das alternative Material erschien vor allem eine Aussage zu der Alternative von Bedeutung. In der Evaluation wurde so häufig die Nennung eines Themenfeldes als Beschreibung verwendet. Die Argumentation für die Annahme der Hypothese H_7 hat dargelegt, dass von einer negativen inhaltlichen, sowie visuellen Darstellung von Materialverbindungen nicht auszugehen ist. Des Weiteren ist vor allem das Vorhandensein von Materialverbindungen, was aus der Bewertung der Nützlichkeit durch Probanden hervorgeht, von Bedeutung. Die vorliegende Darstellung der Materialverbindungen innerhalb der Empfehlungen reicht somit aus. Dies kommt der Beantwortung von F_4 nach.

Die Hypothesen $H_8 - H_{10}$ treffen übergeordnete Annahmen für das entwickelte System:

Hypothese (H_8): Das System leitet innerhalb eines bestimmten Lernkontextes Empfehlungen her, deren Verknüpfungen zwischen bestehenden und neuen, ungenutzten Lernmaterialien eine positiv erkennbare Unterstützung für die Nutzer darstellen.

Hypothese (H_9): Das System leitet innerhalb eines bestimmten Lernkontextes Empfehlungen her, deren Hilfestellungen für bestehende Lernmaterialien eine positiv erkennbare Unterstützung für die Nutzer darstellen.

Hypothese (H_{10}): Es existiert eine positiv erkennbare Akzeptanz zur Nutzung des Systems seitens der Nutzer.

Die Hypothesen H_8 und H_9 sind grundlegend mit den bereits angenommenen Hypothesen H_1 und H_3 vergleichbar. Der zentrale Unterschied besteht allerdings darin, dass die Hypothesen H_1 und H_3 eine Sichtweise auf die implementierte Algorithmik einnehmen und sowohl die Nutzungswahrscheinlichkeit, sowie die Nützlichkeit innerhalb des Empfehlungsprozesses bewerten. Hypothesen H_8 und H_9 sind vielmehr als übergeordnete Annahmen zu sehen, die ein Gesamtbild von der Güte im Empfehlungsprozess über die Darstellung von Empfehlungen bis hin zur Akzeptanz des Systems hinterfragen. Was die Hypothesen in besonderer Form hinterfragen, ist sowohl die Nutzung des Systems als auch die Herleitung von Empfehlungen innerhalb eines bestimmten Lernkontextes. Die Evaluationsszenarien sind bewusst so konstruiert worden, dass sich die Probanden in einer Lernsituation mit konkreten Lernzielen wiederfinden. Vor allem das zweite Evaluationsszenario der online Simulation folgt einem realen Lernsetting. Im Verlauf der Arbeit wurde die kontextbezogene Herleitung von Empfehlungen immer als bedeutsam herausgestellt und so wurden auch die Ergebnisse der Evaluation in einem festgeschriebenen Lernkontext hergeleitet, so dass auch die bislang geführte Argumentation für diesen Lernkontext gilt. Die positive Unterstützung von Empfehlungen in Form von Hilfestellungen und Materialverbindungen bzw. Verknüpfungen zwischen Lernmaterialien wurde in Form einer festgestellten Güte und eine positive Darstellung von Empfehlungen hergeleitet. Die Annahme der Hypothesen H_8 und H_9 ist folglich möglich. Die Hypothese H_{10} folgt ebenfalls einer übergeordneten Annahme, die für die schlussendliche Beantwortung der übergeordneten Forschungsfrage F_5 von zentraler Bedeutung ist. Die Akzeptanz für einen

Empfehlungsdienst bildet die Grundlage zur Nutzung eines etwaigen Systems (vgl. Abschnitt 3.2.5). Zwar könnte die Forschungsfrage F_5 auf Grundlage der bis hierher angenommenen Hypothesen beantwortet werden, allerdings wäre dies ohne die Berücksichtigung der Akzeptanz unvollständig. Das System bedarf der Nutzung durch eine Usergruppe, da ohne Nutzung eine Funktionalität des Systems nicht gewährleistet ist. So bedarf es der Sicherstellung der Akzeptanz, um von einer grundsätzlichen Nutzung des Systems ausgehen zu können. Die Akzeptanz ist grundlegend abhängig von der Qualität der Empfehlungen, sowie des empfundenen Aufwands die Empfehlungen zu erhalten. Die Ergebnisse sind entsprechend der Kategorien in Abschnitt 6.2 für die online Simulation präsentiert. Die Aussagen zur Qualität (Aussage 1-7) wurden durchweg positiv bewertet. Lediglich Aussage 3 und 4 enthalten eine eher mittelmäßige Bewertung. Beide Aussagen erörtern die durch das System verbesserte Produktivität der Lernenden. Bedingt durch die sonst durchweg positive Beurteilung der Aussagen und die gemessene Güte der Empfehlungen, die in gleicher Form ein Qualitätsmaß darstellt, kann übergeordnet von einer klar positiv wahrgenommenen Qualität des Systems ausgegangen werden. Dies gilt in gleicher Form für den Aufwand, den die Lernenden zum Erhalt von Empfehlungen aufwenden müssen. Die Aussagen zur Usability des Systems (Aussagen 8-11) erhalten durchweg positive Beurteilungen. Auch die Darstellung der Empfehlungen - wie bereits zuvor begründet - und ebenso der Zugang zu Empfehlungen erscheint für die Probanden ausreichend. Lediglich die Auswahl an Werkzeugen scheint hinsichtlich des angebotenen Umfangs an Werkzeugen verbesserungswürdig, obgleich auch in diesem Punkt eine positive Tendenz erkennbar ist. Die positive Resonanz wird auch durch die übergeordneten Aussagen 19-22 zur Akzeptanz bestätigt. Die Ergebnisse stellen somit eine positiv erkennbare Akzeptanz für das System dar. Die Hypothese H_{10} kann deshalb ebenfalls angenommen werden.

Die Forschungsfrage F_5 stellt die übergeordnete Forschungsfrage für die vorliegende Arbeit dar. Es wurde ein Empfehlungsdienst vorgestellt, der auf einem Tag-basierten Ansatz beruht und Empfehlungen in Form von inhaltlichen Hilfestellungen und Materialverknüpfungen ableitet. Der Empfehlungsprozess gliedert sich dabei in die Extraktion von inhaltlich relevanten Bereichen und die Herleitung von Empfehlungen. Der entwickelte Empfehlungsdienst wurde unter Berücksichtigung verschiedener Aspekte geprüft, wie die Hypothesen $H_1 - H_{10}$ zeigen. Die Güte von Empfehlungen nahm dabei eine zentral zu prüfende Aussage ein. In diesem

Zusammenhang galt es Zusammenhänge innerhalb des Empfehlungsprozesses zu prüfen, die für die Güte der Empfehlungen von Bedeutung sind. Neben einer positiv empfundenen Darstellung der Empfehlungen wurde ebenso die Akzeptanz für das entwickelte Verfahren hinterfragt und positiv durch die Probanden beantwortet. Die Hypothesen stellten somit alle relevanten und zu prüfenden Punkte zur Diskussion, um eine Beantwortung von Forschungsfrage F_5^1 zu ermöglichen. Bedingt durch die Annahme aller gestellten Hypothesen kann die Forschungsfrage ebenfalls positiv beantwortet werden. Die Interpretation der Ergebnisse hat aber auch gezeigt, dass das System und der implementierte algorithmische Ansatz noch geringfügigen Optimierungsbedarf aufweist, den es trotz der genügenden Ergebnisse zu diskutieren gilt. In diesem Zusammenhang ergibt sich in Kapitel 8 zunächst eine Erörterung der aufgedeckten Schwächen und gibt Lösungsideen, die das Verfahren ergänzen können. Des Weiteren wird die in diesem Kapitel vorliegende Interpretation mit Ergebnissen aus der Literatur gegenübergestellt, um eine Aussage über den Mehrwert des Systems im Vergleich zu weiteren Lösungsansätzen zu treffen.

8 Diskussion

Mit der Interpretation der Ergebnisse in Kapitel 7 konnten die Forschungsfragen und Hypothesen der vorliegenden Arbeit beantwortet werden. Es hat sich gezeigt, dass das implementierte Verfahren positive Ergebnisse erzielt, um eine Unterstützung in Form von inhaltlichen Hilfestellungen auf der einen Seite sowie Verknüpfungen zwischen Lernmaterialien auf der anderen Seite für eine Gruppe Lernender bereitzustellen. Die Präsentation der Ergebnisse hat allerdings in gleicher Weise gezeigt, dass einige Aspekte des Verfahrens diskussionswürdig sind, was im Verlauf des nachfolgenden Kapitels geschehen soll. Des Weiteren bedarf es eines Vergleichs zu bestehenden Ansätzen aus der Literatur, die bereits in Abschnitt 3.5 präsentiert wurden. Es gilt in diesem Zusammenhang zu diskutieren, ob das neu entwickelte Verfahren einen Mehrwert gegenüber bestehenden Verfahren darstellt oder wo bestehende Verfahren eine bessere Funktionsweise anbieten.

Der direkte Vergleich zu bestehenden Verfahren ist dabei nicht ganz trivial. Zwar legt die Literatur einheitliche Möglichkeiten der Evaluation offen, allerdings variieren die Ansätze der Studien, bedingt durch die unterschiedlichen Varianten an Empfehlungsdiensten, deutlich. Grund hierfür sind in erster Linie die unterschiedlichen Rahmenbedingungen und Zielvorgaben eines Systems, die unweigerlich in ebenso unterschiedlichen Lösungsansätzen münden. In Abschnitt 3.5 hat sich bereits gezeigt, dass eine einheitlich Definition so z.B. lediglich für Learning Objects existiert. Sowohl Lernressourcen als auch Kontextinformationen werden hingegen durchweg unterschiedlich interpretiert, so dass eine Vielzahl unterschiedlich zu empfehlender Items existiert. Wie im Verlauf der Arbeit bereits dargelegt wurde, schließt sich vor allem ein direkter algorithmischer Vergleich aus. Häufig fehlt auch die Durchführung von Benutzerstudien oder online Simulationen, so dass zwar belegt durch offline Experimente eine sinnvolle Herleitung

von Empfehlungen angenommen, im realen Einsatz dennoch nicht belegt werden kann (vgl. [14, 26, 30, 49, 81, 99, 111, 113]). Die Literatur zeigt allerdings die Bedeutung von aktiver Teilnahme von Probanden auf, um vor allem hinsichtlich der Unterstützung des Systems eine genügende Aussage treffen zu können [57, 101]. Für weitere Ansätze liegt zwar eine schlüssige Argumentation, aber keine belastbaren Ergebnisse für das jeweilige Verfahren vor (vgl. [86, 97, 106, 117, 144, 162]). Für die Implementationen, für die Ergebnisse aus Benutzerstudien oder online Simulationen vorliegen, kann der Versuch eines Vergleichs erfolgen. Die grundlegende Intention der präsentierten Ansätze ist die Unterstützung von Lernenden, wo sich ein gemeinsamer Ankerpunkt zur Implementation der vorliegenden Arbeit ergibt. Wie zu Beginn dargelegt, variieren die Evaluationsmethoden allerdings stark und so bedarf es vielmehr eines interpretativen Vergleichs, als der quantitativen Gegenüberstellung von Ergebnissen. So verfolgt der Ansatz in [126] die Empfehlung von Webressourcen auf Grundlage von Bewertungen. Die in einer Studie gemessene Nützlichkeit – bewertet von 1-5 (gar nicht nützlich=1, sehr nützlich=5) – zeigt mit Wertigkeiten von 2,9, 3,3 und 4,28 einen positiven Trend innerhalb der Ergebnisse. Vor allem für die letzte Wertigkeit kann dem Verfahren eine positive Unterstützung zugesprochen werden. [3] zeigt durch empirische Ergebnisse ebenfalls eine positive Form der Unterstützung für Lernende für die Empfehlung von Videos und Videosequenzen. So trifft die Klassifizierung von relevanten Items in 74% der Fälle zu und zeigt damit eine genügende Nutzungswahrscheinlichkeit. Die Arbeit in [135] zeigt mäßig positive Unterstützung bei der Empfehlung von Materialien in textueller Form. Auf einer Bewertungsskala von 1-5 (stimme gar nicht zu=1, stimme voll zu=5) empfinden die Probanden durch das System eine mäßige Steigerung der Effektivität (3,14), mäßige Steigerung der Effizienz (3,28) und wünschen sich das System bedingt in weiteren Lernszenarien (3,28), obgleich das System einfach zu bedienen ist (4,0). Die Empfehlung von Text Items wurde in [39] beschrieben. Die Evaluationsergebnisse zeigen, dass Benutzerbewertungen für Items von 1-5 (gar nicht relevant=1, sehr relevant=5) unter Betrachtung des Mean Absolute Errors (MAE) nur gering von den Systemschätzungen abweichen (MAE=0,92). Es kann somit von einer guten Nützlichkeit der Empfehlungen des Systems ausgegangen werden. Das gleiche Evaluationsschema ist in [114] beschrieben, wo die Herleitung von bewusst klein konstruierten Lernmaterialien Bewertung findet. Die Ergebnisse zeigen eine grundlegend ähnliche Abweichung (MAE=0,96), weshalb in diesem Zusammenhang ebenfalls eine angemessene Nützlichkeit angenommen werden kann. Für die Gruppe der

Learning Objects zeigt die Arbeit in [167] für die Bewertung von Empfohlenen LOs auf einer Skala von 1-5 (sehr schlecht=1, sehr gut=5) in den Kategorien Präferenz und Nützlichkeit eine durchweg positive Nützlichkeit ($\sim 4,0$), allerdings eine weniger gute Präferenz für die LOs ($\sim 2,3$). Da vor allem die Nützlichkeit eine hohe Wertigkeit erhalten hat, ist von einer positiven Unterstützung durch das System auszugehen. In [24] erfolgt eine Evaluation auf der Empfehlung einer Liste mit zehn LOs, die durch die User eine eigene Sortierung erhält. In Abhängigkeit der Distanz zwischen geschätzter und neuer Position in der Liste ergibt sich eine mittlere Distanz in zwei Experimenten von 1,76 (Experiment A) und 13,6 (Experiment B). Vor allem das erste Experiment zeigt äußerst gute Vorhersageergebnisse, so dass von einer klar positiven Nutzung der LOs und einer dementsprechend guten Unterstützung durch das System ausgegangen werden muss. In [149] wurde die Nützlichkeit von LOs innerhalb einer Evaluation bewertet. Die LOs haben eine durchschnittliche Nützlichkeit von 3,2 erhalten. Dies kann als mäßig bis genügende Unterstützung für das System interpretiert werden.

Einige Aspekte fallen innerhalb der präsentierten Evaluationsergebnisse bestehender Verfahren auf. Die Erhebung von Ergebnissen erfolgt häufig außerhalb eines realen Lernsettings, weshalb die Studien z.T. lediglich Teilaspekte eines Lernsettings berücksichtigen. Für die Lernenden bedarf es somit häufig keiner aktiven Teilnahme während des Lernens, sondern vielmehr der kontextbezogenen Einschätzung von Empfehlungen (vgl. z.B. [24]). Des Weiteren sind die Szenarien z.T. stark angepasst, so dass eine Beurteilung vereinfacht wird. [114] beschreibt so z.B. die sinnvolle Herleitung von Empfehlungen für Lernmaterialien, allerdings sind für das Evaluationssetting bewusst sehr kleine Materialien ausgesucht worden, um eine Bewertung durchführen zu können. Lediglich die Arbeiten in [3] und [135] erheben Informationen aus einem realen Lernszenario. Der Mehrwert aller Ergebnisse steht zwar unweigerlich außer Frage, allerdings ergibt sich aus der Literatur, dass vor allem für die Bewertung der Unterstützung eines Systems, die Umsetzung eines möglichst realen Evaluationssettings bedarf, da zu viele Einflussfaktoren auf die Beurteilung der tatsächlichen Unterstützung während des Lernens einwirken (vgl. [101] S. 66). Für die Berücksichtigung ausgewählter Teilaspekte gibt es schlussendlich sachdienliche Gründe. Die Systeme verfolgen häufig eine ausgewählte Funktionalität, die es zu prüfen gilt (z.B. Auffinden gleicher LOs oder gleicher Lernressourcen usw.). Wie im Verlauf der Arbeit argumentiert, verfolgt das Verfahren innerhalb der

vorliegenden Arbeiten einen bewusst ganzheitlichen Ansatz, um den gestellten Zielvorgaben entsprechen zu können. Aus diesem Grund war es auch von Bedeutung, eine Evaluation unter Berücksichtigung eines ganzheitlichen Lernsettings durchzuführen. Betrachtet man allerdings den Empfehlungsprozess der vorliegenden Arbeit, so ergeben sich ebenfalls Teilaspekte, die untersucht wurden und mit den Ergebnissen bestehender Verfahren grundlegend vergleichbar sind. So stellt sich die Extraktion inhaltlicher Bereiche als Klassifikationsschritt dar, wohingegen die Herleitung von Empfehlungen der Bewertung von Nützlichkeit bedarf. Auch wenn eine direkte Gegenüberstellung von Ergebnissen nicht sinnvoll erscheint, so ist dennoch ein grober Vergleich hinsichtlich der Ausprägung der Ergebnisse möglich. Es zeigt sich nämlich, dass eine 100% Unterstützung durch Empfehlungsdienste nicht existiert. Dies ist in Empfehlungsdiensten für Webshops ebenso wenig der Fall wie für Empfehlungsdienste im Kontext des E-Learning. Die Faktoren, die auf eine Beurteilung durch eine Person einwirken sind so vielfältig, dass hiervon auch nicht ausgegangen werden sollte. Bereits die angeführte Argumentation zur Interpretation einer Bewertungsskala aus Abschnitt 3.2 stützt diese Aussage. So interpretiert die eine Person innerhalb einer Evaluation seine/ihre Rolle sehr kritisch und vergibt nur sehr selten vollständig positive Bewertungen, wohingegen eine andere Person häufig positiv bewertet. Die Evaluation für Empfehlungsdienste dient deshalb in der Regel nicht der Darlegung durchweg positiver Ergebnisse, sondern vielmehr der Darlegung eines positiven Trends, über den von einer Unterstützung für eine Vielzahl an Nutzern ausgegangen werden kann. Für Empfehlungsdienste ist deshalb eine Optimierung sinnvoll, die unter Berücksichtigung bestimmter Risiken einen gewissen Grad an Fehlern zulässt. Dieser Sachverhalt wird im Verlauf des vorliegenden Abschnitts noch einmal genauer diskutiert. Auf Grundlage dieser Interpretation lässt sich sagen, dass die Ergebnisse der vorliegenden Arbeit im Vergleich zu bestehenden Verfahren gleichwohl positive Unterstützung belegen. Die Bewertung der Nützlichkeit zeigt sowohl im Vergleich zur Empfehlung von Lernressourcen als auch zur Empfehlung von LOs, dass LAOs mit der Extraktion von inhaltlichen Bereichen und der entsprechenden Verknüpfung von Kontextinformationen und alternativen Lernmaterialien positive Ergebnisse erzielt, ohne dass es einer gezielten Aufbereitung bedarf oder Lernmaterialien nicht berücksichtigt werden können. Die Zielvorgaben, die sich aus Nachteilen bestehender Verfahren ergaben finden somit Berücksichtigung, ohne erkennbare Abstriche hinsichtlich der Unterstützung ausgemacht zu haben. Die Ergebnisse werden dadurch gestärkt, dass das Evaluationssetting ein reales

Lernszenario abbildet und die Ergebnisse deshalb keinen weiteren hypothetischen Annahmen unterliegen.

Es gibt dennoch einige Aspekte, die es im Kontext der vorliegenden Arbeit zu diskutieren gilt. Das vorgestellte Verfahren hat gezeigt, dass die Klassifikation oder auch Extraktion von inhaltlich relevanten Bereichen positiv zu bewerten ist. Die übergeordnete Rate an richtig klassifizierten Bereichen ist mit ~77% (Benutzerstudie) bzw. ~86% (online Simulation) durchaus hoch einzuschätzen. Vor allem die positive Einschätzung von Lernenden (~86%) ist von bedeutender Aussage. Da die Herleitung von Empfehlungen zentral mit der Extraktion relevanter Bereiche zusammenhängt, bedarf es der besonderen Berücksichtigung dieses Prozessschrittes. Betrachtet man vor allem die falsch positiv klassifizierten Bereiche, dann ist das Ergebnis als noch nicht optimal zu bewerten. Zwar ist das Ergebnis durchaus zufrieden stellend, allerdings könnte der Einfluss von fälschlicherweise als relevant erachteten Bereichen einen negativen Einfluss auf die Usergruppe mit sich bringen. Die Optimierung des Klassifikationsprozesses birgt dabei allerdings ein gewisses Risiko. Die Verbesserung des Precision Wertes und somit die implizite Verbesserung der richtig als relevant klassifizierten Bereiche, beeinflusst unweigerlich die Wertigkeit Recall und somit die Klassifizierung von nicht relevanten Bereichen. Es gilt an dieser Stelle abzuwägen, ob es aus Sicht der User sinnvoller ist möglichst viele Bereiche zu präsentieren, unter denen sich nicht relevante Bereiche befinden, oder eine zurückhaltende Variante zu wählen und dabei relevante Bereiche ggf. nicht anzubieten. Es gilt in diesem Zusammenhang das Risiko zu beurteilen, welchen Einfluss z.B. eine falsch positive Klassifikation auf den jeweiligen Nutzer hat (vgl. [129] S. 367 ff.). So gilt es z.B. bei der Steuerung von Maschinen die Betriebsamkeit zu maximieren, so dass die Wertigkeiten unter Berücksichtigung des Ausfalls optimiert werden. Das Risiko für Endanwender z.B. im Bereich des Entertainment (Empfehlungen für Filme oder Artikel) ist bei der Empfehlung eines nicht relevanten oder nützlichen Artikels in der Regel marginal. In diesem Zusammenhang ist es eher von Bedeutung eine Vielzahl von falschen Klassifikationen zu vermeiden, um die Geduld eines Endanwenders nicht unnötig zu strapazieren und eine nicht Akzeptanz gegenüber des Systems zu erwirken. Das Risiko von Empfehlungsdiensten im Bereich des Lernens bewegt sich nach eigener Einschätzung zwischen den beispielhaften Gruppen an Szenarien. Die Auswirkungen von falschen Empfehlungen sind für einen Lernenden zwar nicht so schwerwiegend, als dass er

oder sie dadurch bedeutsam beeinträchtigt werden würde. Allerdings können falsche Empfehlungen während des Lernens zu Verwirrung führen, sofern z.B. ein Bereich fälschlicherweise als relevant klassifiziert ist, der Lerner dies allerdings nicht einzuschätzen vermag. Eine genügende Sicherheit bei der Klassifizierung und somit die Optimierung des Precision Wertes ist deshalb wünschenswert. Die Ergebnisse der Evaluation haben in diesem Zusammenhang bereits eine Möglichkeit der Optimierung aufgezeigt. Die Klassifikation von falsch relevanten Bereichen nimmt ab einer Schwelle von 7-8 Tags deutlich ab, so dass diese Schwelle als Ausgangspunkt für eine Klassifikation nutzbar wäre. Die verbesserten Bewertungen hinsichtlich der Nützlichkeit ab 5 Tags stützt dieses Vorgehen. Die Prüfung dieser Annahmen, sowie die Unterschiede zwischen verschiedenen Schwellwerten im Bereich von 5 bis 8 Tags, bedarf einer weiteren Untersuchung die über die Möglichkeiten der vorliegenden Arbeit hinausgehen.

Betrachtet man das Verfahren noch einen Schritt früher - auf Ebene der Clusterung von benachbarten Tags - so kann in diesem Zusammenhang keine Möglichkeit der grundlegenden Optimierung erkannt werden. Solche Tags, die einen Zusammenschluss zu einer falsch positiven Resource ergaben, galt es aus Entwicklersicht dennoch in gleicher Form zu extrahieren. Bei nachträglicher Einsicht lässt sich feststellen, dass bei ausschließlicher Betrachtung der Tagkoordinaten ein Zusammenschluss sinnvoll war. Abbildung 59 illustriert dies anhand von zwei Beispielen, die auf real erhobenen Daten der Evaluation beruhen. Beispiel 1 (Abbildung 59 links) beruht auf insgesamt neun Tags, die zu einer Resource zusammengeschlossen wurden. Die Nähe der Tags lässt ohne den Inhalt zu berücksichtigen auf einen idealen Zusammenschluss schließen, da die Bündelung der Tags im mittleren Bereich der Seite stark ausgeprägt ist. Die Distanz zwischen den Tags ist des Weiteren recht eng zusammenliegend und sonst sind keine weiteren Tags auf der Seite vertreten. Ähnliches gilt für das zweite Beispiel (Abbildung 59 rechts), wo die Bündelung der Tags am Ende der Seite auftritt. Die Distanz zwischen den Tags ist ebenfalls recht gering, so dass ohne inhaltliche Berücksichtigung eine Extraktion erfolgen sollte.

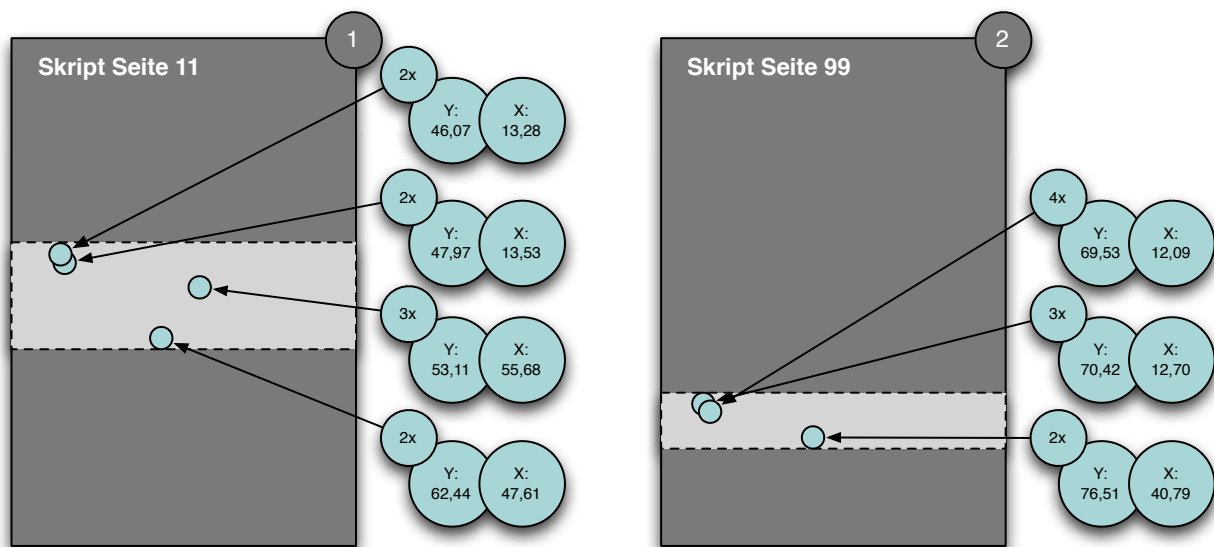


Abbildung 59 Beispiel False Positive Klassifikation

Nicht nur hinsichtlich der Clusterung, sondern für die zuvor vorgeschlagene Tagschwelle, die bereits einige falsch positiv erkannte Bereiche ausschließen würde, sind die Beispiele in Abbildung 59 problematisch. Denn für beide Beispiele greift die vorgeschlagene Tagschwelle nicht. Auf Grundlage der erhobenen Datenbasis wäre dies im Regelfall auch nicht wünschenswert. Eine erste Überlegung wäre die Anzahl der unterschiedlichen User, die Tags innerhalb der Resource gesetzt haben, zu prüfen. In beiden Beispielen haben jeweils drei unterschiedliche User den Bereich mit Tags versehen. Es muss davon ausgegangen werden, dass eine Meinung von drei Personen durchaus Relevanz besitzen kann, so dass eine Prüfung auf der Anzahl unterschiedlicher User wenig sinnvoll erscheint. Auch die Größe eines inhaltlichen Bereichs scheidet als zu prüfender Faktor aus, denn auch sehr kleine Bereiche sind ggf. von Bedeutung. Die User explizit mit in den Entscheidungsprozess mit einzubeziehen ist des Weiteren denkbar. Sofern ein Bereich als nicht relevant oder durch eine Sterne Bewertung schlecht bewertet wurde, könnte es eine Art Quarantäne Option geben, die den Bereich als fraglich einstuft und für die User genügend aussagekräftig so darstellt. In diesem Zusammenhang bedarf es allerdings weiterer Überlegungen, da die Funktionalität ggf. nicht auf Anhieb – also durch eine einmalige schlechte Bewertung – eintreten dürfte. Eine plausible Möglichkeit zur Verbesserung wurde bereits innerhalb der formalen Betrachtung des Algorithmus in Abschnitt 4.4 benannt. Die Berücksichtigung der Verweildauer auf einer Resource oder vielmehr innerhalb eines Bereichs, wo die Resource liegt (z.B. einer Textseite),

als weiteren Relevanzfaktor ist plausibel. Die Verwendung kann innerhalb des Verfahrens über die Betrachtung des impliziten Tags der Verweildauer erfolgen. Der Faktor Verweildauer blieb bislang unberücksichtigt, da nicht schlüssig argumentiert werden konnte, welcher Einfluss für den Einsatz sinnvoll erscheint. Anders als bei der Bestimmung einer Relevanz über die Anhäufung von Tags ist die Zeit kein offensichtlich interpretierbarer Faktor. Eine hohe Verweilzeit scheint eine schlüssige Information dafür zu sein, dass für den Bereich und ebenso für den Inhalt bzw. die jeweilige Resource eine signifikante Relevanz vorliegt. Die entscheidende Frage besteht nun darin, wann von einer Relevanz hinsichtlich der Verweildauer *nicht* ausgegangen werden kann. Eine geringe Verweildauer muss nicht zwangsläufig bedeuten, dass auch eine geringe Relevanz für den Bereich vorliegt. So kann ein Bereich ggf. lediglich aus einem einfach zu interpretierenden Bild oder Text bestehen, so dass es keiner Anstrengung zur Bearbeitung bedarf. Ungeachtet der Gründe, die für eine kurze Verweildauer sprechen, bedarf es der Beurteilung auf Grundlage von stichhaltigen Informationen. Die erhobenen Daten der Evaluation können hierfür herangezogen werden. Abbildung 60 betrachtet hierzu die Klassifikation von richtig (blau) und falsch (rot) relevant klassifizierten Bereichen in Abhängigkeit der Verweildauer für das Skript aus der online Simulation. Die Abbildung zeigt, dass es durchaus eine erkennbare Tendenz dahingehend gibt, dass innerhalb hoch frequentierter Bereiche hinsichtlich der Zeit eine häufigere positive Klassifikation stattfindet. Die Frage ist nun wie bereits gestellt, wo es eine plausible Grenze gemäß eines Schwellwertes zu ziehen gilt. Dies ist wie bereits vermutet nicht trivial zu beantworten. Bedingt durch eine Vielzahl nur geringfügig besuchter Bereiche, ist z.B. die durchschnittliche Verweildauer die analog zum Tagaufkommen Verwendung finden könnte, äußerst hoch. Dies ist für lange Textdokumente auch plausibel, da innerhalb eines Kontextes nicht das ganze Dokument von Bedeutung ist. Zieht man somit für das Beispiel in Abbildung 60 die Grenze bei Wert ~ 13 , so zeigt sich, dass eine Vielzahl richtig als relevant klassifizierter Bereiche keine Berücksichtigung finden würde. Selbst bei einer drastischen Reduzierung auf die Hälfte des Wertes ($\sim 6,5$) findet eine bedeutsame Anzahl an relevanten Bereichen keine Berücksichtigung.

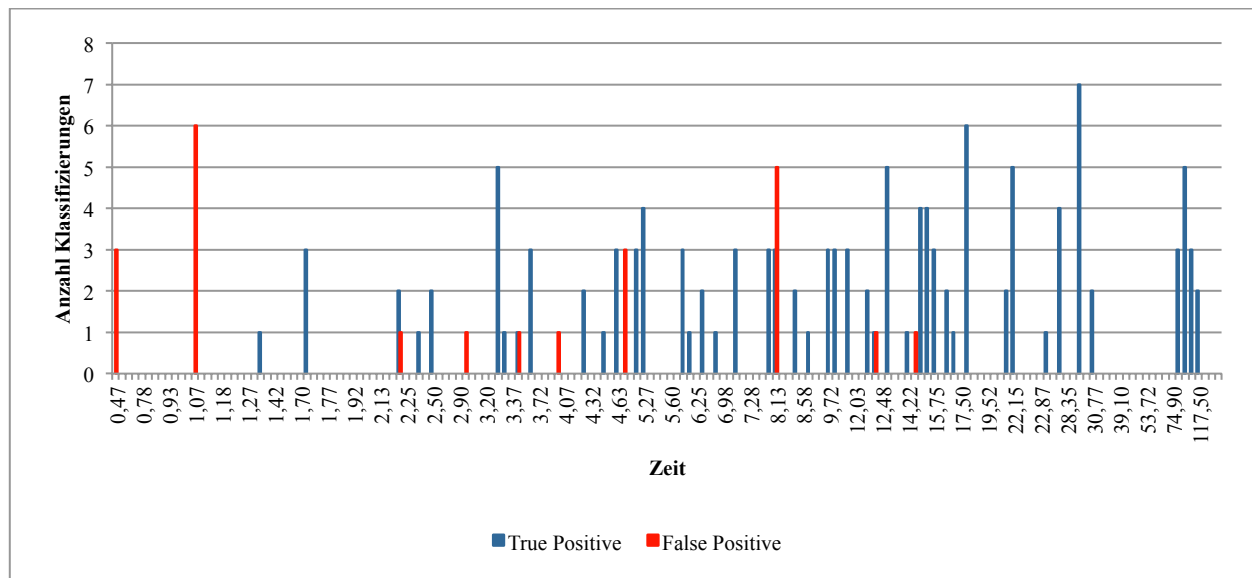


Abbildung 60 Klassifikation auf Grundlage der Verweildauer (Skript, online Simulation)

Die Verweildauer als ein alleiniges Entscheidungskriterium zu nutzen scheint somit nicht denkbar. Da die Klassifikation auf der Grundlage des Tagaufkommens gute Resultate erzielt, wäre die Verweildauer ohnehin als kombinierter Relevanzfaktor sinnvoller einzuschätzen. Es ist denkbar die Verweildauer als Zusatz heranzuziehen, um die Entscheidungskraft des Tagaufkommens zu verbessern. So wurde innerhalb des vorliegenden Abschnittes eine Schwellwertgrenze zwischen 5-8 Tags diskutiert. Für Ressourcen, die ein Aufkommen zwischen oder ggf. geringfügig unter-/oberhalb dieser Grenze aufweisen, könnte eine hohe bzw. geringe Verweildauer eine Entscheidungshilfe für die positive bzw. negative Klassifizierung sein. Abbildung 61 präsentiert für die weitere Argumentation die Klassifizierung von richtig (rot) und falsch (blau) relevant klassifizierten Bereichen unter Berücksichtigung von Verweildauer und Tagaufkommen. Die Abbildung zeigt für die richtig klassifizierten Ressourcen, dass die Verweildauer und das Tagaufkommen i.d.R. proportional stark ansteigen. Dies bedeutet grundlegend, dass richtig klassifizierte Bereiche sowohl eine hohe Verweildauer als auch ein hohes Tagaufkommen aufweisen. Eine Optimierungsmöglichkeit wird in diesem Zusammenhang nicht gesehen.

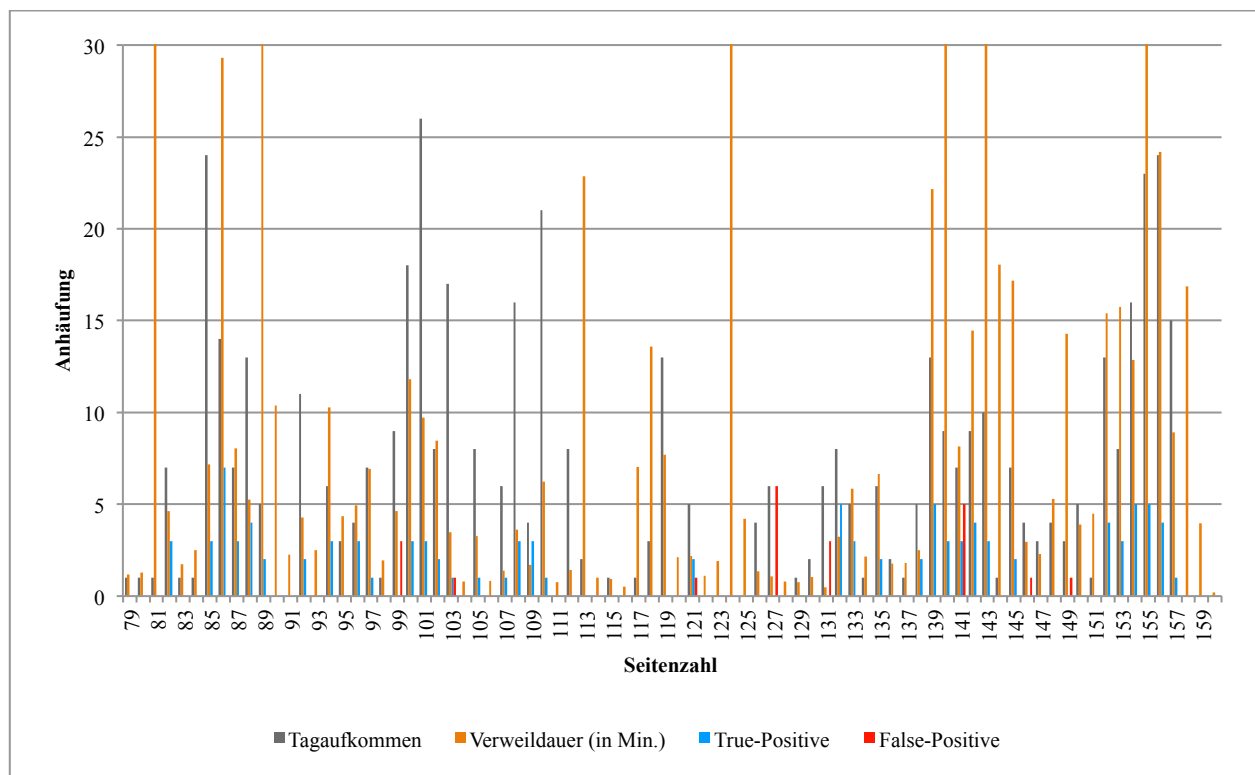


Abbildung 61 Klassifikation auf Grundlage Verweildauer, Tagaufkommen (Skript, online Simulation)⁸⁶

Interessanter ist allerdings die Betrachtung der fälschlicherweise als relevant klassifizierten Bereiche in Abbildung 61. Die Abbildung zeigt, dass Ressourcen aufgrund eines hohen Tagaufkommens falsch als relevant klassifiziert wurden, die Verweildauer für diese Bereiche allerdings besonders gering ist. Tabelle 20 fasst in diesem Zusammenhang die prägnantesten Beispiele aus Abbildung 61 zusammen. Es ist klar zu sehen, dass sich das Tagaufkommen von Ressourcen, die falsche Bereichsklassifizierungen aufweisen, innerhalb oder angrenzend zum angegebenen Schwellwertbereich befindet. Gleichzeitig ist die Verweildauer im Vergleich zu übrigen Dokumentenbereichen äußerst gering. Vor allem die Beispiele auf Seite 127 und 131 sind äußerst prägnant (vgl. Tabelle 20). Das Beispiel von Seite 141 ist hingegen weniger offensichtlich, weist allerdings auch eine zugleich unentschlossene Bewertung durch die Probanden auf. Innerhalb des Beispiels auf Seite 141 sieht man recht deutlich das Entscheidungsproblem, welches im Verlauf des Abschnittes diskutiert wurde.

⁸⁶ Für eine verbesserte Darstellung ist der maximale Wertebereich der X- und Y-Achse bewusst beschränkt, so dass die aussagekräftigsten Bereiche zu sehen sind.

Seite	Tagaufkommen	Verweildauer	True Positive	False Positive
99	9	4,6	0	3
127	6	1,1	0	6
131	6	0,4	0	3
141	7	8,3	3	5

Tabelle 20 Klassifikation auf Grundlage Verweildauer, Tagaufkommen (Auszug)

Man kann mit Bestimmtheit sagen, dass die Hinzunahme der Verweildauer als Relevanzfaktor in erster Linie zur Reduktion falscher Klassifikationen herangezogen werden kann. So gilt es während der Extraktion inhaltlich relevanter Bereiche zu prüfen, in wieweit das Tagaufkommen innerhalb des Schwellwertbereiches liegt und die Verweildauer einer hohen oder niedrigen Wertigkeit entspricht. Denkbar wäre in Kombination mit der Taghäufigkeit die bereits diskutierte mittlere Verweildauer. Die Ergebnisse zeigen zumindest nicht, dass wie zunächst angenommen, richtig klassifizierte Ressourcen hiervon Einfluss nehmen würden. Dennoch bedarf es weiterer Untersuchungen sinnvollerweise unter Berücksichtigung verschiedener Schwellwerte, um Unterschiede ausmachen zu können.

Ein weiterer offener zu diskutierender Aspekt innerhalb des algorithmischen Verfahrens ist ein Modell zur Bestimmung des Userscores. Der Userscore erhebt Einfluss auf die Bestimmung eines Ressourcenscores innerhalb der Datenanalyse des vorgestellten Verfahrens. Wie in Abschnitt 4.4.3 argumentiert ist für die vorliegende Arbeit der Userscore $S_U(u)$ für einen beliebigen User $u \in U$ im System mit $S_U(u) = 1$ bestimmt worden, so dass der Einfluss von Tags für alle User gleich gewichtet ist. Die Herleitung eines Modells zur Bestimmung eines geeigneten Userscores war auf Grundlage der dargelegten Literatur nicht möglich und erweist sich auch auf Grundlage erster erhobener Evaluationsergebnisse als wenig trivial. Eine erste denkbare Möglichkeit ist die Bestimmung eines Userscores auf Grundlage von Aktivitäten im System. Hierzu stellt Abbildung 62 die Ergebnisse der Testaufgaben der Benutzerstudie in Abhängigkeit von Nutzungsaktivitäten in Form von Tags dar⁸⁷. Abbildung 62 zeigt allerdings, dass die Nutzungsaktivität nicht mit den Ergebnissen der Testaufgaben korreliert. Ein hohes

⁸⁷ Lediglich die Testaufgaben in der Benutzerstudie wurden eingesammelt und bewertet, um innerhalb der online Simulation das reale Nutzungsverhalten durch eine Bewertungssituation nicht zu verfälschen.

Tagaufkommen ergibt sich sowohl für Personen mit guten als auch mit nicht ausreichenden Ergebnissen. Dabei wäre es wünschenswert Personen mit einem guten Verständnis für einen Sachverhalt einen höheren Einfluss beizumessen. Es ist somit denkbar die Herleitung eines Userscores auf Grundlage von Testergebnissen durchzuführen (z.B. $S_U(u) = 1$, sofern das Testergebnis von $u \in U$ 100% entspricht). Problematisch ist dieses Vorgehen allerdings dann, wenn Personen mit geringen Testergebnissen einen Bereich z.B. als schwierig einstufen. Dies bewirkt durch eine geringe Einflussnahme der Tags im System ggf. eine zu geringe Berücksichtigung der Rückmeldungen von Personen mit geringen Ergebnissen, deren Probleme mit Materialien allerdings darstellbar sein sollten.

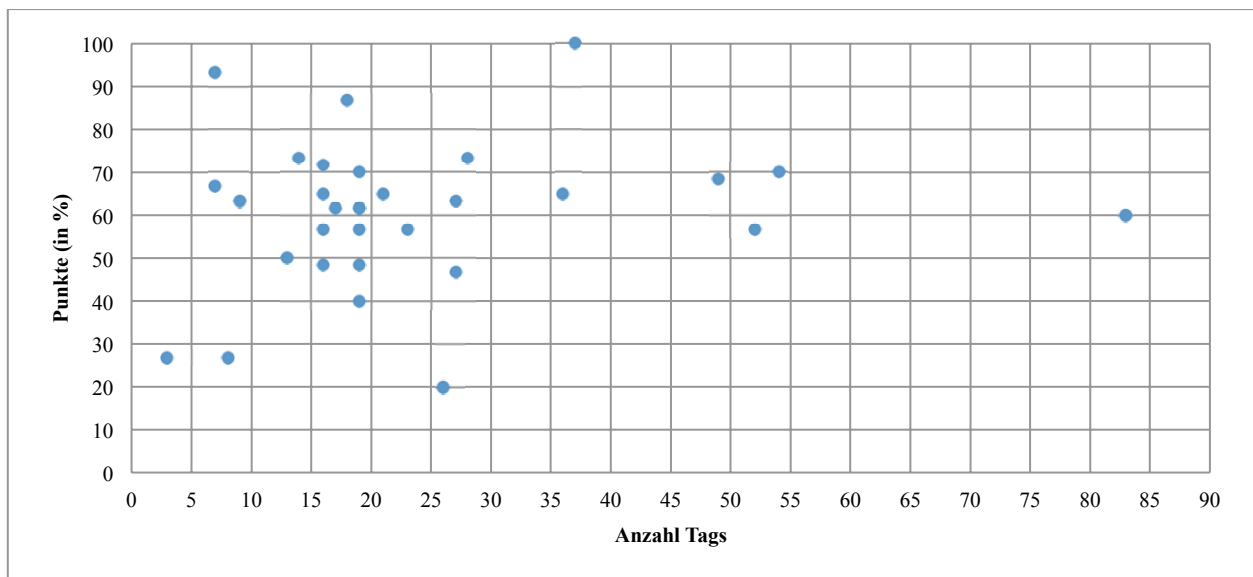


Abbildung 62 Auswertung Testaufgaben Benutzerstudie

Denkbar ist weniger die Herabstufung als die Mehrgewichtung von Personen im System, die sich durch bestimmte Aktivitäten als vertrauenswürdig erwiesen haben. Dieses Vorgehen wurde in der Literatur in ähnlicher Form durch die Herleitung von vertrauenswürdigen Lernpartnern beschrieben, die durch andere Lerner bewertet wurden (vgl. Abschnitt 3.5). Die Herleitung von Lernpartnern bedürfte einer grundlegenden Erweiterung des LAOs Systems. Vielmehr ist auch eine automatisierte Form der Herleitung des Userscores wünschenswert. Eine Möglichkeit, in der alle genannten Faktoren Berücksichtigung finden, ist Lernende aufgrund ihrer Aktivitäten im System (ggf. auch unter Berücksichtigung von weiteren Faktoren wie z.B. Testergebnissen) zu unterschiedlichen Lerngruppen zusammenzufassen und Empfehlungen in Abhängigkeit von

Gruppen zu generieren. Ansätze zur Formation und Bewertung von Lerngruppen existieren in der Literatur (vgl. z.B. [80, 83, 105]). Es wird an dieser Stelle weiterer Forschungsbedarf gesehen, um eine geeignete Adaption für das System LAOs herzuleiten.

Für die weitere Diskussion bedarf es an dieser Stelle einer erneuten Betrachtung der Bewertungen für Empfehlungen. Grundlegend sind die erhobenen Bewertungen positiv zu interpretieren, allerdings existiert auch in diesem Zusammenhang die Möglichkeit zur Optimierung. Die Evaluation hat gezeigt, dass mittelmäßige bis geringe Bewertungen – für als richtig relevant klassifizierte Bereiche – meist auf einem nicht Vorhandensein von Lernmaterialverbindungen oder einer geringen Anzahl an Tags als Datenbasis beruht. Die Herleitung von Lernmaterialverbindungen wird im weiteren Verlauf des Abschnitts noch genauer diskutiert. Die genügende Datenbasis für Empfehlungen wurde zudem bereits zuvor thematisiert und die Nutzung eines Schwellwertes genauer betrachtet. Des Weiteren ließe sich die Aussagekraft von Empfehlungen diskutieren. Die hergeleiteten Empfehlungen basieren auf der Aggregation von Tag Informationen, die in grafischer und textueller Form vorliegen. Die Inhalte der Empfehlungen wurden, auf Grundlage des Learning Object Review Instrument (LORI) abgeleitet. Die grafischen und textuellen Bestandteile werden durch die Präsentation entsprechender Maßnahmen sowie Materialverbindungen ergänzt. Die Ergebnisse der Evaluation zeigen auf, dass eine positive Darstellung für die Empfehlungen vorliegt. Dennoch wäre eine weitere Untersuchung sinnvoll, die vor allem die inhaltliche Aufbereitung der Empfehlungen bei Lernenden hinterfragt. In diesem Zusammenhang ist auch eine Neubewertung der Tagscores und der Grenzen zur Bestimmung der Ausprägungen zur Steigerung der Bewertungen denkbar. Die Argumentation zur Festlegung der Definition ergibt sich aus Abschnitt 4.2, wo die Ausprägung eines Tagscores in Zusammenhang mit der Klarheit der Aussage des Tags einhergeht. Die Evaluation zeigt allerdings, dass es vielmehr der Bereitstellung zusätzlicher (inhaltlicher) Informationen bedarf. Hierfür haben sich die hergeleitete Tagdefinition und auch die vordefinierten Grenzen als zweckdienlich erwiesen, um z.B. einen (kategorisierten) Ressourcenscore zu errechnen und Informationen in visueller oder textueller Form darzustellen. Es wird an dieser Stelle weniger der Bedarf einer Neudefinierung von Tagscores oder Grenzen gesehen, als vielmehr die Erhebung neuer, ergänzender Informationen, die es in weiteren Untersuchungen zu bestimmen gilt.

Die Bereitstellung von Lernmaterialverbindungen, die auf Grundlage von User-basierten Tags erhoben werden, ist hingegen ein Aspekt, den zu diskutieren gilt. Ganz offensichtlich ergeben sich für Empfehlungen keine Materialverbindungen, sofern diese nicht explizit durch die Usergruppe erhoben wurden. Die Vorgehensweise erwies sich innerhalb der Evaluations als durchaus praktikabel, allerdings ist von einer gleichen Aktivität Lernender nicht in jedem Lernszenario auszugehen. Es besteht weiterhin die Annahme, dass die Verbreitung von Informationen bzw. Lernmaterialien aus einem sozialen, kontextbezogenen Gefüge – z.B. innerhalb einer Gruppe Studierender – sinnvoll ist. Die Bewertungen für Empfehlungen mit Materialverbindungen lassen diesen Schluss zu und auch weitere Arbeiten befürworten ein etwaiges Vorgehen [30, 111, 114]. Dennoch ist eine alternative Strategie wünschenswert, die ggf. ohne explizite Rückmeldung Lernender auskommt. Ein möglicher Ansatzpunkt wurde für die Vernetzung zwischen bestehenden Materialien bereits im Verlauf der Arbeit benannt. Hierbei handelt es sich um die weitere Analyse des impliziten Nutzerverhaltens. Es besteht die Annahme, dass der häufige Wechsel (auf eine gesamte Gruppe bezogen) zwischen z.B. einer Seite im Skript und einem themenbezogenen Video einen inhaltlichen oder thematischen Zusammenhang darstellt. Auch die Anhäufung der Nutzungsdauer könnte Aufschluss über themenverwandte Materialien geben. Die Herleitung von Materialverbindungen über das implizite Nutzungsverhalten ist zu diesem Zeitpunkt allerdings nur eine These, die es für eine Umsetzung weiter zu untersuchen gilt. Für die Herleitung von Materialverbindungen aus dem Internet bedarf es hingegen der Zunahme weiterer Verfahrensarten, die im Verlauf der Arbeit bereits ausführlich präsentiert wurden. Die Herleitung einer hybriden Verfahrensweise ist ebenfalls ein Anknüpfungspunkt für weitere Forschungstätigkeiten.

Die aktive Nutzung des Systems ist für das beschriebene Verfahren von zentraler Bedeutung. Wie für Materialverbindungen bereits dargelegt, ist der Empfehlungsdienst des Systems ohne aktive Beteiligung der Lernenden nicht funktional. In diesem Zusammenhang handelt es sich um die generelle Nutzung des Systems. Das aktive setzen von Tags dient als Grundlage für das Auffinden relevanter Bereiche und die Herleitung von Empfehlungen. Auch hier gilt weiterhin die Annahme, dass aufgrund kollektiver Intelligenz die sinnvolle Herleitung entsprechender Empfehlungen möglich ist. Die dargelegten Evaluationsergebnisse belegen dies. In diesem Zusammenhang ist die Umsetzung eines hybriden Verfahrens allerdings *nicht* wünschenswert.

Es wurde sich bewusst dafür entschieden das System auf Grundlage von User-basierten Tags zu implementieren, da so erst die genügende Bewertung von inhaltlichen Bereichen möglich ist. Des Weiteren scheint es durchaus sinnvoll die Lernenden aktiv in das System mit einzubinden, so dass Empfehlungen nicht ausschließlich konsumiert, sondern im besten Fall mitgestaltet werden. Die Einbettung des Systems in ein Lehr-/Lernszenario ist eine Möglichkeit die Nutzung des Systems zu forcieren. Diese Strategie ist zumindest innerhalb der Evaluation aufgegangen. Des Weiteren hat die Akzeptanzbefragung gezeigt, dass Lernende durchaus gewillt sind ein etwaiges System in der Lehre nutzen zu wollen. Wie bereits beschrieben, basiert die aktive Nutzung auf Grundlage von Akzeptanz für ein System, die wiederum von Aufwand und Qualität des Empfehlungsdienstes abhängt. Von einer grundlegenden Akzeptanz wird, bedingt durch die erhobenen Ergebnisse der Evaluation, ausgegangen. Die weitere Steigerung der Akzeptanz ist allerdings ein möglicher Ansatzpunkt, um Lernende stärker zu motivieren das System zu nutzen. So hat sich in diesem Zusammenhang herausgestellt, dass die zur Verfügung gestellten Werkzeuge des Systems einer Verbesserung bedürfen. Innerhalb der Akzeptanzbefragung wurde die Auswahl der Werkzeuge zu einem großen Anteil lediglich neutral bewertet. Es bedarf somit weiterer Untersuchungen, die die Bedürfnisse hinsichtlich zur Verfügung stehender Werkzeuge aus Sicht eines Lerners hinterfragen. Bedingt durch die sonst positiven Rückmeldungen hinsichtlich des Umgangs mit dem System, wird zunächst lediglich die Weiterentwicklung der Werkzeuge als mögliches Handlungsfeld gesehen.

Als weiteren Punkt gilt es die Adaptierbarkeit des Systems zu diskutieren. Innerhalb der Evaluation wurde gezeigt, dass das System im Kontext einer universitären Lehrveranstaltung Verwendung finden kann. Es stellt sich an dieser Stelle die Frage, ob das System außerhalb des Hochschulkontextes ebenso Verwendung finden kann und in wieweit eine fächerübergreifende Adaption möglich ist. Die Verwendung außerhalb der universitären Lehre wird nur eingeschränkt gesehen. Für die Arbeit mit dem System und für die Interpretation der Empfehlungen bedarf es bei den Lernenden ein gewisses Maß an Eigenständigkeit. In Szenarien, in denen es einer gezielten Anleitung der Lehrenden bedarf, ist der Einsatz eines etwaigen Systems ggf. ungeeignet. Denkbar ist allerdings der Einsatz in Weiterbildungsformaten, die nicht zwangsläufig dem Curriculum eines Studiums entstammen. So bietet die Literatur z.B. die Verwendung von Empfehlungsdiensten in MOOC Szenarien an. Der Einsatz des in der

vorliegenden Arbeit präsentierten Verfahrens ist dabei allerdings nicht in jeder Form eines MOOCs sinnvoll. CMOOCs liegt so i.d.R. ein didaktisches Konzept zum Erlernen von Inhalten zu Grunde und so existieren explizit aufbereitete Lernmaterialien, für die es keiner weiteren Aufbereitung bedarf. XMOOCs hingegen sind oftmals wie klassische Vorlesungen aufgebaut und der Teilnehmer beschäftigt sich ggf. grundlegend eigenständig mit den angebotenen Lernmaterialien [155]. Unabhängig des Einsatzszenarios scheint allerdings von großer Bedeutung, dass für die Nutzung immer ein geeigneter Lernkontext vorliegt. Dies wurde bereits mehrfach in der vorliegenden Arbeit beschrieben und entsprechend argumentiert. Die fächerübergreifende Nutzung des Systems ist damit noch nicht beantwortet, allerdings gilt der konkrete Kontextbezug in diesem Zusammenhang ebenfalls als maßgebliche Anforderung. Die mögliche Nutzung ist aus persönlicher Sicht auch nicht grundlegend von einem fachlichen Inhalt abhängig, sondern bedarf vielmehr der Berücksichtigung des durchzuführenden Lernszenarios. Wie die Evaluation zeigt, ist die Nutzung des Systems in einem Lernszenario, in dem ein eigenständiges Arbeiten von den Lernenden verlangt wird, denkbar. Möglichst sinnvoll ist auch die einheitliche Vorgabe für alle Lernenden hinsichtlich des Lernkontextes und der zugrunde liegenden Lernmaterialien, die z.B. in seminaristischen Veranstaltungen ggf. nicht vorliegen.

Final ist zu diskutieren, wann Empfehlungen für User ausgesprochen werden. Innerhalb der Evaluation hat jeder Proband jegliche Empfehlung innerhalb der Lernmaterialien erhalten. Es konnte davon ausgegangen werden, dass die Probanden sich innerhalb der Evaluation mit dem Lehrstoff und den entsprechenden Materialien auseinandergesetzt haben und die Beurteilung der Empfehlungen deshalb möglich war. In einem Szenario, welches beispielsweise über die Dauer eines ganzen Semesters mit einer Vielzahl an Übungen durchgeführt wird, ist eine Regulierung der Empfehlungen sinnvoll. Verschiedene Möglichkeiten sind in diesem Zusammenhang denkbar. Bedeutsam scheint vor allem zu sein, dass sich ein jeweils Lernender mit einem Lerninhalt vertraut gemacht hat, bevor er oder sie Zugriff auf eine Empfehlung erhält. Da das System Informationen darüber besitzt, wie lange ein User sich im System mit einer Stelle im Lernmaterial auseinandergesetzt hat, kann über die impliziten Tags eine entsprechende Abfrage erfolgen. Ob es in diesem Zusammenhang der Berücksichtigung einer konkreten Zeitvorgabe oder lediglich der genügenden Anzahl an Aufrufen z.B. von einer Seite bedarf, bleibt an dieser Stelle noch zu prüfen. Andersherum wäre denkbar, dass ein Lernender eine Empfehlung erhält,

sofern er oder sie sich nicht genügend bzw. gar nicht mit einem bestimmten Bereich auseinandergesetzt hat. In diesem Zusammenhang bleibt zu prüfen, in wieweit sich beide Möglichkeiten gegeneinander ausschließen bzw. ob die Umsetzung beider Optionen sinnvoll ist. Ebenfalls möglich ist die Begutachtung der Tags, die Lernende für inhaltliche Bereiche gesetzt haben. Geht aus dem Tag z.B. eine gegenteilige Aussage gegenüber der restlichen Gruppe hervor, so könnte das System geeignet intervenieren und eine Empfehlung präsentieren. Die Regulierung von Empfehlungen aus Sicht eines Lehrenden ist ebenfalls ein sinnvoller Gedanke. Dies würde allerdings weniger für einzelne User im System als vielmehr für die Gesamtgruppe erfolgen. Ein Lehrender könnte so z.B. die Möglichkeiten erhalten Empfehlungen oder auch alternative Materialien aus dem System auszuschließen, da er oder sie eine Empfehlung oder ein Material ggf. nicht als angemessen empfindet. Diese Variante wird allerdings in erster Linie als optionale Möglichkeit für Lehrende in Betracht gezogen.

Ohnehin gilt es die Perspektive der Lehrenden innerhalb des Systems weiter zu untersuchen. Abschnitt 3.5.4 hat bereits Beispiele aus der aktuellen Literatur präsentiert, die Mehrwerte für Lehrende im Kontext etwaiger Systeme generieren. Die Literatur hat aufgedeckt, dass vor allem das Nutzungsverhalten von Lernenden für Lehrpersonen von Interesse ist. Das System LAOs kommt dieser Anforderung im Kontext von Lernmaterialien nach, weshalb zunächst davon ausgegangen werden kann, dass Lehrende in den zur Verfügung gestellten Informationen einen Mehrwert sehen. Neben der Nutzungshäufigkeit von Materialien bzw. Materialbereichen wird die inhaltliche Einschätzung bzw. Bewertung der Lernenden als wertvolle Information für Lehrende gesehen. Ob und in welchem Maße diese Annahme tatsächlich zutrifft, gilt es in weiteren Untersuchungen zu prüfen. Darüber hinaus besteht das Potential die Menge an bereitgestellten Informationen zu erweitern und diese in kontextbezogener Form (z.B. in Zusammenhang mit Testaufgaben) für Lehrende bereitzustellen. Auch die gezielte (visuelle) Aufbereitung von Informationen könnte ein für Lehrende zweckdienlich sein. Denkbar in diesem Zusammenhang ist z.B. als schwierig bewertete Bereiche in einer Schnellübersicht zu präsentieren, um einen vereinfachten Überblick zu ermöglichen. Gleiches ist ebenfalls für die übrigen Kategorien denkbar. Aber auch in diesem Zusammenhang sind weitere Untersuchungen sinnvoll, die den Bedarf von Lehrenden gezielt aufdecken.

9 Fazit & Ausblick

Wie die vorliegende Arbeit gezeigt hat, ist das Lernen im digitalen Zeitalter nicht immer unproblematisch. Bedingt durch die Menge an verfügbaren digitalen Lernmaterialien im Internet ist die Suche nach alternativen Informationen für Lernende häufig ein schwer zu überwindendes Hindernis. Nicht zuletzt ist auch die Komplexität von Lernmaterialien innerhalb der universitären Lehre ein Grund dafür, dass sich eine Suche im Internet als zeitaufwändig und nicht immer erfolgreich erweist. Der Bedarf an alternativen Lernmaterialien existiert aus Sicht von Lernenden ungeachtet der Tatsache, dass bereits themenspezifische und häufig aufbereitete Lernmaterialien durch jeweils Lehrende zur Verfügung gestellt werden. Die vorliegende Arbeit hat gezeigt, dass ein Bedarf für assistierende Systeme im Bereich des digitalen Lernens existiert, die Lernende bei der Auswahl von geeigneten Inhalten und alternativen Lernmaterialien unterstützen, um sowohl die Nutzung von bestehenden Lernmaterialien als auch das Auffinden alternativer Materialien zu vereinfachen. In diesem Zusammenhang wurde die Entwicklung eines Empfehlungsdienstes präsentiert, der auf Grundlage eines kollaborativen Tagging Ansatzes inhaltliche Bereiche in bestehenden Lernmaterialien extrahiert und diese mit neuen Materialien aus dem Internet verknüpft. Die Entwicklung umfasste dabei sowohl die Implementierung einer Tagging Komponente, welche grundlegend als Userinterface zu betrachten ist, als auch die Implementierung einer Analysekomponenten zur Herleitung von Empfehlungen. Lernende erhalten innerhalb des Systems die Möglichkeit Lernmaterialien online zu bearbeiten und Inhalte dabei zu klassifizieren. Das System stellt hierfür entsprechende Tagging Werkzeuge bereit, die die Klassifizierung sowohl von Text- als auch Videoinhalten ermöglichen. Auf Grundlage der erhobenen Klassifizierung leitet der Ansatz auf einer Folksonomy als Datenbasis in drei Schritten Empfehlungen her. Neben der Extraktion von inhaltlich relevanten Bereichen (Schritt 1) führt der Ansatz eine Bewertung durch (Schritt 2), über die sich die Herleitung von

Empfehlungen inklusive alternativer Materialverknüpfungen ergibt (Schritt 3). Genannte Empfehlungen stehen nach der Herleitung für die Usergruppe im Userinterface zur Verfügung. Des Weiteren stellt das System einen Analysebereich für Lehrende zur Verfügung, in dem die Empfehlungen eingesehen und das Nutzungsverhalten von Lernenden analysiert werden kann.

Die Entwicklung des Ansatzes basiert auf der Darlegung eines Überblicks der aktuellen Forschungsliteratur, der bereits bestehende Verfahren gegenüberstellt. In diesem Zusammenhang wurde gezeigt, dass zwar verschiedene Verfahren mit ähnlichen Intentionen existieren, die allerdings nicht den vollen Zielvorgaben der vorliegenden Arbeit genügen. Diese bestanden im Aufdecken relevanter Inhalte in bestehenden Lernmaterialien und der Verknüpfung von Alternativen aus dem Internet. Für relevante Inhalte sollten zudem Kontextinformationen bzw. Hilfestellungen ableitbar sein. Des Weiteren lag der Fokus auf der unkomplizierten Eingliederung des Systems in bestehende Lernszenarien. Es bedurfte somit einer Entwicklung, die verschiedene Teilaspekte aus der aktuellen Forschungsliteratur kombiniert, um den Zielvorgaben der Arbeit zu genügen. Eine Evaluation bestehend aus zwei Szenarien hat dabei gezeigt, dass das System unterstützende Empfehlungen für Lernende innerhalb eines wohldefinierten Lernkontextes herleitet, ohne dabei Abstriche hinsichtlich der Zielvorgabe ausgemacht zu haben. Die Evaluation konnte zudem Abhängigkeiten zwischen den verschiedenen Schritten innerhalb des Empfehlungsprozesses ausmachen. Das Aufdecken der Abhängigkeiten ermöglichte die Herleitung von Schwellwerten, die es fortan für die Herleitung von Empfehlungen zu berücksichtigen gilt. Durch den Nachweis einer grundlegenden Akzeptanz für das System ist des Weiteren ein wesentlicher Grundstein für die tatsächliche Nutzung des Systems geebnet worden. Sowohl die Qualität der Empfehlungen als auch der Aufwand sie zu erlangen fanden innerhalb der Akzeptanzbefragung Berücksichtigung.

Die Diskussion der Arbeit hat dennoch einige Möglichkeiten zur Optimierung offen gelegt, die vor allem in der Anpassung des Systems hinsichtlich der Bedürfnisse Lernender zu sehen sind. Die Optimierung des Klassifikationsprozesses für inhaltlich relevante Bereiche sowie die Optimierung des Herleitungsprozesses für Empfehlungen wurde ebenfalls diskutiert. In diesem Zusammenhang ist der Bedarf bzw. die Möglichkeit der Optimierung deutlich eingeschränkt, da die Ergebnisse der Evaluation bereits ausgeprägte positive Ergebnisse nachgewiesen haben.

Hierbei bedarf es vielmehr der Durchführung weiterer Studien, die die Ergebnisse unter Berücksichtigung der erhobenen Schwellwerte bestätigen. Lediglich für die Herleitung von Empfehlungen wurden verschiedene Aspekte diskutiert, die es in weiteren Forschungsaktivitäten zu untersuchen gilt. Vor allem die Adaption zur Herleitung eines Userscores ist in diesem Zusammenhang von großer Bedeutung. Verschiedene Möglichkeiten wurden dafür in der Diskussion der Arbeit vorgeschlagen.

Die vorliegende Arbeit hat sowohl Aspekte der bestehenden Forschungsliteratur bestätigt als auch Möglichkeiten für die weitere Forschung eröffnet. So konnte durch die erhobenen Evaluationsergebnisse der sinnvolle Einsatz von Tags zur Klassifizierung von (Lern-)Inhalten aus bestehenden Forschungsarbeiten bestätigt werden. Auch der Einsatz von Tags zur Verknüpfung von Lerninhalten hat sich als sinnvoll erwiesen. Des Weiteren hat die Evaluation weitere Erkenntnisse über das Nutzungsverhalten Lernender mit Lernmaterialien erbracht und so wurden die erhobenen Erkenntnisse sowohl aus verwandter Literatur als auch aus der Vorstudie in Kapitel 2 bestätigt. Im Zusammenhang mit dem Nutzungsverhalten hat sich ebenso gezeigt, dass die Verwendung von impliziten Nutzerinformationen einen wertvollen Mehrwert bei der Analyse von Daten insbesondere zur Herleitung von Empfehlungen mit sich bringt. Auch dieser Sachverhalt entspricht bereits diskutierten Erkenntnissen aus der Literatur.

Gänzlich neue Möglichkeiten ergeben sich durch das implementierte Verfahren. Die vorliegende Arbeit hat gezeigt, dass die hierarchische Clusterung von Tags eine denkbare Möglichkeit ist, um zusammengehörige Informationsgruppen zusammenzuschließen. In grundlegender Form deckt sich dieses Erkenntnis bereits für Datenmengen, die z.B. auf Bewertungen oder Ähnlichkeiten beruhen. Es ist durchaus denkbar, dass das Verfahren für weitere Lösungsansätze im Zusammenhang mit Tags adaptierbar ist. Innerhalb des Verfahrens wurde zudem ein bislang unerprobtes Bewertungsschema zur Klassifizierung von Lerninhalten verwendet. Die Abweichung zu sonst üblichen Tag-basierten Ansätzen besteht darin, dass eine vordefinierte Menge an Tags in unterschiedlichen Formen, basierend auf vier Kategorien zur Verfügung stehen. Für das System hat sich diese Form der Bewertung als zweckdienlich erwiesen und wurde von den Lernenden positiv aufgenommen. Neben der Möglichkeit zur Optimierung des Bewertungsschemas ergeben sich weitere Möglichkeiten das Bewertungsschema für verwandte

Ansätze zu adaptieren, um dargelegte Probleme wie z.B. die Interpretation von Tags zu umgehen. Für das implizite Nutzungsverhalten von Lernenden im Kontext von Lernmaterialien hat die vorliegende Arbeit aufgezeigt, dass neben Indikatoren der Relevanz auch die Herleitung von weiteren Informationen, wie z.B. themenverwandte Stellen zwischen Lernmaterialien, denkbar sind, die es in weiteren Forschungsaktivitäten zu untersuchen gilt. Als weitere Möglichkeit hat die Arbeit die Analyse des Nutzungsverhaltens aus Sicht eines Lehrenden präsentiert. Die bestehende Literatur zeigt, dass Lehrende durchaus Mehrwerte aus der Analyse des Benutzerverhaltens Lernender ziehen. Für das System LAOs ergeben sich dennoch Anknüpfungspunkte für weitere Entwicklungen sowie Untersuchungen, die mögliche Mehrwerte für Lehrende offenlegen und bestätigen.

10 Literaturverzeichnis

- [1] Adomavicius, G. and Tuzhilin, A. 2008. Context-aware Recommender Systems. *Proceedings of the 2008 ACM Conference on Recommender Systems* (New York, NY, USA, 2008), 335–336.
- [2] Adomavicius, G. and Tuzhilin, A. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*. 17, 6 (2005), 734–749.
- [3] Agrawal, A. et al. 2015. YouEDU : Addressing Confusion in MOOC Discussion Forums by Recommending Instructional Video Clips. *Proceedings of the 8th International Conference on Educational Data Mining* (Madrid, Spain, 2015), 297–304.
- [4] Ahn, J. et al. 2007. Open user profiles for adaptive news systems: help or harm? *Proceedings of the International Conference on World Wide Web (WWW '07)* (Banff, Alberta, Canada, 2007), 11–20.
- [5] Akpınar, Y. 2008. Validation of a Learning Object Review Instrument: Relationship between Ratings of Learning Objects and Actual Learning Outcomes. *Interdisciplinary Journal of E-Learning and Learning Objects*. 4, (2008), 291–302.
- [6] Anaya, A.R. et al. 2015. An approach of collaboration analytics in MOOCs using social network analysis and influence diagrams. *Proceedings of the 8th International Conference on Educational Data Mining* (Madrid, Spain, 2015), 492–495.
- [7] Ansari, A. et al. 2000. Internet Recommendation Systems. *Journal of Marketing Research*. 37, 3 (2000), 363–375.
- [8] Baeza-Yates, R. and Ribeiro-Neto, B. 1999. *Modern information retrieval*. Addison-Wesley.
- [9] Balabanović, M. and Shoham, Y. 1997. Fab: content-based, collaborative recommendation. *Communications of the ACM*. 40, 3 (Mar. 1997), 66–72.
- [10] Basu, C. et al. 1998. Recommendation as Classification : Using Social and Content-Based Information in Recommendation. *Proceedings of the 15th National Conference on Artificial Intelligence* (Madison, Wisconsin, 1998), 714–720.
- [11] Billsus, D. and Pazzani, M.J. 1999. A Hybrid User Model for News Story Classification. *Courses and Lectures-International Centre for Mechanical Sciences*. (1999), 99–108.

-
- [12] Billsus, D. and Pazzani, M.J. 1998. Learning collaborative information filters. *Proceedings of the 15th International Conference on Machine Learning* (Madison, Wisconsin, 1998), 46–54.
- [13] Billsus, D. and Pazzani, M.J. 2000. User modeling for adaptive news access. *User Modelling and User-Adapted Interaction*. 10, 2–3 (2000), 147–180.
- [14] Bobadilla, J. et al. 2009. Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems*. 22, 4 (2009), 261–265.
- [15] Bortz, J. and Döring, N. 2006. *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Springer.
- [16] Bouza, A. et al. 2008. SemTree: Ontology-based decision tree algorithm for recommender systems. *Proceedings of the 2007 International Conference on Posters and Demonstrations* (2008), 106–107.
- [17] Breese, J.S. et al. 1998. Empirical analysis of predictive algorithms for collaborative filtering. *Proceedings of the 14th conference on Uncertainty in Artificial Intelligence*. 461, 8 (1998), 43–52.
- [18] Brin, S. and Page, L. 1998. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*. 30, 1–7 (Apr. 1998), 107–117.
- [19] Broisin, J. et al. 2010. A personalized recommendation framework based on cam and document annotations. *Procedia Computer Science* 1. 1, 2 (Jan. 2010), 2839–2848.
- [20] Brooks, C. 2012. *A Data-Assisted Approach to Supporting Instructional Interventions in Technology Enhanced Learning Environments (Dissertationsschrift)*. Department of Computer Science, University of Saskatchewan.
- [21] Bürg, O. and Mandl, H. 2005. Akzeptanz von E-Learning in Unternehmen. *Zeitschrift für Personalpsychologie*. 4, 2 (2005), 75–85.
- [22] Burke, R. 2002. Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*. 12, 4 (2002), 331–370.
- [23] Burke, R. 2000. Knowledge-based recommender systems. *Encyclopedia of library and information systems*. 69, Supplement 32 (2000), 175–186.
- [24] Casali, A. et al. 2011. Recommender System for Personalized Retrieval of Learning Objects. *Book of Educational Recommender Systems and Technologies: Practices and Challenges*. O.C. Santos and J.G. Boticario, eds. IGI Global. 182–210.
- [25] Chen, J.M. et al. 2012. A Hybrid Tag-Based Recommendation Mechanism to Support Prior Knowledge Construction. *In Proceedings of IEEE 12th International Conference on Advanced Learning Technologies (ICALT)* (Rome, Italy, 2012), 23–25.
- [26] Chen, W. et al. 2014. A hybrid recommendation algorithm adapted in e-learning environments. *World Wide Web*. 17, 2 (2014), 271–284.
- [27] Chi, E.H. and Mytkowicz, T. 2008. Understanding the efficiency of social tagging systems using information theory. *Proceedings of the 19th ACM Conference on Hypertext and Hypermedia* (Pittsburgh, PA, 2008), 81–88.

-
- [28] Clark-Wilson, A. and Rothkrantz, L. 2016. Computer Supported Education of the 21st Century: Meeting the Challenges of Today, Supporting Innovations and Developing a Sustainable Future. *Presentation at the 8th International Conference on Computer Supported Education* (Rome, Italy, 2016).
- [29] Claypool, M. et al. 1999. Combining content-based and collaborative filters in an online newspaper. *Proceedings of Recommender Systems Workshop at ACM SIGIR* (Berkeley, CA, 1999), 40–48.
- [30] Cristian, M. et al. 2015. Intelligent Tutor Recommender System for On-Line Educational Environments. *Proceedings of the 8th International Conference on Educational Data Mining* (Madrid, Spain, 2015), 516–519.
- [31] Davis, F.D. 1989. Perceived Usefulness, Perceived Ease Of Use, And User Acceptance of Information Technology. *MIS Quarterly*. 13, 3 (1989), 319–339.
- [32] Deshpande, M. and Karypis, G. 2004. Item-based top-N recommendation algorithms. *ACM Transactions on Information Systems*. 22, 1 (2004), 143–177.
- [33] Desrosiers, C. and Karypis, G. 2011. A Comprehensive Survey of Neighborhood-based Recommendation Methods. *Recommender Systems Handbook*. Springer. 107–144.
- [34] Doush, I.A. et al. 2012. Annotations, Collaborative Tagging, and Searching Mathematics in E-Learning. *International Journal of Advanced Computer Science and Applications*. 2, 4 (2012), 30–39.
- [35] Dron, J. et al. 2000. CoFIND: steps towards a self-organising learning environment. *Proceedings of WebNet World Conference on the WWW and Internet 2000* (Chesapeake, VA, 2000), 146–151.
- [36] Dyckhoff, A.L. et al. 2012. Design and implementation of a mobile application for personal learning analytics. *Educational Technology & Society*. 15, (2012), 58–76.
- [37] Eck, D. et al. 2007. Automatic generation of social tags for music recommendation. *Proceedings of the 20th International Conference on Neural Information Processing Systems* (Vancouver, BC, 2007), 385–392.
- [38] elearnspace - George Siemens Blog: 2016. <http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>. Accessed: 2016-10-25.
- [39] Emadzadeh, E. et al. 2010. Learning Materials Recommendation Using a Hybrid Recommender System with Automated Keyword Extraction. *Word Applied Sciences* 9. 11, (2010), 1260–1271.
- [40] Engelbert, B. et al. 2013. A folksonomy-based recommender system for learning material prediction. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2013* (Victoria, BC, Canada, 2013), 1590–1595.
- [41] Engelbert, B. et al. 2011. A User Supporting Personal Video Recorder Based on a Generic Bayesian Classifier and Social Network Recommendations. *J.J. Park, L.T. Yang, and C. Lee (Eds.): FutureTech 2011, Part II, CCIS 185*. (2011), 1–8.

-
- [42] Engelbert, B. et al. 2011. A user supporting personal video recorder by implementing a generic Bayesian classifier based recommendation system. *In Proceedings of the 2011 IEEE International Conference on Pervasive Computing and Communications Workshops* (Seattle, WA, USA, 2011), 567–571.
- [43] Engelbert, B. et al. 2016. Content Assistance and Recommendations in Learning Material - A Folksonomy-Based Approach. *In Proceedings of the 8th International Conference on Computer Supported Education (CSEDU 2016) - Volume 1* (2016), 456–463.
- [44] Engelbert, B. et al. 2013. The Use of User-based Video-Tags in a Recommender System Scenario for Learning Material Prediction. *In Proceedings of DIVERSE Conference 2013* (Osnabrück, 2013).
- [45] Engelbert, B. et al. 2013. Zwischen Nutzung und Nutzen – Die Suche nach geeigneten Lernmaterialien und deren Mehrwerte im Kontext einer Informatikveranstaltung. *GMW 2014 - Lernräume gestalten - Bildungskontexte vielfältig denken* (Zürich, 2013), 508–519.
- [46] Erdt, M. et al. 2015. Evaluating Recommender Systems for Technology Enhanced Learning: A Quantitative Survey. *IEEE Transactions on Learning Technologies*. 1382, c (2015), 326–344.
- [47] Farzan, R. and Brusilovsky, P. 2011. Encouraging user participation in a course recommender system: An impact on user behavior. *Computers in Human Behavior*. 27, 1 (2011), 276–284.
- [48] Firan, C.S. et al. 2007. The Benefit of Using Tag-Based Profiles. *Proceedings of the 2007 Latin American Web Conference (LA-WEB 2007)* (Washington, DC, 2007), 32–41.
- [49] Friesen, N. 2004. Three Objections to Learning Objects and E-Learning Standards. *Online Education Using Learning Objects*. R. McGreal, ed. RoutledgeFalmer. 59–70.
- [50] Fu, X. et al. 2007. Video Annotation in a Learning Environment. *Proceedings of the American Society for Information Science and Technology*. 43, 1 (Oct. 2007), 1–22.
- [51] Furnas, G.W. et al. 1987. The Vocabulary Problem in Human-System Communication. *Communications of the ACM*. 30, 11 (1987), 964–971.
- [52] Furnas, G.W. et al. 2006. Why Do Tagging Systems Work? *In Proceedings of CHI'06 extended abstracts on Human factors in Computing Systems* (Montreal, Quebec, Canada, 2006), 36–39.
- [53] Gemmis, M. De et al. 2009. Preference Learning in Recommender Systems. *Preference Learning*. 41, (2009).
- [54] Golder, S. and Huberman, B. 2006. Usage patterns of collaborative tagging systems. *Journal of Information Science*. 32, 2 (2006), 198–208.
- [55] Good, N. et al. 2006. Combining Collaborative Filtering with Personal Agents for Better Recommendations Nathaniel. *Tetrahedron*. 62, 37 (2006), 8805–8813.
- [56] Graf, S. et al. 2011. AAT: a tool for accessing and analysing students' behaviour data in learning systems. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (Banff, Alberta, Canada, 2011), 174–179.

-
- [57] Gunawardana, A. and Shani, G. 2009. A Survey of Accuracy Evaluation Metrics of Recommendation Tasks. *The Journal of Machine Learning Research*. 10, (2009), 2935–2962.
- [58] Guy, I. et al. 2010. Social media recommendation based on people and tags. *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (Geneva, Switzerland, 2010), 194–201.
- [59] Hamborg, K.-C. et al. 2012. Akzeptanz von Lehrveranstaltungsaufzeichnungen – Befunde aus zwei empirischen Studien. *Desel, J., Haake, J. M. & Spannagel, C. (Hrsg.), DeLFI 2012 – Die 10. e-Learning Fachtagung Informatik*. (Hagen und Heidelberg, 2012), 63–74.
- [60] Helou, S. El et al. 2010. The 3A Personalized, Contextual and Relation-based Recommender System. *Journal of Universal Computer Science*. 16, 16 (2010), 2179–2195.
- [61] Herlocker, J.L. et al. 2004. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*. 22, 1 (2004), 5–53.
- [62] Hill, W. et al. 1995. Recommending and evaluating choices in a virtual community of use. *Proceedings of the SIGCHI conference on Human factors in computing systems*. (1995), 194–201.
- [63] Hofmann, T. 1999. Probabilistic latent semantic indexing. *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '99* (1999), 50–57.
- [64] Höhfeld, S. 2007. Empfehlungssysteme aus Informationswissenschaftlicher Sicht – State of the Art. *Information Wissenschaft und Praxis*. 58, 5 (2007), 265–276.
- [65] Hotho, A. et al. 2006. FolkRank: A ranking algorithm for folksonomies. *Workshop Information Retrieval 2006 of the Special Interest Group Information Retrieval (FGIR 2006)* (Hildesheim, Germany, 2006), 111–114.
- [66] Hotho, A. et al. 2006. Information retrieval in folksonomies: Search and ranking. *Proceedings of the 3rd European conference on The Semantic Web: research and applications* (Budva, Montenegro, 2006), 411–426.
- [67] Höver, K.M. et al. 2010. Studierende, das Web und Vorlesungsaufzeichnungen. *DeLFI 2010 8 Jahrestagung der Fachgruppe E-Learning der Gesellschaft für Informatik e.V.* (2010), 121–132.
- [68] Hu, R. and Pu, P. 2009. Acceptance issues of personality-based recommender systems. *Proceedings of the 3rd ACM Conference on Recommender Systems (RecSys 2009)* (New York, NY, 2009), 221–224.
- [69] IEEE Learning Technology Standardization Committee 2002. IEEE Standard for Learning Object Metadata (Draft). IEEE Standard 1484.12.1.
- [70] Jannach, D. 2008. Finding preferred query relaxations in content-based recommenders. *Studies in Computational Intelligence*. 109, September (2008), 81–97.
- [71] Jäschke, R. et al. 2007. Tag Recommendations in Folksonomies. *Knowledge Discovery in*

- Databases: PKDD 2007*. 4702, (2007), 506–514.
- [72] Jawaheer, G. et al. 2010. Comparison of implicit and explicit feedback from an online music recommendation service. *Proceedings of the 1st International Workshop on Information Heterogeneity and Fusion in Recommender Systems - HetRec '10*. (2010), 47–51.
- [73] Ji, A. et al. 2007. Collaborative Tagging in Recommender Systems. *AI 2007: Advances in Artificial Intelligence*. 4830, (2007), 377–386.
- [74] John, A. et al. 2006. Collaborative tagging and expertise in the enterprise. *In Proceedings of the Collaborative Web Tagging Workshop Held in Conjunction With WWW 2006* (Edinburgh, Scotland, 2006).
- [75] Johnson, L. et al. 2012. *NMC Horizon Report: 2012 Higher Education Edition*.
- [76] Johnson, M. et al. 2011. The EDM Vis Tool. *Proceedings of the 4th International Conference on Educational Data Mining* (Eindhoven, Netherlands, 2011), 349–350.
- [77] Jones, N. and Pu, P. 2009. User Acceptance Issues In Music Recommender Systems. *EPFL Technical Report HCI-REPORT-2009-001*. (2009).
- [78] Kennedy, G. et al. 2014. Completing the loop: Returning learning analytics data to teachers. *Rhetoric to Reality: Critical Perspectives on Educational Technology. Proceedings Ascilite* (2014), 436–440.
- [79] Khalil, M. et al. 2016. Engaging Learning Analytics in MOOCs: the good, the bad, and the ugly. *In proceedings of the International Conference on Education and new Developments* (Ljubljana, Slovenia, 2016).
- [80] Khalil, M. and Ebner, M. 2016. Clustering patterns of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to reveal student categories. *Journal of Computing in Higher Education*. (2016), 1–9.
- [81] Khribi, M.K. et al. 2009. Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval. *Journal of Educational Technology & Society*. 12, 4 (2009), 30–42.
- [82] Klahold, A. 2009. *Empfehlungssysteme*. Vieweg+Teubner.
- [83] Konert, J. et al. 2013. GroupAL: ein Algorithmus zur Formation und Qualitätsbewertung von Lerngruppen in E-Learning-Szenarien mittels n-dimensionaler Gütekriterien. *DeLFI 2013 – Die 11. e-Learning Fachtagung Informatik* (Bonn, 2013), 71–82.
- [84] Konstan, J. a. et al. 1997. GroupLens: applying collaborative filtering to Usenet news. *Communications of the ACM*. 40, 3 (1997), 77–87.
- [85] Koren, Y. et al. 2009. Matrix Factorization Techniques for Recommender Systems. *Computer*. 42, 8 (2009), 42–49.
- [86] Koutrika, G. et al. 2008. Flexible recommendations over rich data. *Proceedings of the 2008 ACM Conference on Recommender Systems (RecSys 2008)* (Lausanne, Switzerland, 2008), 203–210.

-
- [87] Kumar, A.D. and Radhika, V. 2014. A Survey on Predicting Student Performance. *International Journal of Computer Science and Information Technologies*. 5, 5 (2014), 6147–6149.
- [88] Lang, K. 1995. NewsWeeder: Learning to Filter Netnews. *Proceedings of the 12th International Conference on Machine Learning* (Tahoe City, California, 1995), 331–339.
- [89] Leacock, T.L. and Nesbit, J.C. 2007. A Framework for Evaluating the Quality of Multimedia Learning Resources. *Educational Technology & Society*. 10, (2007), 44–59.
- [90] Lemire, D. and Boley, H. 2003. RACOFI: A Rule-Aplying Collaborative Filtering System. *Proceedings of the IEEE Conference WIC COLA* (Halifax, Canada, 2003), 13–24.
- [91] Lesovskis, A. et al. 2012. Integration of the Tagging Mechanism in the Collaborative e-Learning System. *Applied Computer Systems*. 13, (2012), 54–60.
- [92] Li, P. and Yamada, S. 2004. A movie recommender system based on inductive learning. *In Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems, 2004*. (Singapore, 2004), 1–3.
- [93] Lieberman, H. 1995. Letizia : An Agent That Assists Web Browsing. *Proceedings of the International Joint Conference on Artificial Intelligence*. (1995), 6.
- [94] Linden, G. et al. 2003. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*. 7, 1 (2003), 76–80.
- [95] Lis, B. and Korchmar, S. 2013. *Digitales Empfehlungsmarketing*. Gabler Verlag.
- [96] Littlestone, N. and Warmuth, M.K. 1994. The Weighted Majority Algorithm. *Information and Computation*. 108, 2 (1994), 212–261.
- [97] Lu, J. 2004. A Personalized e-Learning Material Recommender System. *Proceedings of the 2nd IEEE International Conference on Information Technology for Application (ICITA 2004)* (Harbin, China, 2004), 374–380.
- [98] Luo, G. and Pang, Y. 2010. Video annotation for enhancing blended learning of physical education. *In Proceedings of 2010 International Conference on Artificial Intelligence and Education (ICAIE)* (Hangzhou, China, 2010), 761–764.
- [99] Maâtallah, M. and Seridi-Bouchelaghem, H. 2012. Enhanced Collaborative Filtering to Recommender Systems of Technology Enhanced Learning. *In Proceedings of 4th International Conference on Web and Information Technologies ICWIT 2012* (Sidi Bel-Abbes, Algeria, 2012), 129–138.
- [100] Machardy, Z. and Pardos, Z.A. 2015. Evaluating The Relevance of Educational Videos using BKT and Big Data. *Proceedings of the 8th International Conference on Educational Data Mining* (Madrid, Spain, 2015), 424–427.
- [101] Manouselis, N. et al. 2013. *Recommender Systems for Learning*. Springer.
- [102] Marino, O. and Paquette, G. 2010. A competency-driven advisor system for multi-actor learning environments. *Proceedings of the 1st Workshop on Recommender Systems for Technology Enhanced Learning (RecSysTEL 2010)* (2010), 2871–2876.

-
- [103] Mazza, R. and Dimitrova, V. 2004. Visualising student tracking data to support instructors in web-based distance education. *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters (WWW Alt. 2004)*.
- [104] Metsis, V. et al. 2006. Spam filtering with naive bayes-which naive bayes? *Proceedings of the 3rd Conference on Email and Anti-Spam* (Mountain View, California, 2006).
- [105] Mihaescu, M.C. et al. 2015. Intelligent Tutor Recommender System for OnLine Educational Environments. *In Proceedings of the 8th International Conference on Educational Data Mining* (Madrid, Spain, 2015), 516–519.
- [106] Mohsin, S.F. 2010. Web based Multimedia Recommendation System for e-Learning Website. *Journal of Advanced Networking and Applications*. 223, (2010), 217–223.
- [107] Montaner, M. et al. 2003. A taxonomy of recommender agents on the internet. *Artificial Intelligence Review*. 19, 4 (2003), 285–330.
- [108] Mooney, R.J. and Roy, L. 2000. Content-Based Book Recommending Using Learning for Text Categorization. *In Proceedings of the 5th ACM Conference on Digital Libraries* (2000), 195–204.
- [109] Mostow, J. et al. 2005. An Educational Data Mining Tool to Browse Tutor-Student Interactions : Time Will Tell. *Proceedings of the Workshop on Educational Data Mining (2005)*. (2005), 15–22.
- [110] Neven, F. and Duval, E. 2002. Reusable Learning Objects : a Survey of LOM-Based Repositories. *Proceedings of the 10th ACM International Conference on Multimedia* (2002), 291–294.
- [111] Nicolajsen, P.H. and Plank, B. 2015. Using knowledge components for collaborative filtering in adaptive tutoring systems. *Proceedings of the 8th International Conference on Educational Data Mining* (Madrid, Spain, 2015), 536–539.
- [112] Niemann, K. 2015. Increasing the accessibility of learning objects by automatic tagging. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15*. (2015), 414–415.
- [113] Niemann, K. and Wolpers, M. 2013. Usage context-boosted filtering for recommender systems in TEL. *Lecture Notes in Computer Science*. 8095, (2013), 246–259.
- [114] Nurjanah, D. 2016. Good and Similar Learners' Recommendation in Adaptive Learning Systems. *In Proceedings of the 8th International Conference on Computer Supported Education (CSEDU 2016) - Volume 1* (2016), 434–440.
- [115] Oard, D.W. and Kim, J. 1998. Implicit Feedback for Recommender Systems. *Proceedings of the AAAI Workshop on Recommender Systems* (1998), 81–83.
- [116] Olcott, D. 2012. OER perspectives: emerging issues for universities. *Distance Education*. 33, 2 (2012), 283–290.
- [117] Pan, W. and Hawryszkiewicz, I. 2004. A method of defining learning processes. *Proceedings of the 21st ASCILITE Conference* (2004), 734–742.
- [118] Pazzani, M. and Billsus, D. 1997. Learning and Revising User Profiles: The Identification

- of Interesting Web Sites. *Machine Learning*. 27, 3 (1997), 313–331.
- [119] Pazzani, M.J. 1999. A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*. 13, 5 (1999), 393–408.
- [120] Peters, I. 2009. *Folksonomies: Indexing and Retrieval in Web 2.0*. De Gruyter Saur.
- [121] Purwitasari, D. et al. 2011. Ontology-based annotation recommender for learning material using contextual analysis. *Proceedings of the Conference on International Engineering and Technology Education (IETEC 2011)* (Kuala Lumpur, Malaysia, 2011).
- [122] Rafaeli, S. et al. 2004. QSIA - A web-based environment for learning, assessing and knowledge sharing in communities. *Computers and Education*. 43, 3 (2004), 273–289.
- [123] Rajaraman, A. and Ullman, J.D. 2011. *Recommendation Systems*. Cambridge University Press.
- [124] Rashid, A.M. et al. 2002. Getting to Know You: Learning New User Preferences in Recommender Systems. *Proceedings of the 7th International Intelligent User Interfaces Conference*. (2002), 127–134.
- [125] Rawashdeh, M. et al. 2013. Folksonomy link prediction based on a tripartite graph for tag recommendation. *Journal of Intelligent Information Systems*. 40, 2 (2013), 307–325.
- [126] Recker, M.M. et al. 2003. What do you recommend? Implementation and analyses of collaborative information filtering of web resources for education. *Instructional Science*. 31, 4–5 (2003), 299–316.
- [127] Resnick, P. et al. 1994. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. (1994), 175–186.
- [128] Resnick, P. and Varian, H.R. 1997. Recommender systems. *Communications of the ACM*. 40, 3 (1997), 56–58.
- [129] Ricci, F. et al. 2011. *Recommender Systems Handbook*. Springer.
- [130] Romero, C. and Ventura, S. 2010. Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*. 40, 6 (2010), 601–618.
- [131] Romero, C. and Ventura, S. 2007. Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*. 33, 1 (2007), 135–146.
- [132] Rust, I. and Krüger, M. 2011. Der Mehrwert von Vorlesungsaufzeichnungen als Ergänzungsangebot zur Präsenzlehre. *Wissensgemeinschaften: Digitale Medien- Öffnung und Offenheit in Forschung und Lehre*. (2011), 229–239.
- [133] Sack, H. and Waitelonis, J. 2007. OSOTIS - Kollaborative inhaltsbasierte Video-Suche. *DeLFI 2007: 5. E-Learning Fachtagung Informatik* (Bonn, 2007), 281–292.
- [134] Said, A. et al. 2009. A hybrid PLSA approach for warmer cold start in folksonomy recommendation. *In Proceedings of the RecSys 2009 Workshop on Recommender Systems & The Social Web* (2009), 87–90.

-
- [135] Sajjacholapunt, P. and Joy, M. 2016. SRECMATs - An Intelligent Tutoring System to Deliver Online Materials for Student Revision. *In Proceedings of the 8th International Conference on Computer Supported Education (CSEDU 2016) - Volume 1* (2016), 67–74.
- [136] Sarwar, B. et al. 2001. Item-based collaborative filtering recommendation algorithms. *Proceedings of the 10th international conference on World Wide Web* (2001), 285–295.
- [137] Sarwar, B.M. et al. 2000. Application of dimensionality reduction in recommender systems: a case study. *In ACM WebKDD Workshop*. 67, (2000), 12.
- [138] Schafer, J. Ben et al. 2007. Collaborative Filtering Recommender Systems. *The Adaptive Web*. P. Brusilovsky et al., eds. Springer. 291–324.
- [139] Schulmeister, R. 2013. Der Beginn und das Ende von OPEN. Chronologie der MOOC-Entwicklung. *In: MOOCs – Massive Open Online Courses. Offene Bildung oder Geschäftsmodell?*. R. Schulmeister, ed. Waxmann. 17–62.
- [140] Schulmeister, R. 2009. Studierende, Internet, E-Learning und Web 2.0. *E-Learning 2009. Lernen im digitalen Zeitalter*. N. Apostolopoulos et al., eds. Waxmann. 129–140.
- [141] Schulze, P. and Porath, D. 2012. *Statistik: mit Datenanalyse und ökonomischen Grundlagen*. Oldenbourg Wissenschaftsverlag.
- [142] Segaran, T. 2007. *Programming Collective Intelligence: Building Smart Web 2.0 Applications*. O’Reilly Media.
- [143] Shardanand, U. and Maes, P. 1995. Social information filtering: algorithms for automating “word of mouth.” *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1, (1995), 210–217.
- [144] Shelton, B.E. et al. 2010. Linking open course wares and open education resources: creating an effective search and recommendation system. *Procedia Computer Science*. 1, 2 (2010), 2865–2870.
- [145] Sheshadri, V. et al. 2014. InVis: An EDM Tool For Graphical Rendering And Analysis Of Student Interaction Data. *Proceedings of the Workshops held at Educational Data Mining 2014* (2014), 65–69.
- [146] Sicilia, M.-Á. et al. 2010. Exploring user-based recommender results in large learning object repositories: the case of MERLOT. *Procedia Computer Science*. 1, 2 (2010), 2859–2864.
- [147] Siemens, G. et al. 2011. Open Learning Analytics: an integrated & modularized platform. *Knowledge Creation Diffusion Utilization*.
- [148] Sinclair, J. and Cardew-Hall, M. 2007. The folksonomy tag cloud: when is it useful? *Journal of Information Science*. 34, 1 (2007), 15–29.
- [149] Singh, T. and Khanna, S. 2014. Reinforcement learning approach towards effective content recommendation in MOOC environments. *IEEE International Conference on MOOC, Innovation and Technology in Education (MITE), 2014* (2014), 285–289.
- [150] Smith, G. 2008. *Tagging: People-Powered Metadata for the Social Web*. New Riders.

-
- [151] Snelson, C. 2011. YouTube across the Disciplines : A Review of the Literature. *Journal of Online Learning and Teaching*. 7, 1 (2011), 159–169.
- [152] Soboroff, I. and Nicholas, C. 1999. Combining content and collaboration in text filtering. *International Joint Conferences on Artificial Intelligence* (1999), 86–91.
- [153] Su, X. and Khoshgoftaar, T.M. 2009. A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence*. Section 3 (2009), 1–19.
- [154] Szekely, B. and Torres, E. 2005. *Ranking Bookmarks and Bistros: Intelligent Community and Folksonomy Development*.
- [155] Taraghi, B. et al. 2013. Offene und partizipative Lernkonzepte. E-Portfolios, MOOCs und Flipped Classrooms. *Lehrbuch für Lernen und Lehren mit Technologien (L3T)*. M. Ebner and S. Sandra, eds.
- [156] Tatu, M. et al. 2008. RSDC’08: Tag Recommendations using Bookmark Content. *Proceedings of ECML PKDD Discovery Challenge* (2008), 96–107.
- [157] Terveen, L. et al. 1997. PHOAKS: a system for sharing recommendations. *Communications of the ACM*. 40, 3 (1997), 59–62.
- [158] Tibély, G. et al. 2012. Ontologies and tag-statistics. *New Journal of Physics*. 14, (2012), 1–22.
- [159] Tillmann, A. et al. 2012. Einsatz von E-Lectures als Ergänzungsangebot zur Präsenzlehre. *Digitale Medien – Werkzeuge für exzellente Forschung und Lehre*. G. Csanyi et al., eds. Waxmann. 235–249.
- [160] Tran, T. and Cohen, R. 2000. Hybrid Recommender Systems for Electronic Commerce. *AAAI Technical Report*. (2000), 78–84.
- [161] Trant, J. 2009. Studying social tagging and folksonomy: A review and framework. *Journal of Digital Information*. 10, 1 (2009), 1–44.
- [162] Tsai, K.H. et al. 2006. A Learning Objects Recommendation Model based on the Preference and Ontological Approaches. *Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies (ICALT 2006)* (2006), 36–40.
- [163] Tso-sutter, K. et al. 2008. Tag-aware recommender systems by fusion of collaborative filtering algorithms. In *Proceedings of the 2008 ACM Symposium on Applied Computing* (2008), 1995–1999.
- [164] Vargo, J. et al. 2003. Learning object evaluation: Computer mediated collaboration and inter-rater reliability. *International Journal of Computers and Applications*. 25, 3 (2003), 198–205.
- [165] Verbert, K. et al. 2012. Context-aware recommender systems for learning: A survey and future challenges. *IEEE Transactions on Learning Technologies*. 5, 4 (2012), 318–335.
- [166] Wachtler, J. et al. 2016. On Using Learning Analytics to Track the Activity of Interactive MOOC Videos. In *Proceedings of LAK 2016 Workshop on Smart Environments and Analytics in Video-Based Learning* (Edinburgh, United Kingdom, 2016), 8–17.

- [167] Wang, T.I. et al. 2007. Personalized Learning Objects Recommendation based on the Semantic- Aware Discovery and the Learner Preference Pattern. *Educational Technology & Society*. 10, 3 (2007), 84–105.
- [168] Wichelhaus, S. et al. 2008. More than Podcasting: An evaluation of an integrated blended learning scenario. *Ed-Media - World Conference on Educational Multimedia, Hypermedia & Telecommunications* (Vienna, Austria, 2008), 4468–4475.
- [169] Xu, R. and Wunsch, D.C. 2005. Survey of Clustering Algorithms. *IEEE Transactions on Neural Networks*. 16, 3 (2005), 645–678.
- [170] Yanbe, Y. et al. 2007. Can social bookmarking enhance search in the web? *Proceedings of the 2007 Conference on Digital libraries (JCDL 2007)* (2007), 107–116.
- [171] Yeung, C.A. and Iwata, T. 2011. Strength of Social Influence in Trust Networks in Product Review Sites. *Proceedings of the 4th ACM International Conference on Web Search and Data Mining* (2011), 495–504.
- [172] Yu, L. and Li, Q. 2009. Personal Media Data Organization and Retrieval in e-Learning : A Collaborative Tagging Based Approach. *Proceedings of the 1st ACM international workshop on Multimedia technologies for distance learning* (2009), 1–7.
- [173] Zhang, H. et al. 2007. Moodog: Tracking Students' Online Learning Activities. *In Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications (ED-Media 2007)* (2007), 4415–4422.

11 Anhang

Anhang A **Entwicklungsdokumentation**

Die Entwicklung des Systems Learning Assistance Osnabrück (LAOs) untergliedert sich in eine Tagging- sowie Serverkomponente. Die Taggingkomponente stellt das Userinterface des Systems dar, welches die Interaktion zwischen User und System ermöglicht. Die Serverkomponente implementiert sowohl das Datenmanagement als auch jegliche Funktionalität des Empfehlungsdienstes. Durch die getrennten Komponenten ergibt sich eine klare funktionale Trennung, wodurch sich gleichwohl ein struktureller Aufbau des Systems ergibt. Des Weiteren konnte so die flexible Auswahl bzw. der flexible Einsatz von Technologien auf Client- und Serverseite gewährleistet werden. Beide Komponenten sehen einen objektorientierten Ansatz vor, so dass zwischen den Komponenten ein Austausch von Objekten mittels REST (Representational State Transfer) im Austauschformat JSON (JavaScript Object Notation) erfolgen kann. Zur Kommunikation zwischen den Komponenten implementiert der Server einen Webservice, der eine Kommunikation mittels REST erlaubt. Nachfolgend ist sowohl die Umsetzung beider Komponenten als auch die Kommunikation zwischen den Komponenten beschrieben. Des Weiteren geben nachfolgende Abschnitte einen Überblick über die Technologien, die zur Umsetzung herangezogen wurden.

Taggingkomponente

Der LAOs Webclient bzw. die LAOs Taggingkomponente ist strukturell in nachfolgender Abbildung A.1 gezeigt.

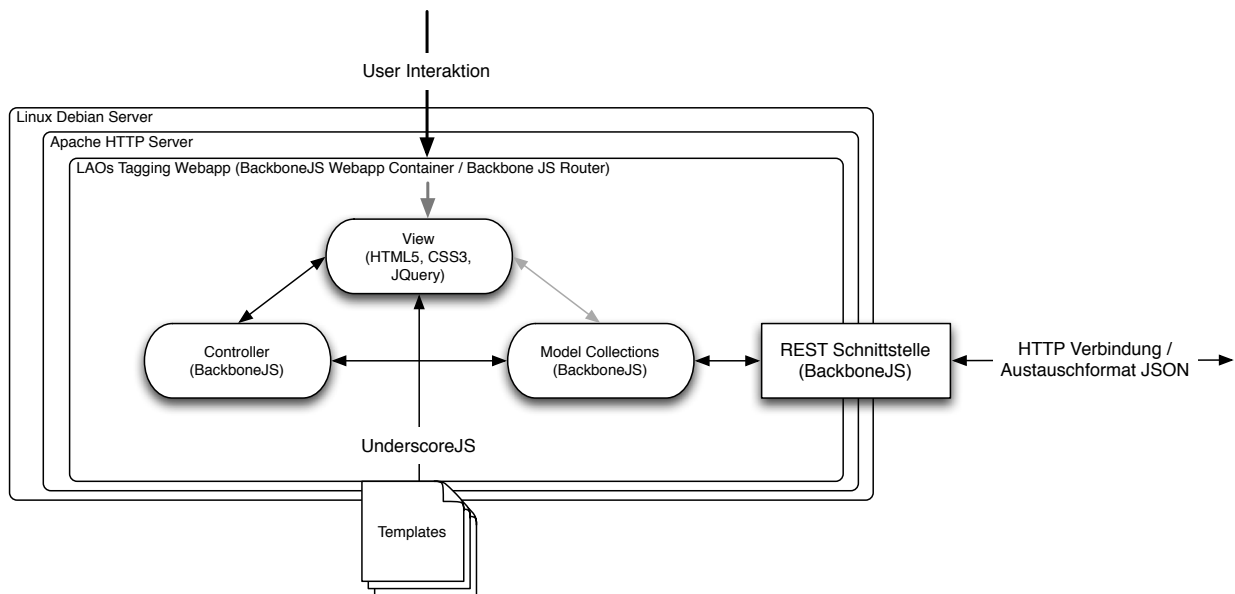


Abbildung A.1 Technische Übersicht Taggingkomponente

Eine erklärende Übersicht über alle verwendeten Technologien ist in nachfolgender Auflistung präsentiert.

- *Apache HTTP Server* (<https://httpd.apache.org/>): Webserver zur Auslieferung von Webseiten
- *HTML5*: Grundlegende Auszeichnungssprache zur Strukturierung von Webseiten und Templates.
- *CSS3*: Grundlegende Auszeichnungssprache zur grafischen Gestaltung von HTML Objekten
- *Bootstrap* (<http://getbootstrap.com/>): Das Bootstrap Framework ist ein Designframework auf Basis von HTML und CSS. Das Framework stellt Gestaltungsvorlagen z.B. für Typografie, Formulare, Buttons und weiteren Oberflächenelementen zur Verfügung. Die Nutzung des Bootstrap Grid-Layouts ermöglicht die Berücksichtigung des Responsive Webdesign Paradigma, um Webseiten ohne gesonderte Implementation auf einer Vielzahl an Endgeräten und Browsern Verfügbar zu machen.
- *Javascript*: Entwicklungssprache für die dynamische Interaktion und Auswertung auf Webseiten.
- *JQuery* (<https://jquery.com/>): JQuery ist eine Javascript Bibliothek zur vereinfachten Umsetzung von dynamischen Webseiten mit Javascript. Die Bibliothek übernimmt dabei

eine Vielzahl an Funktionalität wie DOM-Manipulation, Event Handling, Animationen und AJAX Kommunikation und achtet auf eine browserübergreifende Funktionalität.

- *BackboneJS* (<http://backbonejs.org/>): BackboneJS ist eine Javascript Bibliothek zur Umsetzung von sowohl Singlepage Anwendungen als auch der Umsetzung einer MVC Struktur in Webanwendungen. Durch BackboneJS ist es somit möglich, einer objektorientierten Entwicklung in Webanwendungen nachzugehen. Darüber hinaus implementiert BackboneJS eine REST Schnittstelle zur Daten- und Objektsynchronisation mit einem Server.
- *RequireJS* (<http://requirejs.org/>): Die Bibliothek RequireJS ermöglicht das asynchrone Hinzuladen von Javascript Modulen. Ein Modul kann in RequireJS unterschiedliche Arten von Internetobjekten so z.B. HTML-Seiten oder auch Javascript Bibliotheken enthalten. Der Vorteil von RequireJS ist das bedarfsgerechte Hinzuladen von Modulen, so dass beim Aufruf einer Webapplikation nicht alle, sondern nur benötigte Komponenten der Applikation geladen werden. Die Bibliothek dient somit der Leistungsoptimierung von Webanwendungen.
- *UnderscoreJS* (<http://underscorejs.org/>): UnderscoreJS ist eine Bibliothek die Hilfsfunktionen implementiert, die nativ nicht im Javascript Standard enthalten sind. So ermöglicht die Bibliothek das vereinfachte Arbeiten mit Objekten, Collections und Arrays durch die Umsetzung von z.B. Funktionen zur Veränderung oder Sortierprozeduren.
- *Highcharts* (<http://www.highcharts.com/>): Die Highcharts Bibliothek ist eine Javascript Bibliothek zur interaktiven Visualisierung von Daten (z.B. in Form von Diagrammen), die innerhalb der vorliegenden Arbeit zur Darstellung der aggregierten Useraktivitäten innerhalb der Analysekomponente für Lehrende genutzt wurde.

Die Taggingkomponente ist als eine Single-Page Webanwendung implementiert, die grundlegend nur aus einer Webseite besteht und deren Inhalte dynamisch nachgeladen werden. Die Implementation der Single-Page Anwendung erfolgt über die BackboneJS Bibliothek, die die Event-basierte Erzeugung und Synchronisierung von Webinhalten vornimmt. Die Bibliothek

sieht dabei die Umsetzung eines adaptierten Model-View-Controller⁸⁸ Patterns vor, was eine mühelose Änderbar- und Erweiterbarkeit der Anwendung ermöglicht. Eine Backbone Anwendung besteht grundlegend aus den *Elementen Collection, Model, View* und *Router*, die im weiteren Verlauf noch genauere erläutert werden. Zur Bereitstellung der Webanwendung findet ein Apache Webserver Verwendung. Um einen Einblick in die Struktur der Anwendung liefern zu können, ist nachfolgend die Dokumentenstruktur abgebildet. Die Darstellung berücksichtigt lediglich die wichtigsten Dateien, die z.B. zur Instanziierung der Anwendung genutzt werden, und belässt es für Unterverzeichnisse bei einer übergeordneten Beschreibung der enthaltenen Dateien.

- LAOsWebApp
 - application
 - collection
 - Definition von Backbone Collections (Javascript)
 - model
 - Definition von Backbone Models (Javascript)
 - View
 - Definition von Backbone Views, die ihrerseits aus HTML und Javascript Elementen bestehen.
 - html
 - HTML Gerüst der einzelnen Views
 - js
 - Javascript Funktionalität der einzelnen Views
 - template
 - Template Gerüste
 - laoswebapp.js: Instanziierung des Routers
 - router.js: Implementierung der Routerfunktionalität
 - lib
 - js
 - Javascript Bibliotheken
 - style
 - CSS Bibliotheken
 - index.html: Startseite
 - index.js: Initialisierung der Webapp

⁸⁸ Adaptiert deshalb, da die Controller Funktionalität sowohl im Backbone Router als auch in einer Backbone View implementiert ist.

Als Controller dient innerhalb der Anwendung der BackboneJS Router, der in Abhängigkeit eines aufgerufenen Routenpunktes eine zugehörige View rendert und ausliefert. Der Aufruf eines Routenpunktes erfolgt durch den Zusatz *#Routenpunkt* für die Hauptseite der Applikation innerhalb der URL (z.B. `http://laos.hs-osnabrueck.de/index.html#login`). Das Framework Backbone übernimmt innerhalb des Rendering Prozesses auch die Abfrage von Kontextabhängigen Informationen vom Server (z.B. Formular und Listeninformationen). Die Darstellung der Views ist mit HTML5, CSS3 und dem Framework Bootstrap umgesetzt. Die in der Taggingkomponente implementierten Views sind nachfolgend untergliedert für Lernende und Lehrende dargestellt.

LAOs Views Lernender:

- *Login*: Loginpage zur Anmeldung im System.

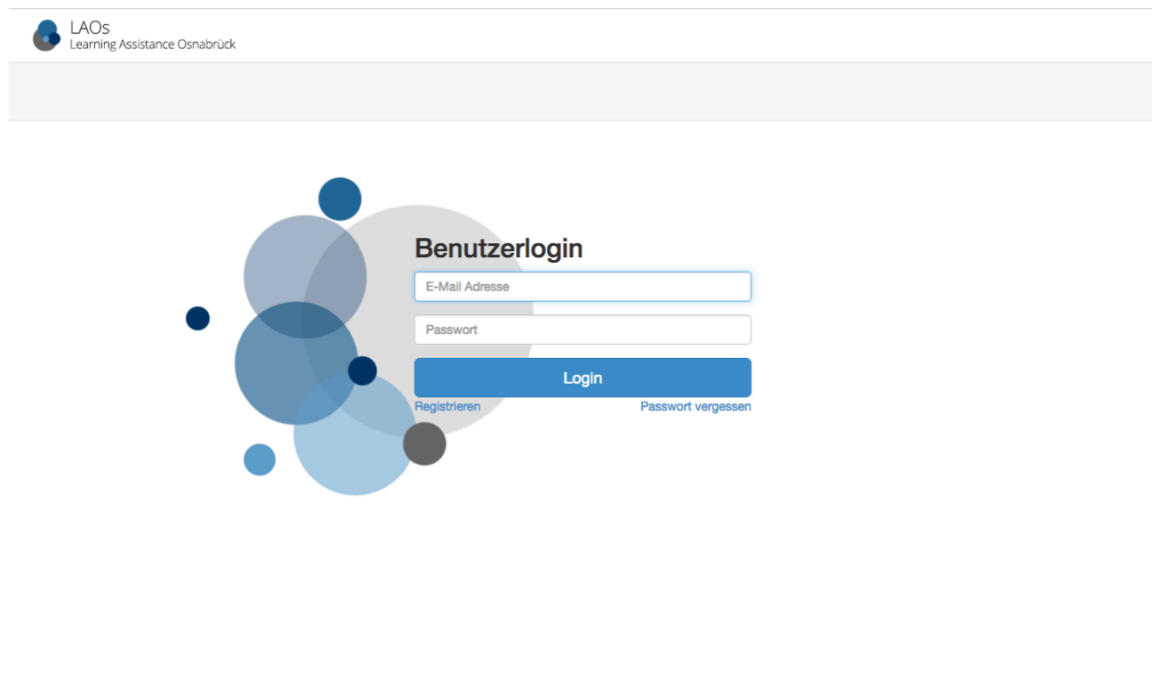


Abbildung A.2 Loginpage

- *Benutzerübersicht*: Übersicht über Benutzerdaten im System (im Prototyp lediglich Logout Funktionalität implementiert).

- *Kursgesamtübersicht*: Übersicht über alle im System vorhandenen Kurse.

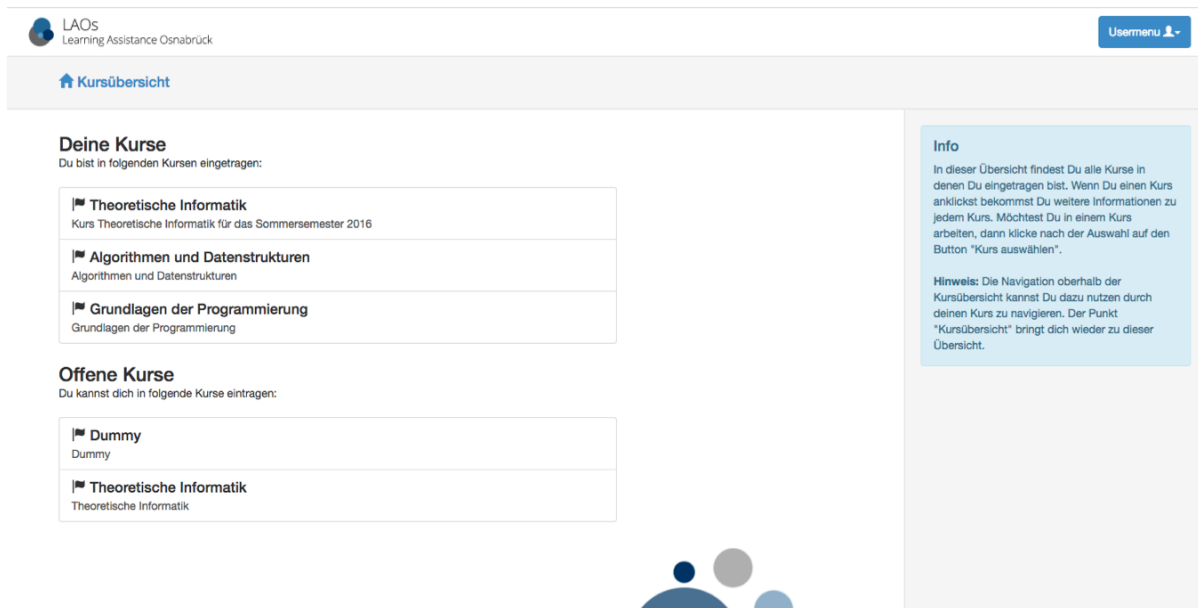


Abbildung A.3 Kursübersicht

- *Kursübersicht*: Übersicht eines einzelnen Kurses, über die ein Zugang zum Übungs- oder Materialpool des ausgewählten Kurses möglich ist.

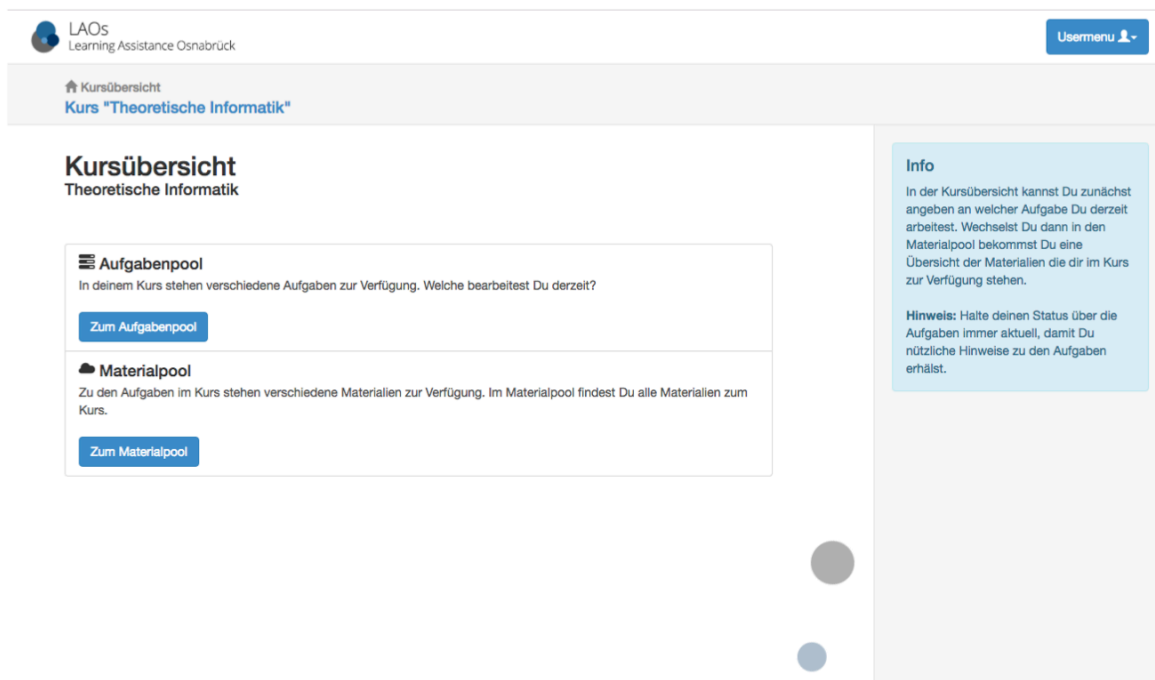


Abbildung A.4 Kursübersicht

- *Übungspool*: Übersicht über verfügbare Übungen innerhalb eines Kurses.

- *Materialpool*: Übersicht über verfügbare Lernmaterialien innerhalb eines Kurses.

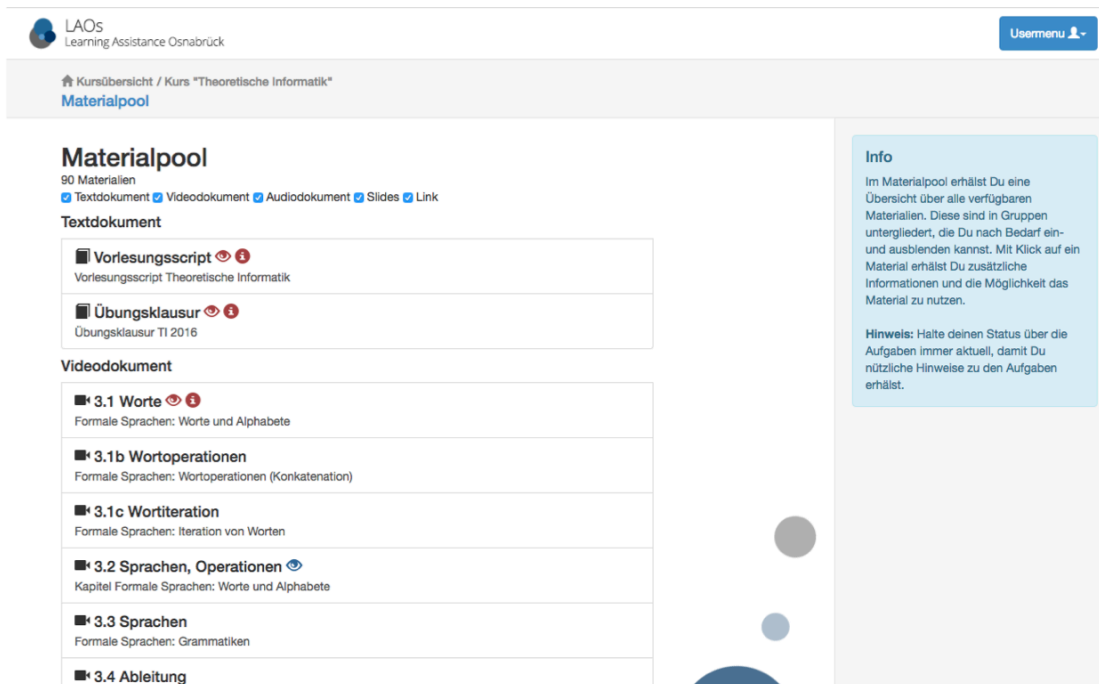


Abbildung A.5 Materialpool

- *(Text-, Video-) Arbeitsbereich*: Der Taggingbereich im System für ein ausgewähltes Lernmaterial (vgl. Abschnitt 4.3).

Nachfolgende Abbildungen zeigen die Taggingwerkzeuge innerhalb des Systems. Die Darstellung der Werkzeuge ist unter Verwendung von HTML5 und CSS3 umgesetzt. Animationen und (z.B. Click) Funktionalitäten werden mit der Javascript Bibliothek JQuery durchgeführt. Das Framework BackboneJS implementiert das Eventhandling der Applikation. Die Formulare der Werkzeuge sind als Templates vorgesehen, die dynamisch beim Öffnen des jeweiligen Werkzeuges mittels der Javascript Bibliothek UnderscoreJS geladen werden.

- *Kontextmenu*: Öffnet sich bei einem Rechtsklick auf das jeweilige Dokument



Abbildung A.6 Kontextmenu

- *Vordefinierter Kommentar*: Mehrere Kommentare können gleichzeitig ausgewählt werden.

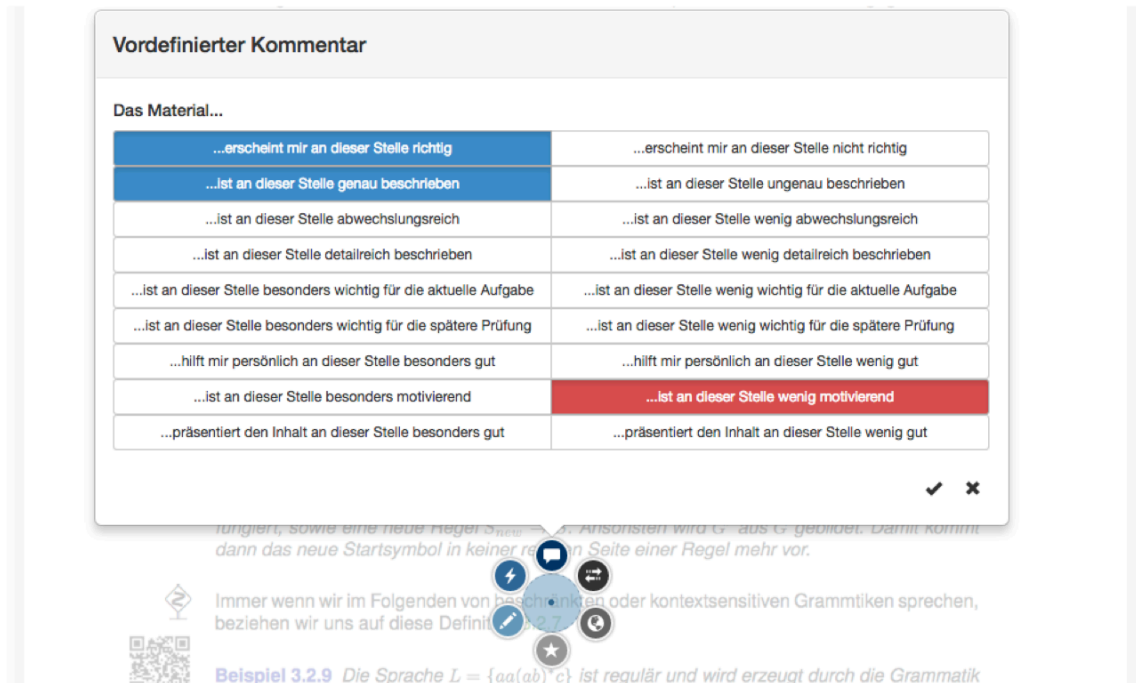


Abbildung A.7 Vordefinierter Kommentar

- *Kategorisierter Quicktage*: vgl. Abschnitt 4.3.
- *Marker*: Marker in Verbindung mit einem vordefinierten Kommentar.

Bemerkung 3.2.8 Jede Grammatik G , die nach der alten Definition 3.2.1 kontextsensitiv oder beschränkt war, kann man einfach zu einer kontextsensitiven bzw. beschränkten Grammatik G' nach neuer Definition machen. Dazu ist in G das Startsymbol auf der rechten Seite der Regeln zu entfernen. Man führt daher eine neue Variable S_{new} ein, die als neue Startvariable fungiert, sowie eine neue Regel $S_{\text{new}} \rightarrow S$. Ansonsten wird G' aus G gebildet. Damit kommt dann das neue Startsymbol in keiner rechten Seite einer Regel mehr vor.

Immer wenn wir im Folgenden von beschränkten oder kontextsensitiven Grammatiken sprechen, beziehen wir uns auf diese Definition 3.2.7.

Beispiel 3.2.9 Die Sprache $L = \{aa(ab)^*c\}$ ist regulär und wird erzeugt durch die Grammatik $G = (\{S, A\}, \{a, b, c\}, R, S)$ mit $R = \{S \rightarrow aaA, A \rightarrow abA \mid c\}$.

Markierung

Das Material...

...erscheint mir an dieser Stelle richtig	...erscheint mir an dieser Stelle nicht richtig
...ist an dieser Stelle genau beschrieben	...ist an dieser Stelle ungenau beschrieben
...ist an dieser Stelle abwechslungsreich	...ist an dieser Stelle wenig abwechslungsreich
...ist an dieser Stelle detailreich beschrieben	...ist an dieser Stelle wenig detailreich beschrieben
...ist an dieser Stelle besonders wichtig für die aktuelle Aufgabe	...ist an dieser Stelle wenig wichtig für die aktuelle Aufgabe
...ist an dieser Stelle besonders wichtig für die spätere Prüfung	...ist an dieser Stelle wenig wichtig für die spätere Prüfung
...hilft mir persönlich an dieser Stelle besonders gut	...hilft mir persönlich an dieser Stelle wenig gut
...ist an dieser Stelle besonders motivierend	...ist an dieser Stelle wenig motivierend
...präsentiert den Inhalt an dieser Stelle besonders gut	...präsentiert den Inhalt an dieser Stelle wenig gut

✓ ✕

Abbildung A.8 Marker

- **Rating:** Bewertungsskala von 1 bis 5.

Bewertung

Bewertung

1 2 3 4 5

✓ ✕

Abbildung A.9 Rating

- *Neues Material*: Verknüpfung zu einem neuen Lernmaterial aus dem Internet.

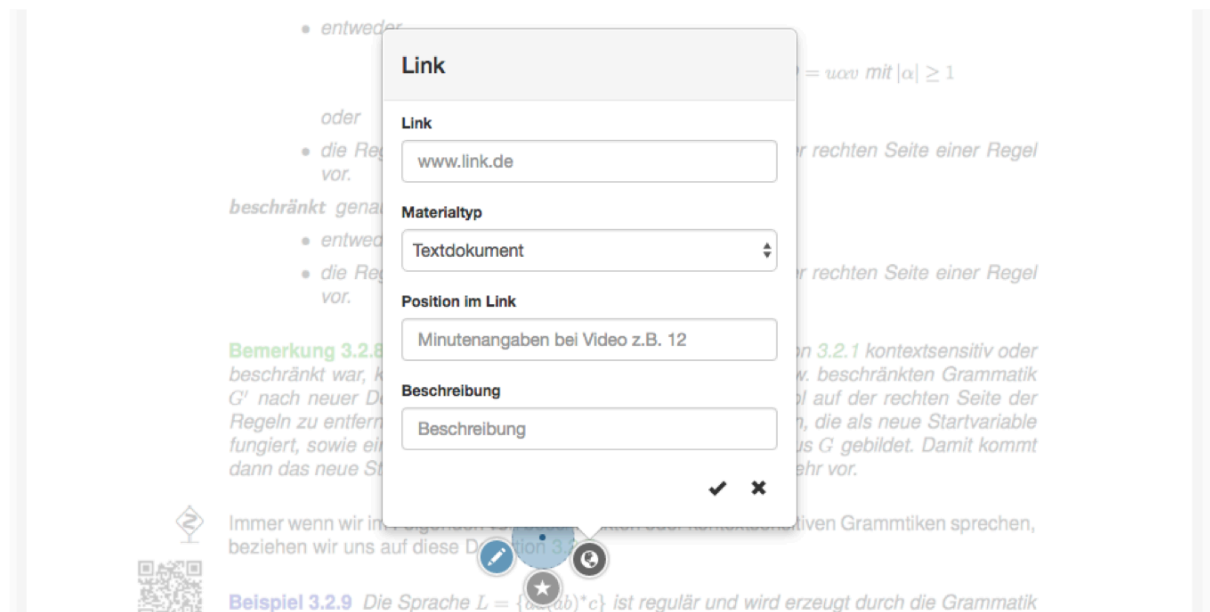


Abbildung A.10 Neues Material

- *Materialverbindung*: Eröffnung einer Materialverbindung (Abbildung A.11); Schließen einer Materialverbindung an einer neuen Position im gleichen oder neuen Material (Abbildung A.12).



Abbildung A.11 Materialverbindung Öffnen

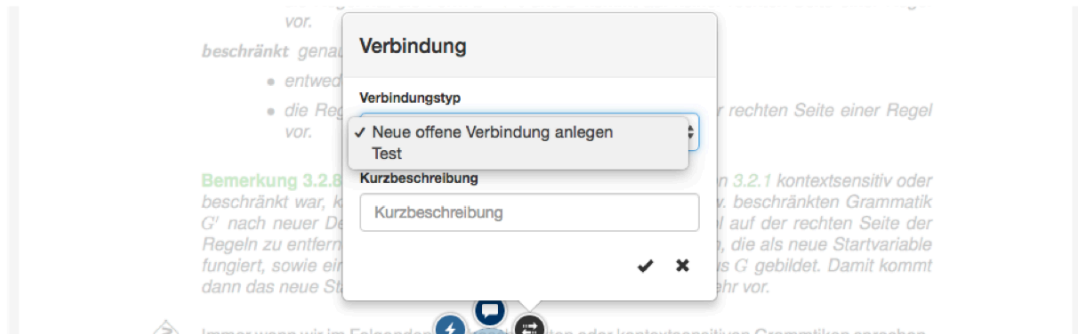


Abbildung A.12 Materialverbindung Schließen

- *Zeitliche Bereichsmarkierung*: Markierung eines zeitlichen Bereichs innerhalb von Audio-/Videomaterial.

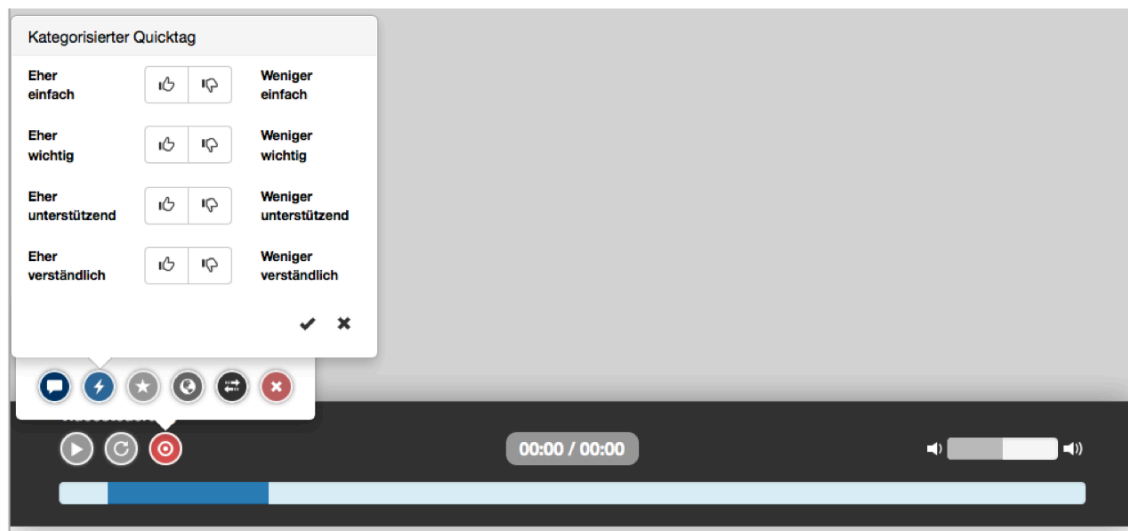


Abbildung A.13 Zeitliche Bereichsmarkierung (Audio/Video)

LAOs Views Lehrender:

- *Login*: Loginpage zur Anmeldung im System (gleich der Loginpage für Lernende).
- *Adminübersicht*: Übersichtsseite zum Anlegen, Löschen oder Auswählen von Kursen.

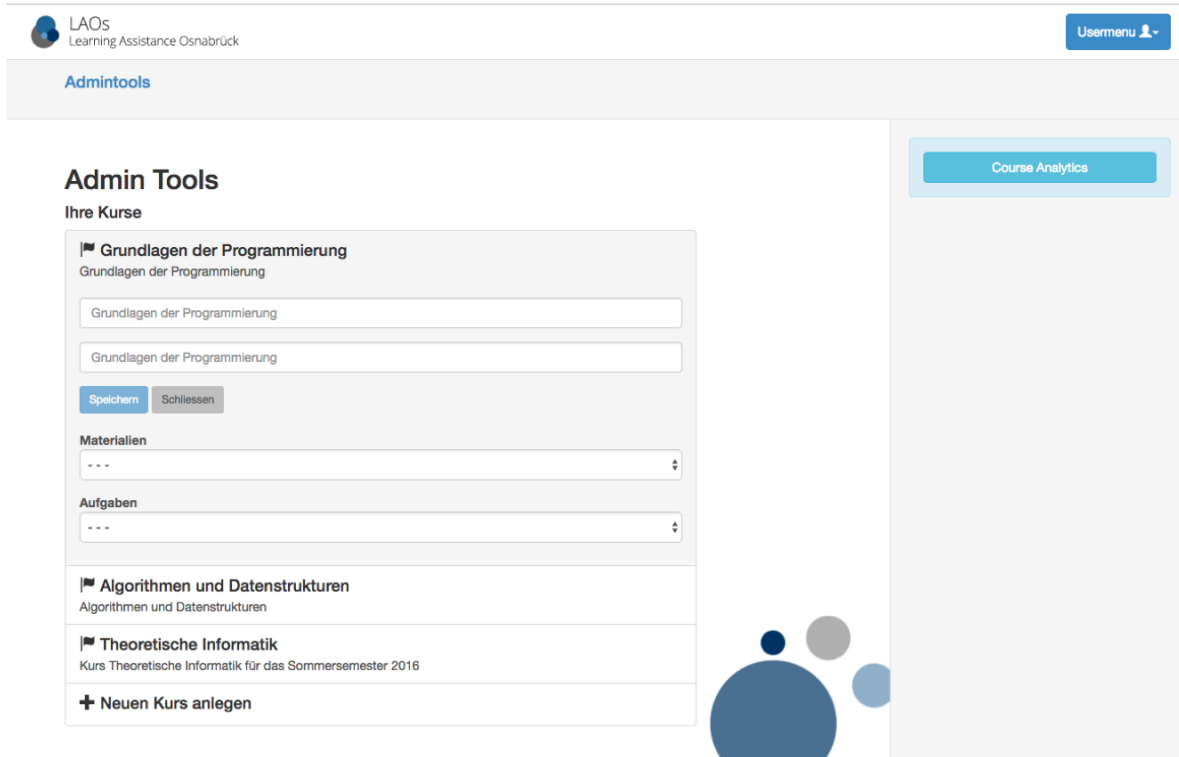


Abbildung A.14 Adminübersicht

- *Kursübersicht*: Übersichtsseite für einen ausgewählten Kurs mit Analysetools für Kursteilnehmer und Kursmaterialien (vgl. Abschnitt 4.4.4.2).
- *Materialanalyse*: Gleich dem Arbeitsbereich eines Lernenden, wobei die Taggingfunktionalität deaktiviert und die Übersicht über alle vorhandenen Daten (Tags, Empfehlungen) aktiviert ist.

Die View wird dabei in Abhängigkeit von Model Collections gerendert, die während des Rendering Prozesses vom Server angefragt werden. BackboneJS übernimmt in diesem Zusammenhang die vollständige Synchronisation mit dem Server mittels REST Schnittstelle. Eine Model Collection besteht dabei aus einer Vielzahl von Objekten, die durch ein Backbone Model beschrieben sind. Die Userinteraktion findet über die jeweilige BackboneJS View statt, über die Daten und Objekte durch den User manipuliert werden können. Auch in diesem

Zusammenhang nimmt Backbone eine direkte Synchronisation des Objektes bzw. Datenbestandes mit dem Server vor.

Die nachfolgende Abbildung A.15 stellt die in der Webapp verwendeten Models dar. Die Abbildung folgt aus Gründen der Übersicht einer deutlich vereinfachten Darstellung und verzichtet deshalb auf Objektattribute oder –methoden.

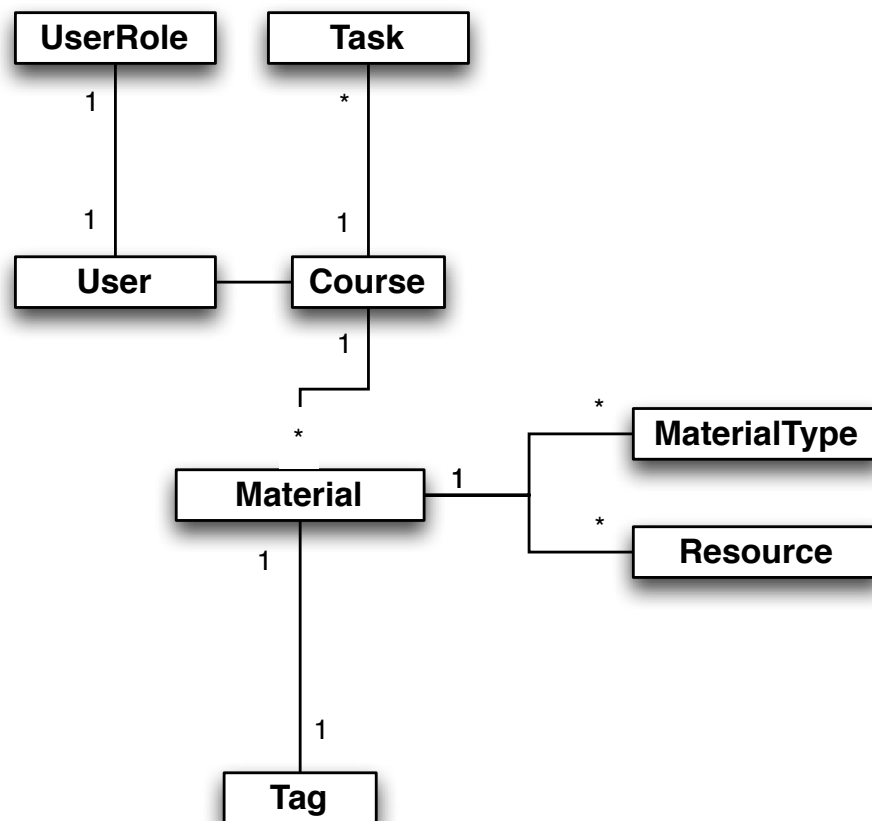


Abbildung A.15 Modelübersicht Taggingkomponente

Serverkomponente

Nachfolgende Abbildung zeigt den strukturellen Aufbau der LAOs Serverkomponente.

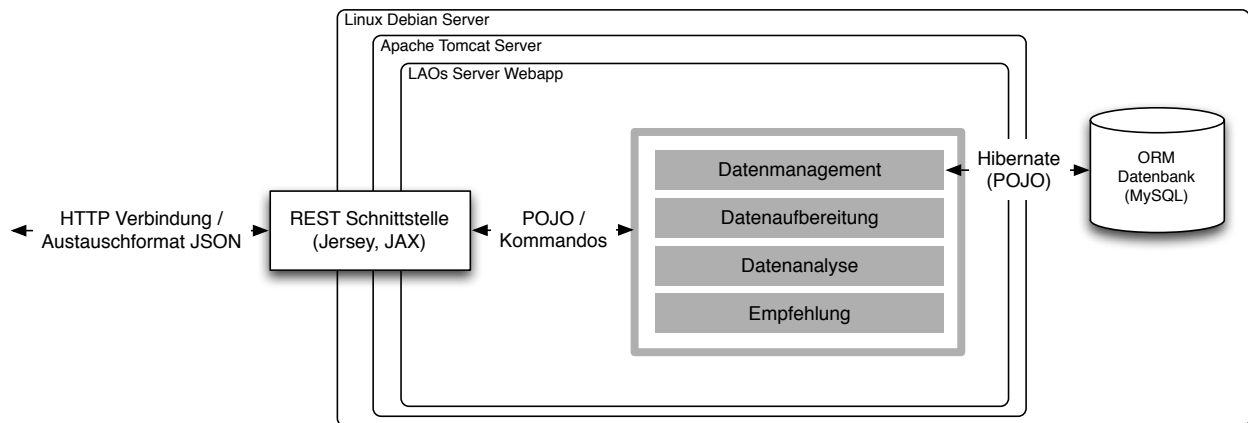


Abbildung A.16 Technische Übersicht Serverkomponente

Nachfolgende Technologien fanden bei der Entwicklung der Serverkomponente Verwendung.

- *Apache Tomcat* (<http://tomcat.apache.org/>): Der Apache Tomcat ist ein Webserver/-container zur Bereitstellung von Java-basierten Web-/Serveranwendungen.
- *Java*: Programmiersprache
- *Hibernate* (<http://hibernate.org/>): Hibernate ist ein Framework zur Persistierung von Daten oder Objekten (Umsetzung eines Objekt Relationalen Mappings ORM) für die Programmiersprache Java.
- *Jersey* (<https://jersey.java.net/>): Jersey ist ein Framework zur Umsetzung von REST Webservices auf Grundlage des JAX-RS Standards in Java. Die Implementation erfolgt im Programmcode über Java Annotationen.
- *JSON*: JSON steht kurz für JavaScript Object Notation und ist ein Austauschformat zwischen Client-/Serveranwendungen
- *MySQL* (<https://www.mysql.de/>): Relationales Datenbankverwaltungssystem
- *MySQL Connector* (<https://dev.mysql.com/downloads/connector/>): Schnittstelle zwischen Java Applikation und MySQL Datenbank

Die Serverkomponente des Systems wurde als Java-Servlet entwickelt, dessen Auslieferung durch einen Apache Tomcat Server übernommen wird. Die Entwicklung als Java-Servlet ermöglicht die Bereitstellung einer Java Anwendung als Webapplikation. Als Kommunikationsmöglichkeit mit der Taggingkomponente bietet der Server einen Webservice an, über den mittels REST Schnittstelle Objekte angefragt, erstellt, aktualisiert oder gelöscht werden können. Die Schnittstelle ist über das Java Framework Jersey implementiert. Der Vorteil einer REST Schnittstelle ist die Kommunikation über das HTTP Protokoll, sowie die Nutzung von HTTP Methoden wie GET, POST, PUT und DELETE, so dass eine Synchronisierung mittels BackboneJS auf Seiten der Taggingkomponente möglich ist. Die Zugriffssteuerung auf die REST Schnittstelle erfolgt in Abhängigkeit des Benutzernamen und Passworts des Anfragenden. Der Taggingclient eines Lernenden verfügt so z.B. lediglich über die Berechtigung nutzereigene Daten zu erfragen sowie Tags anzulegen bzw. nutzereigene Tags zu löschen. Nutzer mit administrativem Zugang erhalten hingegen die Möglichkeit zum Abfragen von vollständigen Kursinformationen (inkl. User, Materialien, Materialressourcen und Tags). Als Austauschformat zwischen Client und Server findet JSON Verwendung.

Neben dem Zugriff auf Daten ist der Server als Implementierungsumgebung des Empfehlungsdienstes umgesetzt. Der Empfehlungsdienstprozess kann ebenfalls über den Webservice des Servers mittels REST angestoßen werden.

Die Persistierung von Daten erfolgt innerhalb einer MySQL Datenbank. Um allerdings eine fortwährende Nutzung eines objektorientierten Ansatzes bzw. der Verwendung von Plain Old Java Objects (POJOs) zu ermöglichen, ist die direkte Persistierung von Objekten in Form eines Objektrelationalen Mappings (ORM) mittels des Java Frameworks Hibernate umgesetzt worden.

Nachfolgende Abbildung A.4 zeigt das Klassendiagramm der Serverkomponente. Auch in dieser Abbildung wurde bewusst auf die Darstellung von Attributen und Methoden verzichtet, um eine genügende Übersicht gewährleisten zu können.

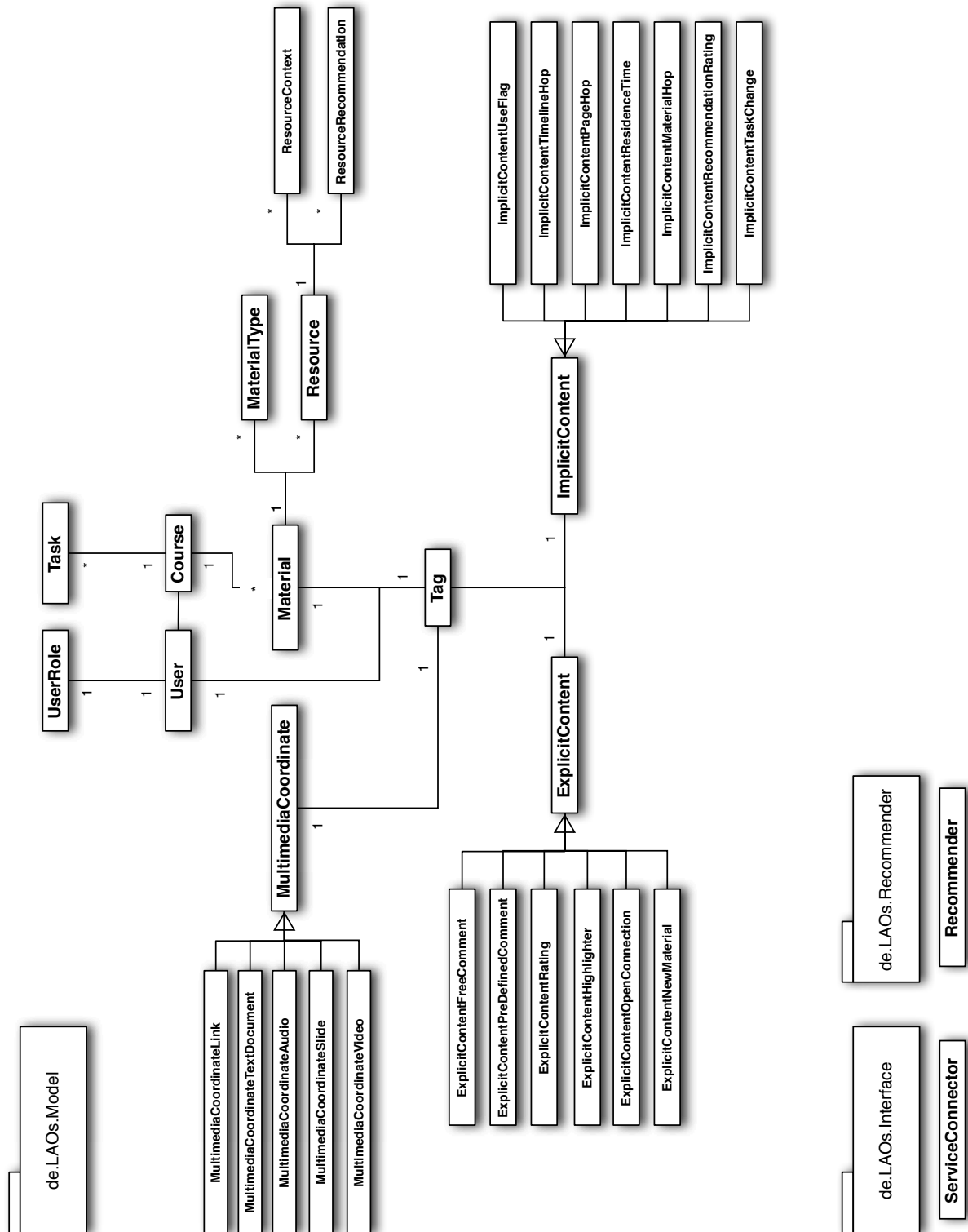


Abbildung A.17 Klassendiagramm Serverkomponente

Nachfolgend ist ein Überblick über die Methoden des severseitig implementierten Webservice und Recommender Moduls gezeigt.

LAOs Server Webservice Methoden:

de.LAOs.Model .Course	addCourse (java.lang.String userMail, java.lang.String userPassword, de.LAOs.Model.Course c) Add a Course to the Database /course/add/{userMail}/{userPassword}/
de.LAOs.Model .Material	addMaterial (java.lang.String userMail, java.lang.String userPassword, de.LAOs.Model.Material m) Add a Material to the Database URL: /material/{userMail}/{userPassword}/
de.LAOs.Model .Tag	addOpenConnectionReferenceTag (java.lang.String userMail, java.lang.String userPassword, de.LAOs.Model.Tag t) Add an OpenConnection Reference to the Database URL: /tag/open/connection/reference/{userMail}/{userPassword}/
de.LAOs.Model .Tag	addTag (java.lang.String userMail, java.lang.String userPassword, de.LAOs.Model.Tag t) Add a Tag to the database URL: /tag/{userMail}/{userPassword}/
java.util.List<de.LAOs.Model.Tag>	addTagList (java.lang.String userMail, java.lang.String userPassword, java.util.List<de.LAOs.Model.Tag> t) Add a List of Tags to the Database URL: /tag/list/{userMail}/{userPassword}/
de.LAOs.Model .Task	addTask (java.lang.String userMail, java.lang.String userPassword, int courseId, de.LAOs.Model.Task t) Add a Task to the Database
de.LAOs.Model .User	addUser (de.LAOs.Model.User u) Add a User to the Database /user/{userMail}/{userPassword}/
java.util.List<de.LAOs.Model.UserRole>	addUserRole (java.lang.String userMail, java.lang.String userPassword, de.LAOs.Model.UserRole ur) Add a UserRole (admin level only) /userrole/add/{userMail}/{userPassword}/
de.LAOs.Model .Course	addUserToCourse (java.lang.String userMail, java.lang.String userPassword, int courseId, de.LAOs.Model.Course c) Add a User to Course relation /user/course/{id}/{userMail}/{userPassword}/
java.util.List<de.LAOs.Model.Resource>	calculateResourcesByCoordinateId (int materialId, java.lang.String userMail, java.lang.String userPassword, int strictFlag)
java.lang.String	deleteMaterial (java.lang.String userMail, java.lang.String userPassword, int id) Delete a certain Material (admin level only) URL: /material/{id}/{userMail}/{userPassword}/
java.lang.String	deleteTag (java.lang.String userMail, java.lang.String userPassword, int id) Delete a certain Tag URL: /tag/{id}/{userMail}/{userPassword}/

java.lang.String	deleteUser (java.lang.String userMail, java.lang.String userPassword, int id) Delete a certain User (admin level only) /user/{id}/{userMail}/{userPassword}/
java.lang.String	deleteUserRole (java.lang.String userMail, java.lang.String userPassword, int id) Delete a certain UserRole (admin level only) /userrole/{id}/{userMail}/{userPassword}/
de.LAOs.Model.Course	deleteUserToCourse (java.lang.String userMail, java.lang.String userPassword, int courseId, de.LAOs.Model.Course c) Add a User to Course relation /user/course/{id}/{userMail}/{userPassword}/
de.LAOs.Model.Course	getCourse (java.lang.String userMail, java.lang.String userPassword, int id) Get a certain Course by ID /course/{id}/{userMail}/{userPassword}/
de.LAOs.Model.Course	getCourseFromMaterial (java.lang.String userMail, java.lang.String userPassword, int id) Get a certain Course from Material ID /course/from/material/{materialId}/{userMail}/{userPassword}/
java.util.List<de.LAOs.Model.Course>	getCourseList ()
java.util.List<de.LAOs.Model.Course>	getCourseList (java.lang.String userMail, java.lang.String userPassword) Get a List of Courses /course/list/{userMail}/{userPassword}/
java.util.List<de.LAOs.Model.Tag>	getCourseMaterialConnections (java.lang.String userMail, java.lang.String userPassword, int materialId)
java.util.List<de.LAOs.Model.Tag>	getCourseMaterialLinks (java.lang.String userMail, java.lang.String userPassword, int materialId)
java.util.List<de.LAOs.Tools.AnalyticsWrapperObject>	getCourseMaterialRelevance (java.lang.String userMail, java.lang.String userPassword, int materialId, int useRealTime)
java.util.List<de.LAOs.Tools.AnalyticsWrapperObject>	getCourseMaterialTagAppearance (java.lang.String userMail, java.lang.String userPassword, int materialId, int useRealTime)
java.util.List<de.LAOs.Tools.AnalyticsWrapperObject>	getCourseMaterialUseByUser (java.lang.String userMail, java.lang.String userPassword, int userId)
java.util.List<de.LAOs.Tools.AnalyticsWrapperObject>	getCourseMaterialUser (java.lang.String userMail, java.lang.String userPassword, int courseId)
java.util.List<de.LAOs.Model.Course>	getCourseOpenList (java.lang.String userMail, java.lang.String userPassword) Get a List of Courses /course/list/{userMail}/{userPassword}/
java.util.List<de.LAOs.Tools.AnalyticsWrapperObject>	getCourseStatistics (java.lang.String userMail, java.lang.String userPassword, int courseId)

ppperObject>	
java.util.List <de.LAOs.Tools.AnalyticsWrapperObject>	getCourseTagOverview (java.lang.String userMail, java.lang.String userPassword, int courseId)
de.LAOs.Model.Tag	getCurrentRecommendationRating (java.lang.String userMail, java.lang.String userPassword, int resourceId) Get a users rating for a resource/recommendation
java.util.List <de.LAOs.Model.Tag>	getExplicitTagByMaterial (int materialId, java.lang.String userMail, java.lang.String userPassword) Get a List of explicit Tags for a certain Material AND User URL: /tag/list/explicit/{materialId}/{userMail}/{userPassword}/
java.util.List <de.LAOs.Model.Tag>	getFullTagByMaterial (int materialId, java.lang.String userMail, java.lang.String userPassword) Get a List of Tags for certain Material Admin Level URL: /tag/list/{materialId}/{userMail}/{userPassword}/
de.LAOs.Model.Material	getMaterialById (java.lang.String userMail, java.lang.String userPassword, int id) Get a certain Material by ID URL: /material/{id}/{userMail}/{userPassword}/
java.util.List <de.LAOs.Model.Material>	getMaterialList (java.lang.String userMail, java.lang.String userPassword, int courseId) Get List of a Material for a certain Course URL: /material/list/{id}/{userMail}/{userPassword}/
java.util.List <de.LAOs.Model.MaterialType>	getMaterialTypeList (java.lang.String userMail, java.lang.String userPassword) Get List of Material Types URL: /materialtype/list/{userMail}/{userPassword}/
java.util.List <de.LAOs.Model.Tag>	getOpenConnectionTagByMaterial (int materialId, java.lang.String userMail, java.lang.String userPassword) Get a List of OpenConnection Tags for a certain Material AND User URL: /tag/list/open/connection/{materialId}/{userMail}/{userPassword}/
java.util.List <de.LAOs.Model.Course>	getOwnerCourseList (java.lang.String userMail, java.lang.String userPassword) Get a List of Courses where a certain User is Owner /course/owner/list/{userMail}/{userPassword}/
java.util.List <de.LAOs.Model.Tag>	getPageTagByMaterial (int materialId, int pageId, java.lang.String userMail, java.lang.String userPassword) Get a List of Tags for certain Material and a certain MultimediaCoordinate Admin Level URL: /tag/list/{materialId}/{pageId}/{userMail}/{userPassword}/
java.util.List <de.LAOs.Model.Resource>	getResourcesByCoordinateId (int materialId, int coordinateId, java.lang.String userMail, java.lang.String userPassword)
java.lang.String	getStatus () Test Method if the Server is running
de.LAOs.Model.Tag	getTagById (java.lang.String userMail, java.lang.String userPassword, int id) Get a certain Tag Object by Id URL: /tag/{id}/{userMail}/{userPassword}/

java.util.List t<de.LAOs.Model 1.Tag>	getTagByMaterial (int materialId, java.lang.String userMail, java.lang.String userPassword) Get a List of Tags for certain Material AND a certain User URL: /tag/list/{materialId}/{userMail}/{userPassword}/
de.LAOs.Model .Tag	getTagLastTask (java.lang.String userMail, java.lang.String userPassword) Get the last Task as Tag Type for a certain User URL: /tag/last/task/{userMail}/{userPassword}/
java.util.List t<de.LAOs.Model 1.Tag>	getTagList (java.lang.String userMail, java.lang.String userPassword) Get a List of Tags for a certain User; Admin Level Only; URL: /tag/list/{userMail}/{userPassword}/
java.util.List t<de.LAOs.Model 1.Tag>	getTagListByUser (java.lang.String userMail, java.lang.String userPassword) Get a List of Tags for a certain User; User Level; URL: /tag/list/user/{userMail}/{userPassword}/
java.util.List t<de.LAOs.Model 1.Tag>	getTagListTask (java.lang.String userMail, java.lang.String userPassword, int courseId) Get a List of Tasks for a certain User URL: /tag/list/task/{userMail}/{userPassword}/
java.util.List t<de.LAOs.Model 1.Tag>	getTagRecommendation (java.lang.String userMail, java.lang.String userPassword, int materialId) Get Recommendation Tags for a certain user and a certain material URL: /tag/last/task/{userMail}/{userPassword}/
de.LAOs.Model .Task	getTask (java.lang.String userMail, java.lang.String userPassword, int id) Get a certain Task by ID /task/{id}/{userMail}/{userPassword}/
java.util.List t<de.LAOs.Model 1.Task>	getTaskList (java.lang.String userMail, java.lang.String userPassword, int id) Get a List of Tasks for a certain Course /task/list/{id}/{userMail}/{userPassword}/
de.LAOs.Model .User	getUserById (java.lang.String userMail, java.lang.String userPassword, int id) Get a certain User by ID URL: /user/{id}/{userMail}/{userPassword}/
de.LAOs.Model .User	getUserByLogin (java.lang.String userMail, java.lang.String userPassword) Get a certain User by Login Data /user/{userMail}/{userPassword}/
java.util.List t<de.LAOs.Model 1.User>	getUserList (java.lang.String userMail, java.lang.String userPassword) Get a List of all Users (admin level only) /user/list/{userMail}/{userPassword}/
java.util.List t<java.lang.Object>	getUserListByCourse (int courseId, java.lang.String userMail, java.lang.String userPassword) Get a List of all Course Users (admin level only) /user/list/{userMail}/{userPassword}/
de.LAOs.Model .UserRole	getUserRole (java.lang.String userMail, java.lang.String userPassword, int id) Get a certain UserRole by ID (admin level only) /userrole/{id}/{userMail}/{userPassword}/
java.util.List t<de.LAOs.Model 1.UserRole>	getUserRoleList (java.lang.String userMail, java.lang.String userPassword) Get a List of User Roles (admin level only) /userrole/{userMail}/{userPassword}/
de.LAOs.Model .Course	updateCourse (java.lang.String userMail, java.lang.String userPassword, de.LAOs.Model.Course c) Update Course Information /course/update/{userMail}/{userPassword}/
de.LAOs.Model .Tag	updateCurrentRecommendationRating (java.lang.String userMail, java.lang.String userPassword, de.LAOs.Model.Tag t) Add or Update Recommendation / Resource Rating

de.LAOs.Model .Material	updateMaterial (java.lang.String userMail, java.lang.String userPassword, de.LAOs.Model.Material m) Update Material Information /material/update/{userMail}/{userPassword}/
de.LAOs.Model .Tag	updateTag (java.lang.String userMail, java.lang.String userPassword, de.LAOs.Model.Tag t) Update a Tag URL: /tag/update/{userMail}/{userPassword}/
de.LAOs.Model .Task	updateTask (java.lang.String userMail, java.lang.String userPassword, de.LAOs.Model.Task t) Update Task Information

LAOs Server Recommender Methoden:

static boolean []	calculatePageRelevance (java.util.ArrayList<de.LAOs.Tools.AnalyticsWrapperStringObject> data, boolean strictFlag) Calculate different types of thresholds for page relevance in Text documents For the use to discard the calculation of resources in non relevant pages Use strictFlag == TRUE to incorporate Tagappearance AND Used Time Use strictFlag == FALSE to ensure Tagappearance OR Used Time needs to respected
static java.util.List<de.LAOs.Model.ResourceRecommendation>	calculateResourceAdditionalResources (de.LAOs.Model.Resource r, java.util.List<de.LAOs.Model.Tag> eco)
static java.util.List<de.LAOs.Model.ResourceContext>	calculateResourceContext (de.LAOs.Model.Resource r)
static java.util.List<de.LAOs.Tools.AnalyticsWrapperDoubleObject>	calculateResourceCategoryScore (de.LAOs.Model.Resource r, java.util.List<de.LAOs.Configuration.ExplicitContentConfiguration> ecc) Calculate Resource Score for the fields of Difficulty, Importance, Assistance, Understanding
static double	calculateResourceScore (de.LAOs.Model.Resource r) Calculate a general Resource Score including Difficulty Score, Importance Score, Assistance Score, Understanding Score
static double	calculateTagCollectionScore (java.util.List<de.LAOs.Model.Tag> tags)
static boolean []	calculateTimeRelevance (java.util.ArrayList<de.LAOs.Tools.AnalyticsWrapperStringObject> data, boolean strictFlag) Calculate different types of thresholds for time relevance in Video documents For the use to discard the calculation of resources in non relevant pages Use strictFlag == TRUE to incorporate Tagappearance AND Used Time Use strictFlag == FALSE to ensure Tagappearance OR Used Time needs to respected
static void	calculateTimeRelevance (java.util.ArrayList<de.LAOs.Model.Tag> tags)
static double	defineTagScore (java.lang.String configurationCategory, java.lang.String configurationContent, java.lang.String configurationContentCategory,

	<pre>java.util.List<de.LAOs.Configuration.ExplicitContentConfiguration> conf)</pre> <p>Calculate a Tag Score on the basis of a given ExplicitContentConfiguration and Tag values</p>
<pre>static java.util.ArrayList<de.LAOs.Model.Resource></pre>	<pre>findResourcesText(java.util.List<de.LAOs.Model.Tag> tags, boolean[] relevanceArray, boolean strictFlag)</pre> <p>Find Resources in a Text Document on the basis of a Tag Collection</p>
<pre>static java.util.ArrayList<de.LAOs.Model.Resource></pre>	<pre>findResourcesVideo(java.util.List<de.LAOs.Model.Tag> tags, boolean[] relevanceArray, boolean strictFlag)</pre> <p>Find Resources in a Video Document on the basis of a Tag Collection</p>
<pre>static java.util.ArrayList<de.LAOs.Model.Tag></pre>	<pre>getCategoryTagCollection(java.util.List<de.LAOs.Model.Tag> tags, java.util.List<de.LAOs.Configuration.ExplicitContentConfiguration> conf, java.lang.String category)</pre> <p>Get Tag Collection for a specific category type (importance, assistance, understanding, difficulty)</p>
<pre>static java.lang.String</pre>	<pre>getCurrentTagCategory(java.lang.String tagContent, java.util.List<de.LAOs.Configuration.ExplicitContentConfiguration> conf)</pre>
<pre>static java.util.ArrayList<de.LAOs.Model.Tag></pre>	<pre>setTagScore(java.util.List<de.LAOs.Model.Tag> tags, java.util.List<de.LAOs.Configuration.ExplicitContentConfiguration> conf, java.util.List<de.LAOs.Configuration.ContentRuleChain> rules)</pre> <p>Set Tag Scores for a given Tag Collection</p>

Anhang B Detailübersicht Tag Konfiguration

Legende:		Value	Vordefinierter Tagscore			RA	Ausprägung A für Handlungsempfehlung				
		AA	Ausprägung A für Systemaussage			RB	Ausprägung B für Handlungsempfehlung				
		AB	Ausprägung B für Systemaussage			RC	Ausprägung C für Handlungsempfehlung				
		%	Platzhalter für Ausprägung			RD	Ausprägung D für Handlungsempfehlung				
Type	Aussage	Value	Kategorie	Aussage System	Empfehlung System	AA	AB	RA	RB	RC	RD
Rating	5 Sterne	0,5									
Rating	4 Sterne	0,3									
Rating	3 Sterne	0,1									
Rating	2 Sterne	-0,3									
Rating	1 Stern	-0,5									
Quicktag	Eher Wichtig	0.7	W	% relevant	% wichtig ist	sehr		sehr		ggf.	ggf.
Quicktag	Eher Unwichtig	-0.7	W	% irrelevant	an einer anderen Stelle bzw. als Ergänzung % wichtiger	sehr	eher	deutlich	eher	ggf.	ggf.
Quicktag	Eher Schwieriger	0.7	S	% schwierig	% schwierig ist	gar nicht	wenig	gar nicht	wenig	ggf.	ggf.
Quicktag	Weniger Schwierig	-0.7	S	% komplex	investiere % Zeit zum erarbeiten	sehr	eher	sehr viel	viel	ggf.	ggf.
Quicktag	Eher Verständlich	0.7	V	% eingängig	inhaltlich % gut aufbereitet ist	sehr		sehr		ggf.	ggf.
Quicktag	Eher Unverständlich	-0.7	V	% unverständlich	betrachte % aufmerksam	sehr	eher	unbedingt	besser	ggf.	ggf.
Quicktag	Eher Unterstützend	0.7	U	% hilfreich	% zum lernen anregt	sehr		sehr		ggf.	ggf.
Quicktag	Weniger Unterstützend	-0.7	U	% ineffektiv	ziehe % alternative Materialien zum Lernen heran	sehr	eher	unbedingt	besser	ggf.	ggf.
Freitext	Das Material erscheint mir an dieser Stelle richtig	0.9	V	% fehlerfrei	inhaltlich % gut aufbereitet ist	vollständig		sehr		ggf.	ggf.
Freitext	Das Material erscheint mir an dieser Stelle nicht richtig	-0.9	V	% missverständlich	betrachte % aufmerksam	sehr	eher	unbedingt	besser	ggf.	ggf.
Freitext	Das Material ist an dieser Stelle genau beschrieben	0.8	V	% präzise	inhaltlich % gut aufbereitet ist	sehr		sehr		ggf.	ggf.
Freitext	Das Material ist an dieser Stelle ungenau beschrieben	-0.8	V	% unpräzise	betrachte % aufmerksam	sehr	eher	unbedingt	besser	ggf.	ggf.
Freitext	Das Material ist an dieser Stelle abwechslungsreich	0.7	U	% vielschichtig	% zum lernen anregt	sehr		sehr		ggf.	ggf.
Freitext	Das Material ist an dieser Stelle wenig abwechslungsreich	-0.7	U	% monoton	ziehe % alternative Materialien zum Lernen heran	sehr	eher	unbedingt	besser	ggf.	ggf.
Freitext	Das Material ist an dieser Stelle detailreich beschrieben	0.8	S	% ausführlich	% schwierig ist	sehr		gar nicht	wenig	ggf.	ggf.
Freitext	Das Material ist an dieser Stelle wenig detailreich beschrieben	-0.8	S	% oberflächlich	investiere % Zeit zum erarbeiten	sehr	eher	sehr viel	viel	ggf.	ggf.
Freitext	Das Material ist an dieser Stelle besonders wichtig für die aktuelle Aufgabe	0.9	W	% relevant	% wichtig ist	sehr		sehr		ggf.	ggf.
Freitext	Das Material ist an dieser Stelle wenig wichtig für die aktuelle Aufgabe	-0.9	W	% irrelevant	an einer anderen Stelle bzw. als Ergänzung % wichtiger	sehr	eher	deutlich	eher	ggf.	ggf.
Freitext	Das Material ist an dieser Stelle besonders wichtig für die spätere Prüfung	0.9	W	% relevant	% wichtig ist	sehr		sehr		ggf.	ggf.
Freitext	Das Material ist an dieser Stelle wenig wichtig für die spätere Prüfung	-0.9	W	% irrelevant	an einer anderen Stelle bzw. als Ergänzung % wichtiger	sehr	eher	deutlich	eher	ggf.	ggf.
Freitext	Das Material hilft mir persönlich an dieser Stelle	0.9	U	% nützlich	% zum lernen anregt	sehr		sehr		ggf.	ggf.

	besonders gut										
Freitext	Das Material hilft mir persönlich an dieser Stelle wenig gut	-0.9	U	% unzweckmäßig	ziehe % alternative Materialien zum Lernen heran	sehr	eher	unbedingt	besser	ggf.	ggf.
Freitext	Das Material ist an dieser Stelle besonders motivierend	0.8	U	% motivierend	% zum lernen anregt	sehr		sehr		ggf.	ggf.
Freitext	Das Material ist an dieser Stelle wenig motivierend	-0.8	U	% demotivierend	ziehe % alternative Materialien zum Lernen heran	sehr	eher	unbedingt	besser	ggf.	ggf.
Freitext	Die Präsentation des Materials hilft mir an dieser Stelle Inhalte zu erarbeiten	0.9	U	% anschaulich	% zum lernen anregt	sehr		sehr		ggf.	ggf.
Freitext	Die Präsentation des Materials hilft mir an dieser Stelle wenig Inhalte zu erarbeiten	-0.9	U	% abstrakt	ziehe % alternative Materialien zum Lernen heran	sehr	eher	unbedingt	besser	ggf.	ggf.
Marker	Das Material erscheint mir an dieser Stelle richtig	0.9	V	% fehlerfrei	inhaltlich % gut aufbereitet ist	vollständig		sehr		ggf.	ggf.
Marker	Das Material erscheint mir an dieser Stelle nicht richtig	-0.9	V	% missverständlich	betrachte % aufmerksam	sehr	eher	unbedingt	besser	ggf.	ggf.
Marker	Das Material ist an dieser Stelle genau beschrieben	0.8	V	% präzise	inhaltlich % gut aufbereitet ist	sehr		sehr		ggf.	ggf.
Marker	Das Material ist an dieser Stelle ungenau beschrieben	-0.8	V	% unpräzise	betrachte % aufmerksam	sehr	eher	unbedingt	besser	ggf.	ggf.
Marker	Das Material ist an dieser Stelle abwechslungsreich	0.7	U	% vielschichtig	% zum lernen anregt	sehr		sehr		ggf.	ggf.
Marker	Das Material ist an dieser Stelle wenig abwechslungsreich	-0.7	U	% monoton	ziehe % alternative Materialien zum Lernen heran	sehr	eher	unbedingt	besser	ggf.	ggf.
Marker	Das Material ist an dieser Stelle detailliert beschrieben	0.8	S	% ausführlich	% schwierig ist	sehr		gar nicht	weniger	ggf.	ggf.
Marker	Das Material ist an dieser Stelle wenig detailliert beschrieben	-0.8	S	% oberflächlich	investiere % Zeit zum erarbeiten	sehr	eher	sehr viel	viel	ggf.	ggf.
Marker	Das Material ist an dieser Stelle besonders wichtig für die aktuelle Aufgabe	0.9	W	% relevant	% wichtig ist	sehr		sehr		ggf.	ggf.
Marker	Das Material ist an dieser Stelle wenig wichtig für die aktuelle Aufgabe	-0.9	W	% irrelevant	an einer anderen Stelle bzw. als Ergänzung % wichtiger	sehr	eher	deutlich	eher	ggf.	ggf.
Marker	Das Material ist an dieser Stelle besonders wichtig für die spätere Prüfung	0.9	W	% relevant	% wichtig ist	sehr		sehr		ggf.	ggf.
Marker	Das Material ist an dieser Stelle wenig wichtig für die spätere Prüfung	-0.9	W	% irrelevant	an einer anderen Stelle bzw. als Ergänzung % wichtiger	sehr	eher	deutlich	eher	ggf.	ggf.
Marker	Das Material hilft mir persönlich an dieser Stelle besonders gut	0.9	U	% nützlich	% zum lernen anregt	sehr		sehr		ggf.	ggf.
Marker	Das Material hilft mir persönlich an dieser Stelle wenig gut	-0.9	U	% unzweckmäßig	ziehe % alternative Materialien zum Lernen heran	sehr	eher	unbedingt	besser	ggf.	ggf.
Marker	Das Material ist an dieser Stelle besonders motivierend	0.8	U	% motivierend	% zum lernen anregt	sehr		sehr		ggf.	ggf.
Marker	Das Material ist an dieser Stelle wenig motivierend	-0.8	U	% demotivierend	ziehe % alternative Materialien zum Lernen heran	sehr	eher	unbedingt	besser	ggf.	ggf.
Marker	Die Präsentation des Materials hilft mir an dieser Stelle Inhalte zu erarbeiten	0.9	U	% anschaulich	% zum lernen anregt	sehr		sehr		ggf.	ggf.
Marker	Die Präsentation des Materials hilft mir an dieser Stelle wenig Inhalte zu erarbeiten	-0.9	U	% abstrakt	ziehe % alternative Materialien zum Lernen heran	sehr	eher	unbedingt	besser	ggf.	ggf.

Anhang C Aufgabenblatt Laborversuch

Aufgaben

Die nachfolgenden Aufgaben beschäftigen sich mit dem Themenbereich der Algorithmik. Zum einen beschäftigen wir uns in den Aufgaben mit dem Sortieren von Zahlenketten für deren Lösung es verschiedene Algorithmen gibt. Des Weiteren werden wir uns mit der Problematik kürzester Wege auseinandersetzen, wie man es z.B. aus Navigationssystemen kennt. Hierfür werden Graphenalgorithmus genutzt von denen wir einen genauer betrachten.

1.) Allgemein

- Definiere kurz was man unter einem Algorithmus versteht.
- Gib ein Beispiel für ein Handlungsgebiet eines Algorithmus an.
- In der Algorithmik gibt es verschiedene Grundbegriffe. Definiere was sich hinter den drei Grundbegriffen Eingabe, Ausgabe und Problem verbirgt.

2.) Insertionsort

Wir betrachten nun ein erstes Sortierproblem, in dem Zahlen in die richtige Reihenfolge gebracht werden sollen. Gegeben sei die Folge an Zahlen $F = [4,2,3,1,6,2]$. Sortiere die Folge unter Verwendung des Insertionsort Algorithmus. Notiere nach jedem Schritt wie die Folge aktuell sortiert ist und markiere das aktuell sortierte Element. Nutze hierfür die nachfolgende Tabelle:

4	2	3	1	6	2

3.) Mergesort

Anders als der Insertionsort Algorithmus ist der Mergesort Algorithmus ein Verfahren, welches nach dem „Divide & Conquer“ Prinzip arbeitet, um Sortierprobleme zu lösen.

- a) Definiere was man unter dem Prinzip eines „Divide & Conquer“ Algorithmus versteht und wo der Vorteil gegenüber eines Verfahrens wie z.B. dem Insertionsort Algorithmus liegt.
- b) Grob lässt sich der Mergesort Algorithmus in zwei Arbeitsschritte gliedern. Beschreibe kurz in eigenen Worten, welche Arbeitsschritte gemeint sind und was die jeweilige Aufgabe ist.
- c) Gegeben sei wiederum die Folge an Zahlen $F = [4,2,3,1,6,2]$. Sortiere die Folge unter Verwendung des Mergesort Algorithmus. Notiere nach jedem Schritt, wie die Folge aktuell sortiert ist. Markiere durch einkreisen oder nummerieren, welche Elemente aktuell zu einer Gruppe gehören. Nutze zum lösen der Aufgabe die nachfolgende Tabelle:

4	2	3	1	6	2

4.) Quicksort

Der Quicksort Algorithmus ist ein weiteres Verfahren zur Lösung von Sortierproblemen. Beantworte die nachfolgenden Fragen zum Quicksort Algorithmus.

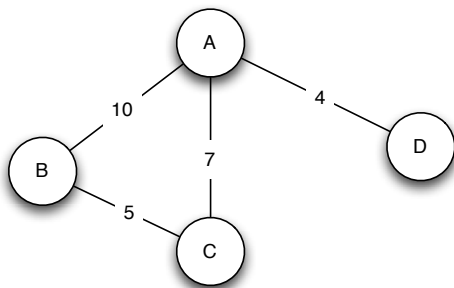
- a) Der Quicksort Algorithmus arbeitet wie der Mergesort Algorithmus nach dem „Divide & Conquer“ Prinzip (kreuze an): Richtig Falsch

- b) Was ist ein Pivotelement im Kontext des Quicksort Algorithmus und wie kann es ausgewählt werden?
- c) Was versteht man beim Quicksort Algorithmus unter „In-Place“ Sortierung?
- d) Welche alternative Form zur „In-Place“ Sortierung wird für Quicksort verwendet? Beschreibe kurz in eigenen Worten, wo der Unterschied zur „In-Place“ Variante liegt.

5.) Kürzeste Wege

Wir beschäftigen uns im Folgenden mit *Graphenalgorithmen* und in diesem Zusammenhang mit dem Problem kürzester Wege.

- a) Welche abstrakte Datenstruktur wird bei Problemen kürzester Wege genutzt? Gib zusätzlich eine formale Definition für die Datenstruktur an.
- b) Beschrifte in der nachfolgenden Abbildung was eine Kante, was ein Kantengewicht und was ein Knoten ist. Beschreibe kurz ob es sich um einen gerichteten oder einen ungerichteten Graphen handelt.

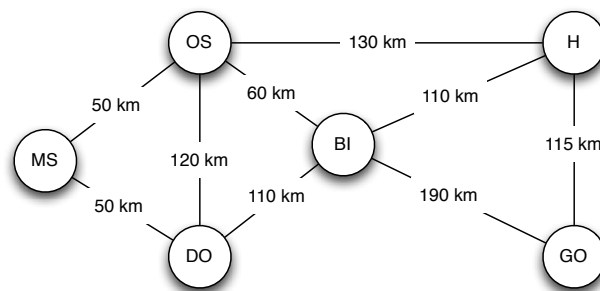


- c) Worin liegt der Unterschied zwischen den Probleme-Varianten Single-Source Shortest-Path (SSSP) und All-Pairs Shortest-Path (APSP)? Beschreibe kurz in eigenen Worten.
- d) Nenne jeweils einen Algorithmus der zur Lösung von SSSP und APSP Problemen genutzt werden kann.

SSSP: _____

APSP: _____

- e) Beschreibe kurz in eigenen Worten die Funktionsweise des Dijkstra Algorithmus.
- f) Was bedeutet es wenn ein Knoten während der Verwendung des Dijkstra Algorithmus den Wert ∞ besitzt?
- g) Im nachfolgenden Graph sind die Städte Osnabrück (OS), Münster (MS), Dortmund (DO), Bielefeld (BI), Hannover (H) und Göttingen (GO) als Knoten und die Entfernung zwischen den Städten (sofern vorhanden) als Kantengewicht dargestellt.

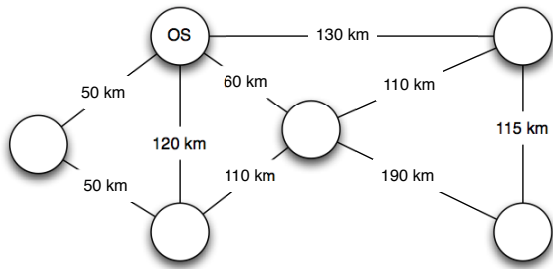


In welcher Reihenfolge werden bei der Verwendung des Dijkstra Algorithmus die Knoten ausgewählt? Schreibe die Knoten ausgehend vom Startpunkt Osnabrück (OS) der Reihe nach auf:

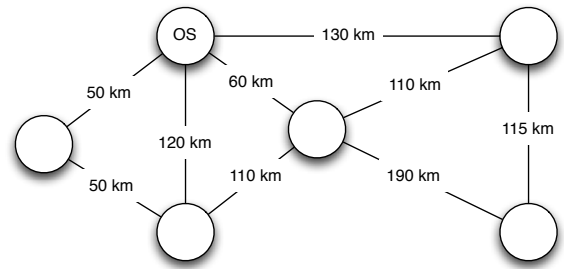
OS → _____

- h) Bestimme die kürzesten Wege von Osnabrück (OS) in alle anderen Städte unter Verwendung des Dijkstra Algorithmus und der Abbildung aus Teilaufgabe g.). Führe hierzu lediglich die ersten zwei Schritte des Algorithmus durch und trage diese in nachfolgende Abbildungen des Graphen ein.

1.)



2.)



Werkzeuge

Die Werkzeuge im LAOs System funktionieren für jedes Material in gleicher Form. Werkzeuge können über die Menuleiste am rechten Dokumentenrand aktiviert werden (Abbildung 1). Nachdem ein Werkzeug aktiviert wurde, öffnet sich nach einem weiteren Klick auf einen Materialbereich das Werkzeug Popup (Abbildung 2).

Alternativ kann über den Rechtsklick der Maus auf einer Materialstelle das Werkzeugmenu auf dem Material geöffnet werden (Abbildung 3). Nach einem weiteren Klick auf das jeweilige Material öffnen sich das Werkzeug Popup.



Abbildung 1

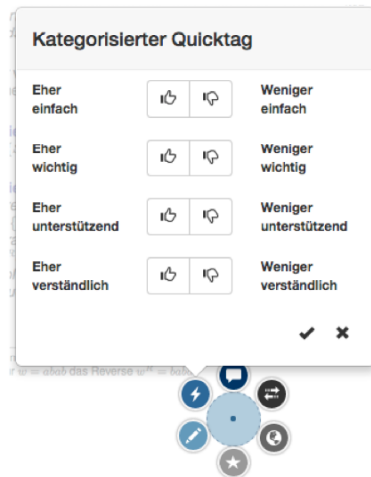


Abbildung 2



Abbildung 3

In jedem Werkzeug Popup hast Du verschiedene Möglichkeiten den Inhalt zu beschreiben (siehe Tabelle unten). Nachdem Du alle gewünschten Informationen eingegeben hast, bestätigst Du die Informationen über den Haken oder verwirfst die Informationen über das x rechts unten im Popup Fenster (vgl. Abbildung 2). Eine Annotation wird in der Tagline zwischen Material und Dokumentennavigation in der jeweiligen Farbe des Werkzeuges angezeigt (siehe Abbildung 4). Mit einem Klick auf die Annotation in der Tagline wird deine Annotation im Material angezeigt (siehe Abbildung 5). Hier hast Du die Möglichkeit über das Mülleimer Symbol die Annotation wieder zu löschen.

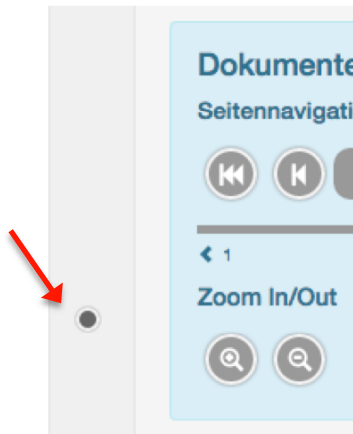


Abbildung 4

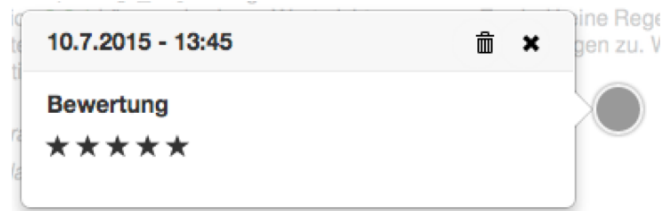


Abbildung 5

In einem Video hast Du zusätzlich die Möglichkeit Annotationen für einen zeitlichen Bereich zu erstellen. Hierzu aktivierst Du das Record Symbol in der Videosteuerung (Abbildung 6) und markierst in der Timeline den gewünschten zeitlichen Bereich (Abbildung 7). Den Bereich kannst Du dann mit einem gewünschten Inhalt annotieren.

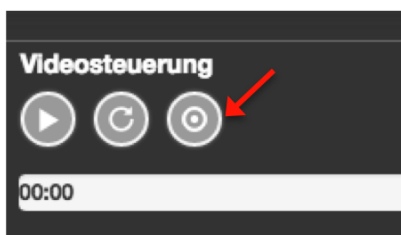


Abbildung 6

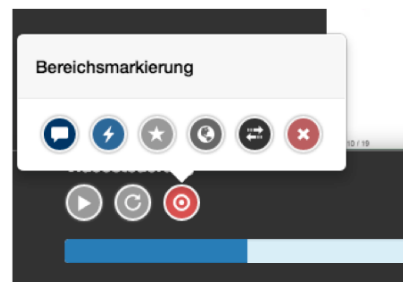








Abbildung 7

Legende Werkzeuge

Vordefinierter Kommentar	
	Mit einem vordefinierten Kommentar kannst Du sehr präzise Aussagen über ein Materialbereich treffen. Die Aussagen stehen dabei in positiver oder negativer Form zur Verfügung. Du kannst mit dem Werkzeuge mehrere Aussagen auf einmal auswählen.
Kategorisierter Quickttag	
	Über das Werkzeug kategorisierter Quickttag stehen dir vier verschiedenen Kategorien zur Verfügung, in die Du ein Materialbereich entweder in positiver oder negativer Form eingliedern kannst (z.B. eher schwierig oder weniger schwierig). Du kannst mehrere Kategorien auf einmal auswählen.
Highlighter	
	Den Highlighter kannst Du wie eine Art Marker benutzen, sofern Du einen ganzen Bereich markieren möchtest. Einen gehighlighteten Bereich kannst Du danach mit einem vordefinierten Kommentar beschreiben.
Bewertung	
	Mit einer Bewertung gibst Du ein generelles Feedback zu einem Materialbereich zwischen einem Stern – sofern dir etwas in einem Bereich nicht gefällt – bis hin zu fünf Sternen, sofern dir etwas besonders gut gefällt.
Neues Material	
	Du kannst auch neues Material mit in das System aufnehmen. Hierzu nutzt Du das Werkzeug neue Referenz. Du gibst einfach den Link zum neuen Material an und verknüpfst so einen Materialbereich mit einem alternativen Material.
Materialverbindung	
	Um Materialien aus dem bestehenden Materialpool miteinander zu verknüpfen nutzt Du das Werkzeug Materialverbindung. Du eröffnest in einem Material einfach eine neue offene Verbindung und kannst diese im selben oder einem anderen Material schließen. So verknüpfst Du zwei Materialstellen miteinander.

Anhang D Aufgabenblatt Online Simulation

Prof. Dr. K. Morisse

Übungsklausur Theoretische Informatik

01. Mai 2016

Name: _____

MatrNr: _____ Scriptversion: _____

Aufgabe	1	2	3	4	5	Summe
max. Punkte	20	15	8	8	10	61
err. Punkte						

Aufgabe 1: _____

Es sei $\Sigma = \{a, b\}$ und es sei

$$L = \{w \in \Sigma^* \mid \#_a(w) = 2i, \#_b(w) = 3j; i, j \geq 0\}$$

- (a) Man konstruiere einen vollständigen deterministischen endlichen Automaten als Zustandsdiagramm für die Sprache L [5]
- (b) Erläutern Sie kurz die Idee Ihrer Konstruktion ihre Konstruktion [5]
- (c) Untersuchen Sie, ob der von Ihnen konstruierte Automat M_L minimal ist. [5]
- (d) Konstruieren Sie eine Grammatik G mit $L(G) = L$ und erzeugen Sie das Wort *abbababbab* mit Hilfe der Grammatik. [5]

Aufgabe 1: 20

Aufgabe 2: _____

Es sei L ein Sprache über einem endlichen Alphabet Σ .

- (a) Beschreiben Sie kurz, wie sich allgemein der Nachweis führen lässt: [5]

$$L \in \mathcal{L}_1 \setminus \mathcal{L}_3.$$

- (b) Zeigen Sie dies am Beispiel der Sprache [10]

$$L = \{w \in \Sigma \mid \#_b(w) > 3 \cdot \#_a(w)\}$$

Aufgabe 2: 15

Aufgabe 3: _____

Welche der folgenden Behauptungen ist richtig bzw. falsch? Geben Sie einen Nachweis oder benennen Sie ein Gegenbeispiel.

- (a) Jede Teilmenge einer regulären Sprache ist regulär. [4]
- (b) Jede Obermenge einer regulären Sprache ist regulär. [4]

Aufgabe 3: 8

Aufgabe 4: _____ 8 Punkte

Es sei M der PDA mit der folgenden Übergangsfunktion:

$$\begin{aligned} \delta(q, 0, Z_0) &= \{(q, XZ_0)\} & \delta(p, \epsilon, X) &= \{(p, \epsilon)\} \\ \delta(q, 0, X) &= \{(q, XX)\} & \delta(p, 1, X) &= \{(p, XX)\} \\ \delta(q, 1, X) &= \{(q, X)\} & \delta(p, 1, Z_0) &= \{(p, \epsilon)\} \\ \delta(q, \epsilon, X) &= \{(p, \epsilon)\} & & \end{aligned}$$

Welche der nachfolgenden Konfigurationen kann von der Konfiguration $(p, 1101, XXZ_0)$ aus **nicht** erreicht werden?

- $(p, 1101, XZ_0)$
- $(p, 101, \epsilon)$
- $(p, 101, XXXXZ_0)$
- $(p, 101, XXXZ_0)$

Aufgabe 5: _____ 10 Punkte

Es sei G eine kontextfreie Grammatik in CNF mit den folgenden Regeln:

$$\begin{aligned} S &\rightarrow AB|BC \\ A &\rightarrow BA|a \\ B &\rightarrow CC|b \\ C &\rightarrow AB|a \end{aligned}$$

Prüfen Sie mittels des CYK-Algorithmus, ob das Wort $w = ababaa$ mit G erzeugt werden kann.

Anhang E Lernmaterialien Online Simulation

Lernmaterial	Größe
Vorlesungsskript Theoretische Informatik	261 Seiten
Übungsblatt	2 Seiten
Lernvideo Worte	04:52 Minuten
Lernvideo Wortoperationen	05:43 Minuten
Lernvideo Wortiteration	05:03 Minuten
Lernvideo Sprachen, Operationen	06:54 Minuten
Lernvideo Sprachen	06:23 Minuten
Lernvideo Ableitung	05:11 Minuten
Lernvideo Sprache einer Grammatik	02:06 Minuten
Lernvideo Sprache einer Grammatik Beispiel	08:47 Minuten
Lernvideo Rechtslineare Grammatik	04:41 Minuten
Lernvideo Kontextfreie Grammatik	01:08 Minuten
Lernvideo Kontextsensitive Grammatik	01:53 Minuten
Lernvideo Beschränkte Grammatik	07:00 Minuten
Lernvideo Beispiel rechtslineare Grammatik	03:58 Minuten
Lernvideo Beispiel kontextfreie Grammatik	04:57 Minuten
Lernvideo Audience Response Einsatz (Clicker)	08:48 Minuten
Lernvideo Chomsky Hierarchie	04:29 Minuten
Lernvideo Definition Endlicher Automat	04:37 Minuten
Lernvideo Beispiel endlicher Automat	06:24 Minuten
Lernvideo Sprache eines endlichen Automaten	05:55 Minuten
Lernvideo Automat als Graph	03:31 Minuten
Lernvideo Minuten Zusammenhang Sprache und Graph	04:32 Minuten
Lernvideo Beispiel Deterministischer endlicher Automat	14:34 Minuten
Lernvideo Beispiel DEA	05:18 Minuten
Lernvideo Definition NDEA	05:38 Minuten
Lernvideo Sprache NDEA	04:25 Minuten

Lernvideo Beispiel Clicker 1	07:32 Minuten
Lernvideo Beispiel Clicker 2	03:01 Minuten
Lernvideo eNDEA	05:01 Minuten
Lernvideo eNDEA Sprache	00:55 Minuten
Lernvideo Beispiel eNDEA	04:10 Minuten
Lernvideo Satz: Äquivalenz von eNDEA und NDEA	08:46 Minuten
Lernvideo Beispiel: Umwandlung eNDEA in NDEA	08:24 Minuten
Lernvideo Rationale Sprachen und rechtslineare Grammatiken	12:53 Minuten
Lernvideo Umwandlung rechtslineare Grammatik in endlichen Automaten (e-NDEA)	02:44 Minuten
Lernvideo Abschlusseigenschaften L3 (rationale Sprachen)	06:23 Minuten
Lernvideo Abschluss L3 gegenüber Negation	03:19 Minuten
Lernvideo Abschluss L3 gegenüber Vereinigung	04:52 Minuten
Lernvideo Abschluss L3 gegenüber Konkatenation	02:07 Minuten
Lernvideo Abschluss L3 gegenüber Kleene-Stern	02:22 Minuten
Lernvideo Abschluss L3 gegenüber Durchschnitt	01:09 Minuten
Lernvideo Produktautomat	16:28 Minuten
Lernvideo Abschluss L3 gegenüber Homomorphismus	03:59 Minuten
Lernvideo Abschluss der L3 Sprachen unter Homomorphismus	04:43 Minuten
Lernvideo Reguläre Ausdrücke	04:14 Minuten
Lernvideo Reguläre Ausdrücke: Sprache eines RegExp	03:58 Minuten
Lernvideo Beispiele regulärer Ausdrücke	05:39 Minuten
Lernvideo Konstruktion Automat aus regulärem Ausdruck	11:47 Minuten
Lernvideo Satz von Kleene: Rationale Sprachen = Reguläre Sprachen	28:22 Minuten
Lernvideo Beispiel Satz von Kleene	06:51 Minuten
Lernvideo Pumping Lemma für rationale Sprachen	07:31 Minuten
Lernvideo Analyse Pumping-Lemma	02:05 Minuten
Lernvideo Beispielanwendung Pumping-Lemma für rationale Sprachen	09:19 Minuten
Lernvideo Checkliste zur Anwendung des Pumping-Lemmas	01:37 Minuten

Lernvideo Relation Wort-Äquivalenz	03:54 Minuten
Lernvideo Beispiel zur Relation zwischen Worten bzgl. Sprache	05:16 Minuten
Lernvideo Beispiel für Äquivalenz bzgl. Sprache $(ab+ba)^*$	12:18 Minuten
Lernvideo Satz von Myhill-Nerode	05:02 Minuten
Lernvideo Zustandsäquivalenz in endlichen Automaten	03:39 Minuten
Lernvideo Zustandsäquivalenz endlicher Wortlänge	01:17 Minuten
Lernvideo Automatenminimierung	03:32 Minuten
Lernvideo Minimierung endlicher Automaten - Beispiel	11:37 Minuten
Lernvideo Entscheidbare Probleme für rationale Sprachen	05:18 Minuten
Lernvideo CFG - Kontextfreie Grammatiken	03:37 Minuten
Lernvideo Document Type Definition	08:29 Minuten
Lernvideo Ableitungsbaum für kontextfreie Sprachen	04:05 Minuten
Lernvideo Definition Ableitungsbaum	03:32 Minuten
Lernvideo Ordnung Ableitungsbaum	04:21 Minuten
Lernvideo Ordnung Ableitungsbaum (korrigiert)	04:21 Minuten
Lernvideo Zusammenhang Ableitungsbaum und Ableitung	01:12 Minuten
Lernvideo Eindeutigkeit von Ableitungen	02:35 Minuten
Lernvideo Links/Rechtsableitung	01:00 Minuten
Lernvideo Mehrdeutigkeit von Ableitungen	04:18 Minuten
Lernvideo Umformung CFG	02:51 Minuten
Lernvideo Chomsky-Normalform	03:22 Minuten
Lernvideo Greibach-Normalform	01:44 Minuten
Lernvideo Abschlusseigenschaften L2	08:19 Minuten
Lernvideo Einführung Kellerautomaten	06:57 Minuten
Lernvideo Formale Definition PDA	04:46 Minuten
Lernvideo Konfiguration und Arbeitsweise des PDA	04:46 Minuten
Lernvideo Grafische Darstellung Kellerautomat	05:45 Minuten
Lernvideo Akzeptierte Sprache eines Kellerautomaten	10:22 Minuten
Lernvideo Akzeptanz PDA: Finaler Zustand in leerer Keller	07:04 Minuten
Lernvideo Akzeptanz PDA: Leerer Keller in finaler Zustand	03:43 Minuten

Lernvideo PDA = L2 (kontextfreie Sprachen)	11:07 Minuten
Lernvideo Pumping-Lemma für kontextfreie Sprachen	07:39 Minuten
Lernvideo Anwendungsbeispiel Pumping-Lemma kontextfreie Sprachen	03:06 Minuten
Lernvideo Wortproblem in L2 und Idee CYK-Algorithmus	13:03 Minuten
Lernvideo Anwendungsbeispiel CYK-Algorithmus	08:31 Minuten

Anhang F Auswertung Benutzerstudie

	Teilnehmer	33						
	Materialien	7						
		Größe	Nutzungsda uer Materialien	Implizite Tags	Explizite Tags	Ressourcen	Drops	Drops (gezählt)
5	Skript	160	1667	4724	722	26	45	4
6	Buch	384	312	1529	38	10	0	0
7	Video Mergesort	04:55	88	212	28	2	7	0
8	Video Quicksort	04:30	84	282	32	3	6	0
9	Video Quicksort In-Place	07:35	31	224	13	1	1	0
10	Video Dijkstra	24:48	42	193	24	1	5	0
11	Video Dijkstra APSP	21:46	9	43	1	0	0	0
	Summe		2233	7207	858	43	64	4
	Durschnitt pro User		67,67	218,39	26			

Anhang G Auswertung Klassifikation Benutzerstudie

R-ID	Klassifikation (System)	Anzahl Tags	Alternat. Materialien	Relevant (User)	N. Relevant (User)	Abgleich
1161	0	42	0	1	0	TP
1162	0	72	1	1	0	TP
1163	0	99	1	1	0	TP
1164	0	34	2	1	0	TP
1165	0	5	1	1	0	TP
1166	0	24	2	1	0	TP
1167	0	20	1	1	0	TP
1168	0	37	3	1	0	TP
1169	0	65	3	1	0	TP
1170	0	35	1	1	0	TP
1171	0	18	1	1	0	TP
1172	0	5	1	0	1	FP
1173	0	3	0	0	1	FP
1174	0	15	1	1	0	TP
1175	0	3	0	0	1	FP
1176	0	3	0	0	1	FP
1177	0	5	0	1	0	TP
1178	0	20	1	1	0	TP
1179	0	8	0	1	0	TP
1180	0	9	0	1	0	TP
1181	0	7	0	1	0	TP
1182	0	4	2	1	0	TP
1183	0	5	0	1	0	TP
1184	0	8	0	1	0	TP
1185	0	4	1	0	1	FP
1186	0	6	0	0	1	FP
1187	0	7	0	1	0	TP
1188	0	4	1	1	0	TP
1189	0	6	0	1	0	TP
1190	0	6	0	1	0	TP
1191	0	2	0	0	1	FP
1192	0	4	0	1	0	TP
1193	0	2	0	1	0	TP
1194	0	2	0	0	1	FP
1195	0	4	0	1	0	TP
1196	0	4	0	0	1	FP
1198	0	16	0	1	0	TP
1199	0	11	0	1	0	TP

1200	0	3	0	0	1	FP
1201	0	11	0	1	0	TP
1202	0	8	0	1	0	TP
1203	0	22	0	1	0	TP
1204	0	9	1	1	0	TP
1205	1	3	0	0	1	TN
1206	1	2	0	0	1	TN
1207	1	3	0	0	1	TN
1208	1	8	0	1	0	FN
1209	1	3	0	0	1	TN
1210	1	3	0	0	1	TN
1211	1	2	2	0	1	TN
1212	1	2	0	0	1	TN
1213	1	2	0	0	1	TN
1214	1	2	1	1	0	FN

Anhang H Auswertung online Simulation

	Teilnehmer	90								
	Teilnehmer angemeldet	56								
	Teilnehmer mit geringfügiger Anmeldung	23								
	Teilnehmer mit Bewertungen	25								
	Materialien	90								
		Größe	Nutzungsdauer	Implizite Tags	Implizite Tags (Datenanalyse)	Explizite Tags	Resource n	Bewertungen	Drops	Drops (gezählt)
12	Skript	261	6749	5590	5221	988	105	198	78	11
13	Worte	04:52	1478	290	217	5	2	12	1	0
14	Wortoperationen	05:43	83	85	53	7	0	0	1	0
15	Wortiteration	05:03	113	58	33	3	0	0	3	0
16	Sprachen, Operationen	06:54	412	106	63	3	0	0	1	0
17	Sprachen	06:23	4810	82	60	8	0	0	1	0
18	Ableitung	05:11	57	79	45	1	0	0	1	0
19	Sprache einer Grammatik	02:06	166	39	30	1	0	0	1	0
20	Sprache einer Grammatik Beispiel	08:47	186	88	74	9	0	0	2	0
21	Rechtslineare Grammatik	04:41	310	75	27	2	0	0	1	0
22	Kontextfreie Grammatik	01:08	23	59	36	4	0	0	1	0
23	Kontextsensitive Grammatik	01:53	19	40	24	3	0	0	1	0
24	Beschränkte Grammatik	07:00	33	80	31	1	0	0	1	0
25	Beispiel rechtslineare Grammatik	03:58	4	33	0	0	0	0	0	0
26	Beispiel kontextfreie Grammatik	04:57	6	25	0	0	0	0	0	0
27	Audience Response Einsatz (Clicker)	08:48	18	130	18	0	0	0	0	0
28	Chomsky Hierarchie	04:29	64	185	97	3	0	0	1	0
29	Definition Endlicher Automat	04:37	73	95	77	5	0	0	2	0
30	Beispiel endlicher Automat	06:24	20	106	87	1	0	0	1	0
31	Sprache eines endlichen Automaten	05:55	56	66	34	2	0	0	1	0
32	Automat als Graph	03:31	11	71	35	3	0	0	1	0
33	Zusammenhang Sprache und Graph	04:32	371	121	58	0	0	0	0	0
34	Beispiel Deterministischer endlicher Automat	14:34	960	134	100	5	0	0	2	0
35	Beispiel DEA	05:18	115	92	44	5	0	0	1	0
36	Definition NDEA	05:38	16	15	7	0	0	0	0	0
37	Sprache NDEA	04:25	13	28	0	0	0	0	0	0
38	Beispiel Clicker 1	07:32	43	22	2	0	0	0	0	0
39	Beispiel Clicker 2	03:01	3	16	0	0	0	0	0	0
40	eNDEA	05:01	24	24	0	0	0	0	0	0
41	eNDEA Sprache	00:55	3	6	0	0	0	0	0	0

42	Beispiel eNDEA	04:10	53	25	10	1	0	0	1	0
43	Satz: Äquivalenz von eNDEA und NDEA	08:46	507	115	31	1	0	0	1	0
44	Beispiel: Umwandlung eNDEA in NDEA	08:24	37	44	6	0	0	0	0	0
45	Rationale Sprachen und rechtslineare Grammatiken	12:53	120	63	44	2	0	0	1	0
46	Umwandlung rechtslineare Grammatik in endlichen Automaten (e-NDEA)	02:44	153	20	7	0	0	0	0	0
47	Abschlusseigenschaften L3 (rationale Sprachen)	06:23	78	49	8	0	0	0	0	0
48	Abschluss L3 gegenüber Negation	03:19	248	17	2	0	0	0	0	0
49	Abschluss L3 gegenüber Vereinigung	04:52	23	45	0	0	0	0	0	0
50	Abschluss L3 gegenüber Konkatenation	02:07	11	22	0	0	0	0	0	0
51	Abschluss L3 gegenüber Kleene-Stern	02:22	11	24	0	0	0	0	0	0
52	Abschluss L3 gegenüber Durchschnitt	01:09	9	11	0	0	0	0	0	0
53	Produktautomat	16:28	651	99	7	0	0	0	0	0
54	Abschluss L3 gegenüber Homomorphismus	03:59	19	36	8	1	0	0	1	0
55	Abschluss der L3 Sprachen unter Homomorphismus	04:43	24	43	7	0	0	0	0	0
56	Reguläre Ausdrücke	04:14	95	67	23	2	0	0	1	0
57	Reguläre Ausdrücke: Sprache eines RegExp	03:58	17	16	0	0	0	0	0	0
58	Beispiele regulärer Ausdrücke	05:39	16	52	16	0	0	0	0	0
59	Konstruktion Automat aus regulärem Ausdruck	11:47	127	96	73	0	0	0	0	0
60	Satz von Kleene: Rationale Sprachen = Reguläre Sprachen	0:28:22	704	213	162	2	0	0	1	0
61	Beispiel Satz von Kleene	06:51	32	56	1	0	0	0	0	0
62	Pumping Lemma für rationale Sprachen	07:31	83	130	31	1	0	0	1	0
63	Analyse Pumping-Lemma	02:05	34	24	0	0	0	0	0	0
64	Beispielanwendung Pumping-Lemma für rationale Sprachen	09:19	55	50	0	0	0	0	0	0
65	Checkliste zur Anwendung des Pumping-Lemmas	01:37	438	94	43	2	0	0	1	0
66	Relation Wort-Äquivalenz	03:54	25	23	14	2	0	0	1	0
67	Beispiel zur Relation zwischen Worten bzgl. Sprache L	05:16	24	23	0	0	0	0	0	0
68	Beispiel für Äquivalenz bzgl. Sprache $(ab+ba)^*$	12:18	41	28	0	0	0	0	0	0
69	Satz von Myhill-Nerode	05:02	75	55	6	0	0	0	0	0
70	Zustandsäquivalenz in endlichen Automaten	03:39	102	14	8	0	0	0	0	0
71	Zustandsäquivalenz endlicher Wortlänge	01:17	90	20	7	0	0	0	0	0
72	Automatenminimierung	03:32	314	340	255	10	1	3	2	0
73	Minimierung endlicher Automaten - Beispiel	11:37	1213	432	382	34	2	8	7	0
74	Entscheidbare Probleme für rationale Sprachen	05:18	44	25	3	0	0	0	0	0
75	CFG - Kontextfreie Grammatiken	03:37	17	47	37	1	0	0	1	0
76	Document Type Definition	08:29	12	27	0	0	0	0	0	0
77	Ableitungsbaum für kontextfreie Sprachen	04:05	44	25	0	0	0	0	0	0
78	Definition Ableitungsbaum	03:32	11	19	0	0	0	0	0	0
79	Ordnung Ableitungsbaum	04:21	10	13	0	0	0	0	0	0
80	Ordnung Ableitungsbaum (korrigiert)	04:21	8	16	1	0	0	0	0	0
81	Zusammenhang Ableitungsbaum und Ableitung	01:12	4	5	0	0	0	0	0	0
82	Eindeutigkeit von Ableitungen	02:35	6	7	0	0	0	0	0	0

83	Links/Rechtsableitung	01:00	3	3	0	0	0	0	0	0
84	Mehrdeutigkeit von Ableitungen	04:18	13	23	0	0	0	0	0	0
85	Umformung CFG	02:51	14	16	2	0	0	0	0	0
86	Chomsky-Normalform	03:22	17	27	12	0	0	0	0	0
87	Greibach-Normalform	01:44	10	15	2	0	0	0	0	0
88	Abschlusseigenschaften L2	08:19	30	37	2	0	0	0	0	0
89	Einführung Kellerautomaten	06:57	38	29	9	0	0	0	0	0
90	Formale Definition PDA	04:46	128	26	13	0	0	0	0	0
91	Konfiguration und Arbeitsweise des PDA	04:46	106	105	79	0	0	0	0	0
92	Grafische Darstellung Kellerautomat	05:45	127	80	61	0	0	0	0	0
93	Akzeptierte Sprache eines Kellerautomaten	10:22	45	34	2	0	0	0	0	0
94	Akzeptanz PDA: Finaler Zustand in leerer Keller	07:04	19	27	0	0	0	0	0	0
95	Akzeptanz PDA: Leerer Keller in finaler Zustand	03:43	11	5	0	0	0	0	0	0
96	PDA = L2 (kontextfreie Sprachen)	11:07	79	67	8	0	0	0	0	0
97	Pumping-Lemma für kontextfreie Sprachen	07:39	104	41	0	0	0	0	0	0
98	Anwendungsbeispiel Pumping-Lemma kontextfreie Sprachen	03:06	43	56	38	0	0	0	0	0
99	Wortproblem in L2 und Idee CYK-Algorithmus	13:03	2454	605	428	61	3	27	14	0
100	Anwendungsbeispiel CYK-Algorithmus	08:31	2155	1066	682	67	4	22	12	0
101	Übungsklausur	2	21995	852	769	123	5	54	7	0
	Summe		49215	13484	9772	1369	122	324	154	11
	Durchschnitt pro User		878,8392857	240,7857143	174,5	24,44642857		12,96		

Anhang I Auswertung Klassifikation online Simulation

R-ID	Klassifikation (System)	Anzahl Tags	Alternativen	Relevant (User)	N. Relevant (User)	N. Klassifiziert	Bew. 2	Bew. 1	Bew. 0	Bew. -1	Bew. -2	N. Bewertet	Bew. Mittelwert
1382	0	41	8	8	0	5	4	5	1	1	0	2	4,090909091
1383	0	32	2	6	0	4	5	1	1	0	0	3	4,571428571
1384	0	10	2	3	2	6	2	2	2	0	1	4	3,571428571
1385	0	12	5	8	0	3	5	4	1	0	0	1	4,4
1386	0	27	7	6	0	3	3	4	0	0	0	2	4,428571429
2020	0	4	1	4	0	1	1	0	2	0	1	1	3
2021	0	4	0	2	0	2	0	0	0	0	2	2	1
2022	0	12	0	1	1	0	0	0	1	0	1	0	2
2024	0	6	1	1	0	0	1	0	0	0	0	0	5
2025	0	8	0	1	0	0	0	0	0	0	0	1	-
2026	0	7	0	1	0	1	1	0	0	0	0	1	5
2027	0	24	1	0	0	1	0	0	0	0	0	1	-
2028	0	41	0	0	0	1	0	0	0	0	0	1	-
2029	0	14	0	0	0	1	0	0	0	0	0	1	-
2030	0	3	0	0	0	1	0	0	0	0	0	1	-
2031	0	23	4	0	0	1	0	0	0	0	0	1	-
2032	0	8	2	0	0	1	0	0	0	0	0	1	-
2033	0	10	0	0	0	1	0	0	0	0	0	1	-
2039	0	11	0	0	1	1	0	0	0	0	1	1	1
2040	0	14	0	1	0	0	0	0	1	0	0	0	3
2041	0	18	0	1	0	0	0	1	0	0	0	0	4
2043	0	12	3	1	0	0	1	0	0	0	0	0	5
2044	0	17	0	1	0	0	0	1	0	0	0	0	4
2066	0	6	1	2	0	0	2	0	0	0	0	0	5
2067	0	12	1	2	0	0	1	1	0	0	0	0	4,5
2068	0	9	2	1	0	0	1	0	0	0	0	0	5
2069	0	6	1	3	0	0	2	1	0	0	0	0	4,666666667
2070	0	30	0	3	0	0	1	1	1	0	0	0	4
2071	0	7	0	3	0	0	0	1	2	0	0	0	3,333333333
2072	0	11	1	4	0	0	1	3	0	0	0	0	4,25
2073	0	9	0	3	0	0	2	0	0	0	0	1	5
2074	0	17	0	4	0	0	0	0	4	0	0	0	3
2075	0	5	0	2	0	2	0	1	2	0	0	1	3,333333333
2076	0	14	1	2	0	0	1	1	0	0	0	0	4,5
2077	0	7	2	3	0	0	2	1	0	0	0	0	4,666666667
2078	0	4	0	3	0	0	1	0	1	0	0	1	4

2079	0	7	0	1	0	0	0	0	0	0	0	1	-
2080	0	11	0	0	3	0	0	0	2	0	0	1	3
2081	0	20	0	3	0	0	3	0	0	0	0	0	5
2083	0	14	0	3	0	0	0	0	2	0	0	1	3
2084	0	10	0	2	0	0	2	0	0	0	0	0	5
2085	0	14	0	1	1	1	0	0	1	0	0	2	3
2086	0	12	0	1	0	1	0	1	0	0	0	1	4
2087	0	6	0	1	0	0	1	0	0	0	0	0	5
2088	0	19	0	3	0	0	2	1	0	0	0	0	4,666666 667
2089	0	2	0	3	0	0	0	0	2	0	1	0	2,333333 333
2091	0	30	0	1	0	0	1	0	0	0	0	0	5
2096	0	3	0	2	1	0	0	0	0	0	3	0	1
2097	0	4	0	0	0	1	0	0	0	0	0	1	#DIV/0!
2098	0	3	0	0	3	0	0	0	0	0	3	0	1
2099	0	3	0	0	3	0	0	0	1	0	2	0	1,666666 667
2100	0	4	0	0	3	0	0	0	0	3	0	0	2
2101	0	10	1	5	0	0	4	1	0	0	0	0	4,8
2102	0	5	0	3	0	0	2	0	0	0	0	1	5
2103	0	2	0	1	1	0	0	0	1	0	0	1	3
2104	0	3	0	1	1	0	0	0	1	0	0	1	3
2105	0	3	0	2	0	0	0	0	1	0	1	0	2
2106	0	16	2	5	0	0	4	1	0	0	0	0	4,8
2107	0	7	0	3	0	0	1	1	1	0	0	0	4
2108	0	4	0	0	2	0	0	0	1	0	4	0	1,4
2109	0	3	0	3	0	2	0	0	2	0	1	2	2,333333 333
2110	0	5	0	3	0	0	0	0	2	0	1	0	2,333333 333
2111	0	4	0	1	0	0	0	0	1	0	0	0	3
2112	0	13	1	3	0	0	2	1	0	0	0	0	4,666666 667
2113	0	6	0	2	0	0	2	0	0	0	0	0	5
2114	0	3	0	0	1	1	0	0	1	0	0	1	3
2115	0	5	0	0	0	2	0	0	0	0	0	2	-
2116	0	2	0	0	1	1	0	0	0	0	1	1	1
2117	0	3	0	0	0	1	0	0	0	0	0	1	-
2118	0	16	0	4	0	1	1	1	2	0	0	1	3,75
2119	0	9	0	3	0	1	2	0	1	0	0	1	4,333333 333
2120	0	19	1	5	0	1	0	5	0	0	0	1	4
2121	0	32	0	5	0	0	3	2	0	0	0	0	4,6
2122	0	24	0	4	0	1	2	1	1	0	0	1	4,25
2123	0	23	0	1	0	1	0	1	0	0	0	1	4
2124	0	26	0	4	0	1	1	1	2	0	0	1	3,75
2125	0	6	0	2	2	1	0	0	2	0	1	2	2,333333 333

2126	0	18	2	4	1	1	1	3	0	1	0	1	3,8
2127	0	20	0	4	0	2	2	1	1	0	0	2	4,25
2131	0	9	1	8	0	1	5	2	1	0	0	1	4,5
2132	0	16	1	7	2	1	4	3	1	1	0	1	4,111111 111
2133	0	19	0	4	2	2	0	1	2	0	4	1	2
2136	0	17	0	3	0	1	0	0	2	0	0	2	3
2137	0	11	1	3	0	1	3	0	0	0	0	1	5
2138	0	8	1	3	0	0	3	0	0	0	0	0	5
2139	0	4	0	2	2	2	0	0	2	0	2	2	2
2140	0	3	0	5	1	0	3	0	1	0	1	1	3,8
2141	1	2	0	1	4	0	1	0	0	2	2	0	2,2
2142	1	2	0	0	5	0	0	0	0	2	3	0	1,4
2143	1	3	0	1	1	0	0	0	1	0	1	0	2
2144	1	6	0	0	2	0	0	0	0	1	0	1	2
2145	1	4	0	0	2	1	0	0	0	0	2	1	1
2146	1	4	0	0	1	1	0	0	0	0	1	1	1
2147	1	2	0	0	1	0	0	0	0	0	1	0	1

Anhang J Auswertung Akzeptanzbefragung

ID	Quality	Trifft voll zu (5)	Trifft zu (4)	Neutral (3)	Trifft nicht zu (2)	Trifft gar nicht zu (1)
1	Die Empfehlungen waren für die Suche nach Informationen wertvoll	0	12	7	1	1
2	Ich hätte die Bereiche der Empfehlungen ebenfalls markiert	10	7	2	1	1
3	Die Empfehlungen würde meine Produktivität erhöhen	0	8	10	2	1
4	Die Empfehlungen würde es mir erleichtern Inhalte zu erlernen	0	11	7	2	1
5	Mir würde es gefallen, dass das System meine Stärken und Schwächen erkennt	3	9	7	2	0
6	Ich habe durch das System Inhalte entdeckt, die ich vorher nicht kannte	1	13	5	1	1
7	Ich habe durch das System neue Materialien entdeckt, die ich vorher nicht kannte	2	15	1	2	1
	Effort					
8	Ich habe mich im System schnell zurecht gefunden	6	12	1	1	1
9	Mir fiel es leicht durch das Textdokument zu navigieren	5	14	1	1	0
10	Mir fiel es leicht durch das Videodokument zu navigieren	2	15	3	1	0
11	Mir fiel es leicht Werkzeuge im System zu nutzen und Annotationen zu setzen	4	15	0	1	1
12	Die Auswahl der Werkzeuge war übersichtlich	0	16	3	0	2
13	Die Auswahl der Werkzeuge war verständlich	0	16	3	1	1
14	Die Menge an Werkzeugen war ausreichend	1	6	13	1	0
15	Die Empfehlungen waren leicht zugänglich	2	11	7	1	0
16	Die Darstellung der Empfehlungen war verständlich	0	18	1	0	2
17	Das System würde mich beim Lernen einschränken (negativ)	1	3	3	6	8
18	Ich würde durch das System länger für die Bearbeitung von Aufgaben benötigen (negativ)	1	3	4	5	8
	Acceptance					
19	Mir würde es helfen zu wissen welche Informationen in Lernmaterialien wichtig, unterstützend, schwierig oder verständlich sind	3	15	2	0	1
20	Mir würde es helfen zu wissen wie andere Studierende Lerninhalte einschätzen	4	11	5	0	1
21	Ich würde das System in der regulären Lehre nutzen wollen	2	13	4	1	1
22	Mir würde es helfen alternative Lerninhalte vorgeschlagen zu bekommen	3	12	4	1	1