

Orientierungsunterstützung in Ressourcenräumen mithilfe interaktiver Visualisierungen

Marcel Kliemann, Sebastian Groß, Niels Pinkwart

Institut für Informatik, Humboldt-Universität zu Berlin

Zusammenfassung

In diesem Beitrag wird untersucht, wie die Orientierung in umfangreichen Sammlungen digitaler Ressourcen mithilfe interaktiver Visualisierungen unterstützt werden kann. Solche Ressourcenräume können dabei aus einer heterogenen Menge an digitalen Ressourcen wie z. B. Wikipedia-Artikel, Lehrbücher und Videos bestehen. Da heutzutage ein einfacher Zugang zu einer Fülle von Ressourcen zu jedem beliebigen Thema möglich ist, kann es bei der manuellen Auswertung schnell zu einer Informationsüberflutung kommen. Darum werden effiziente Techniken benötigt, solche Ressourcen automatisiert zu analysieren und zu strukturieren. Dazu wurde ein prototypisches System entwickelt, welches eine Visualisierungspipeline zur automatisierten Extraktion von Informationsdimensionen aus Ressourcen ermöglicht, diese in eine Abhängigkeitsstruktur gruppiert und schließlich in Form einer TreeMap bzw. eines Wurzelbaums visualisiert und dem Nutzer zugänglich macht. Im Rahmen einer Laborstudie konnte gezeigt werden, dass eine Strukturierung der Ressourcen basierend auf einer Stichwörter-Hierarchie als Unterstützung zur Orientierung in Ressourcenräumen geeignet ist.

1 Einleitung

Im Zeitalter stetig ansteigender Datenmengen (Big Data) und sozialen Medien hat sich das Finden und Publizieren von digitalen Informations- und Lernressourcen grundlegend verändert. Durch Suchmaschinen und Datenbanken mit Millionen von Büchern und wissenschaftlichen Publikationen besteht heutzutage ein fast uneingeschränkter Zugang zu solchen Ressourcensammlungen. Neben den klassischen Publikationswegen sind mit Kollaborationsprojekten wie Wikipedia, Blogs und sozialen Netzwerken neue Quellen entstanden, die die Menge digitaler Ressourcen stetig wachsen lassen.

Zusammenhängende Ressourcen bilden einen Ressourcenraum. Möchte sich jemand ein Thema erschließen, lässt sich heutzutage schnell ein Ressourcenraum bestehend z. B. aus

Veröffentlicht durch die Gesellschaft für Informatik e.V. 2016 in
W. Prinz, J. Borchers, M. Ziefle (Hrsg.):
Mensch und Computer 2016 – Tagungsband, 4. - 7. September 2016, Aachen.
Copyright © 2016 bei den Autoren.
<http://dx.doi.org/10.18420/muc2016-mci-0050>

Wikipedia- und Blogartikeln, Lehrbüchern, Vorlesungsfolien etc. zusammensuchen. Um mit diesen arbeiten zu können, müssen die einzelnen Ressourcen analysiert und der gesamte Ressourcenraum nach Abhängigkeiten strukturiert werden.

In diesem Artikel wird untersucht, wie Nutzer bei der Orientierung in digitalen Ressourcenräumen unterstützt werden können. Basierend auf Informationen, die aus den Inhalten bzw. Metainformationen der Ressourcen gewonnen wurden, kann dann ein solcher Ressourcenraum strukturiert visualisiert werden. In Abschnitt 2 betrachten wir zunächst den Stand der Forschung. Abschnitt 3 skizziert den Ansatz zur automatischen Analyse und der Visualisierung von Ressourcenräumen. In einer Laborstudie wurde der Ansatz evaluiert. Die Ergebnisse stellen wir in Abschnitt 4 dar. Abschließend ziehen wir ein Fazit und geben einen Ausblick auf zukünftige Forschungen in Abschnitt 4.4.

2 Stand der Forschung

Um einen effizienten Zugang zu einem Ressourcenraum zu ermöglichen, eignet es sich die Strukturen visuell aufzubereiten. Hierdurch kann das Verständnis von Inhalten und Zusammenhängen gefördert werden (Nesbit und Adesope 2006). Ein Software-Konzept hierfür ist „CmapTools“ (Cañas und Novak, 2014), mit dem sich allgemein Concept-Maps visuell erzeugen lassen und so händisch eine Strukturierung von Ressourcen vorgenommen werden kann. Eine halb-automatisierte Strukturierung erfolgt bei dem Konzept „Nestor“ (Zeiliger und Esnault, 2014) durch die visuelle Aufbereitung des Verlaufes der genutzten Web-Ressourcen bei einer Recherche, welche händisch durch den Benutzer ergänzt werden können. Weiter existieren Konzepte, mit denen sich die bereits gewichteten Suchmaschinenergebnisse visuell aufbereiten lassen (Chen et al. 2010). Diesen statischen Visualisierungen mangelt es jedoch an effizienten Orientierungstechniken. Eine Ressource enthält typischerweise mehrere Informationsdimensionen (Inhalt, Quelle, Autoren, Qualitätsbewertung, etc.) besitzen, so dass bei einer größer werdenden Anzahl von Ressourcen die Informationsdichte stark zunehmen und eine Informationsüberflutung eintreten kann. Andere Ansätze untersuchen, wie sich Strukturen aus Benutzerpräferenzen bzw. Empfehlungen von anderen Benutzern finden lassen (Hsieh und Wang 2010). Die Informationen für diese Art der Strukturierung stehen jedoch in den meisten Fällen nicht von Beginn an zur Verfügung, womit hier eine vollständige Automatisierung des Prozesses nicht stattfinden kann.

Der Vorgang der Analyse und Strukturierung stellt sich mit dem Anstieg der zur Verfügung stehenden Mengen an Ressourcen zunehmend als schwierig dar. Das effiziente Identifizieren von relevanten Informationen und den Abhängigkeiten zwischen diesen kann ab einem bestimmten Punkt nicht mehr händisch geschehen. Im Forschungsfeld des Text Mining existieren hierfür Konzepte, wie sich Mengen von Textdokumenten automatisiert analysieren und in Strukturen bringen lassen (Aggarwal und Zhai 2012). Allerdings geht es in diesen Ansätzen mehr darum, Zusammenhänge in (sehr großen) Ressourcenraum zu finden und weniger darum, den Zugang zu einzelnen relevanten Ressourcen zu ermöglichen.

In diesem Beitrag stellen wir einen Ansatz vor, der Ressourcenräume basierend auf Abhängigkeitsstrukturen zwischen Ressourcen visualisiert, um dadurch eine Unterstützung zur Orientierung zu bieten. Hierzu wird zunächst eine Verkettung von Verfahren vorgestellt, die eine inhaltliche hierarchische Abhängigkeitsstruktur zwischen Ressourcen identifiziert und ableitet. Jede Ebene der Hierarchie wird durch eine Menge von Stichwörtern beschrieben, die den darunterliegenden Teil-Ressourcenraum möglichst kompakt inhaltlich beschreibt. Zur Visualisierung dieser Abhängigkeitsstrukturen werden anschließend zwei verschiedene Darstellungsformen vorgestellt, die in einer Laborstudie hinsichtlich ihrer Unterstützungseigenschaften untersucht werden.

3 Analyse und Visualisierung von Ressourcenräumen

Der Ausgangspunkt eines Ressourcenraumes ist eine unstrukturierte Menge von digitalen Ressourcen, die in unterschiedlichsten Formaten (z. B. als Textdokument, HTML-Quelltext einer Website, PDF, Audio, Video etc.) und Formen (z. B. Fachliteratur, Vorlesungsmaterial, etc.) vorliegen können. Auf der Visualisierungspipeline von Card (Card et al. 1999) aufbauend, wird im Folgenden eine Verfahrensverkettung dargestellt, welche automatisiert eine Analyse und Strukturierung vornimmt. Das Resultat dient dann als Grundlage für eine Visualisierung.

3.1 Identifizieren von Abhängigkeitsstrukturen

Ein Ressourcenraum z. B. zu dem Thema „Software Engineering“ gliedert sich u. a. in die Teil-Ressourcenräume „Vorgehensmodell“ und „Softwarearchitektur“ und letzterer wiederum u. a. in „Grobentwurf“ und „Feinentwurf“. Sei eine Menge von Ressourcen für einen Ressourcenraum gegeben, existiert quasi bereits eine natürliche hierarchische Struktur von Teil-Ressourcenräumen über diesen. Wird diese Struktur als Wurzelbaum repräsentiert, liegen an den Blättern die Ressourcen. Inhaltlich ähnliche Ressourcen bilden dann einen Teil-Ressourcenraum und können zu einem inneren Knoten zusammengefasst werden. Mehrere innere Knoten können nach demselben Prinzip wieder zu einem Elternknoten zusammengefasst werden, bis die Wurzel als Gesamt-Ressourcenraum erreicht wurde. Diese Hierarchie besitzt die Eigenschaft, dass sie in jeder Ebene des Baumes einen unterschiedlichen Grad der Konkretisierung des Inhalts besitzt, der von der Wurzel ausgehend zunimmt. Orientiert sich ein Benutzer anhand dieser Hierarchie, kann er die Menge an Ressourcen mit dem zunehmenden Abstieg in den Ebenen verkleinern und auf für ihn relevante Informationen reduzieren. Die inneren Knoten des Baumes dienen demnach als Orientierung in dem Ressourcenraum. Diese Knoten benötigen eine kompakte Beschreibung des Teil-Ressourcenraumes, anhand der eine effiziente Orientierung ermöglicht werden kann. Eine solche Beschreibung ergibt sich aus einer Menge von Stichwörtern, da diese schnell erfasst werden können. Um diese für jeden Knoten zu erhalten, müssten diese zunächst aus den Inhalten der zugrundeliegenden Ressourcen extrahiert werden. Aus diesem Konzept der Abhängigkeitsstrukturen ergibt sich die Anforderung, dass die Ressourcen für die Gruppierung zu Knoten inhaltlich verglichen werden müssen. Die inhaltliche Beschreibung der Ressourcen muss dann in der Knotenhierarchie nach oben weitergeben werden.

Der erste Schritt der Visualisierungspipeline erfordert zunächst eine Datenmodellierung, durch die alle Ressourcen in eine einheitliche Datenstruktur gebracht werden. Auf dieser aufbauend kann im zweiten Schritt eine Abhängigkeitsstruktur erzeugt werden. Bei der Modellierung sollten alle Informationsdimensionen in eine Form überführt werden, die einen Vergleich dieser Dimensionen zulässt. Für die Dimensionen, die Metainformationen enthalten (Autoren, Bewertungen, Anzahl der Zitierungen etc.), ist dies z. B. durch das Auslesen aus Literaturdatenbanken möglich. Der Inhalt der Ressourcen liegt jedoch meist in einer heterogenen Menge von unterschiedlichen Arten vor.

Die Stichwörter für die zu generierende Struktur müssen aus den Ressourcen gewonnen werden und den Inhalt möglichst kompakt repräsentieren. Um diese zu extrahieren, werden sämtliche textuellen Inhalte mithilfe des n -Gramm-Modells (Rajaraman und Ullman 2012) in Wortgruppen (Terme) bestehend aus n Wörtern zerlegt. Da nicht alle Terme eines Textes repräsentativ für dessen Inhalt stehen, müssen diese gewichtet werden. Mithilfe des Tf-idf-Maßes (Manning et al. 2008) werden Terme, die häufig in vielen Ressourcen vorkommen, niedrig gewichtet und Terme, die häufig in wenigen Ressourcen vorkommen, höher gewichtet. Wird eine Teilmenge, bestehend aus den höchst gewichteten Termen entnommen, besitzt damit jede Ressource eine einheitliche Repräsentation der Beschreibung ihres Inhaltes in Form einer Menge von Stichwörtern.

Nachdem die Ressourcen in einer einheitlichen Datenstruktur vorliegen, kann im zweiten Schritt der Visualisierungspipeline die Abhängigkeitsstruktur generiert werden. Mithilfe eines hierarchisch agglomerativen Clusterverfahren (Manning et al. 2008) lassen sich nun zunächst die Ressourcen nach ihrer inhaltlichen Ähnlichkeit (Huang 2008) zu Clustern (innere Knoten der Baumstruktur) gruppieren. Ähnliche Cluster lassen sich für eine neue Ebene der Hierarchie wieder zu neuen Clustern zusammenführen, bis schließlich ein einzelnes Cluster übrigbleibt (der Wurzelknoten). Aus der Schnittmenge der Stichwörter aller in einem Cluster liegenden Ressourcen ergeben sich die inhaltlichen Bezeichnungen für die Cluster.

3.2 Visualisierungstechniken für eine prototypische Umsetzung

Mit der gewonnenen Abhängigkeitsstruktur sind die Ressourcen nicht direkt miteinander verbunden, sondern indirekt über die Knoten des über ihnen aufgebauten hierarchischen Wurzelbaumes. Für die Darstellung solcher hierarchischen Strukturen finden sich in der Literatur viele Ansätze zur Visualisierung (Holten 2006, Herr et al. 2010). Aus diesen wurden für die Untersuchung eine Visualisierung als Wurzelbaum (siehe Abbildung 1) und eine Visualisierung als TreeMap (siehe Abbildung 2) ausgewählt. Diese beiden Darstellungsformen unterscheiden sich dahingehend, dass bei einer Visualisierung als Wurzelbaum eine frontale Ansicht auf die hierarchische Struktur erfolgt und bei der TreeMap eine Ansicht von oben. Die Darstellung als Wurzelbaum wurde um die Technik des Semantic Zoom (Spence 2001) erweitert. Bei einer niedrigen Zoomstufe wird dabei die gesamte Struktur des Ressourcenraumes auf dem Bildschirm angezeigt, damit der Nutzer sich einen Überblick verschaffen kann. Mit jeder Zoomstufe verfeinern sich die Ansicht und die dargestellten Informationen.

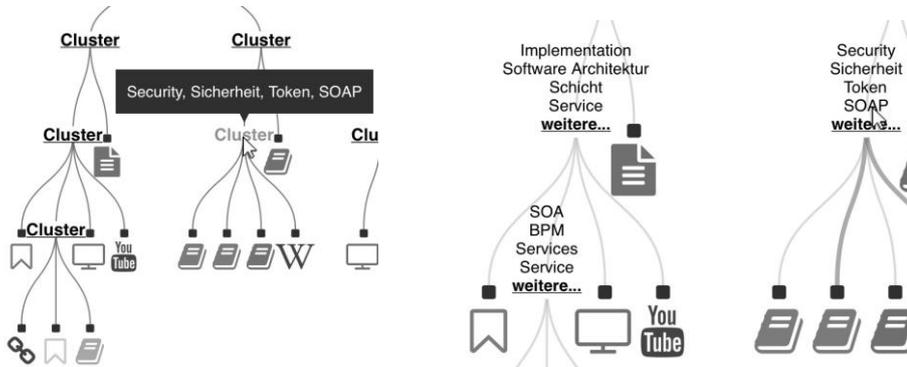


Abbildung 1: Ausschnitte aus der Visualisierung als Wurzelbaum mit einer niedrigen Zoom-Stufe (links) und einer hohen Zoom-Stufe (rechts).

Die Orientierungsunterstützung in Form von Stichwörtern in den inneren Knoten wird hier ausgeblendet und muss explizit aufgerufen werden. Beim Hineinzoomen in einen bestimmten Bereich der Struktur wird nur noch ein Teil-Ressourcenraum angezeigt und dafür die Stichwörter direkt in der Visualisierung ausgegeben.

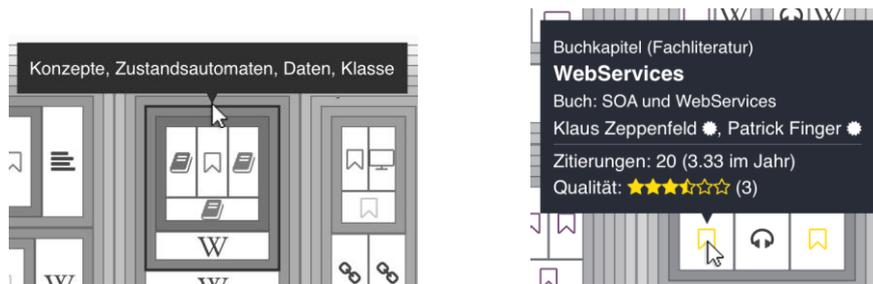


Abbildung 2: Ausschnitte aus der Visualisierung der TreeMap mit Darstellung der Informationsdimensionen.

Im Rahmen einer Laborstudie wurde untersucht, ob der hier vorgestellte Ansatz der Visualisierung einer Stichwörter-Hierarchie als Abhängigkeitsstruktur eines Ressourcenraumes zu einer effektiven und effizienten Orientierung in einem Ressourcenraum verhelfen kann. Für die Studie wurde ein Ressourcenraum zum Thema „Software Engineering“ zusammengestellt. Um ein möglichst realistisch und vergleichbares Szenario mit einer heute üblichen Internet-Recherche zu ermöglichen, wurden insgesamt 137 Ressourcen aus einem weiten Spektrum an digitalen Ressourcen unterschiedlichster Form und Art zur Verfügung zusammengestellt (68 Lehrbücher/Buchkapitel, 33 Wikipedia-Artikel, 19 Vorlesungsfolien, 10 Webseiten/Videos, 4 Podcasts/Skripts, 3 weitere Texte). Diese Ressourcen wurden durch Suchmaschine bzw. Literaturdatenbank gefunden und stellen eine etwa 80 Minuten Recherchearbeit zu dem Thema des Ressourcenraumes dar.

Die automatisierte Analyse und Strukturierung der Ressourcen wurde vor der Studie durch eine prototypische Implementation der in Abschnitt 3.1 aufgezeigten Transformationsschritte

der Visualisierungspipeline durchgeführt. Die Visualisierung der gewonnenen Strukturen erfolgte über das JavaScript-Framework „D3.js“.

4 Evaluation

4.1 Studiendesign

An der Studie nahmen insgesamt 44 Teilnehmer teil, die sich auf eine Ausschreibung über mehrere Universitäts-E-Mail-Verteiler und Social Networks gemeldet hatten. Die ausgewählten Teilnehmer, von denen kein thematisches Vorwissen verlangt wurde, wurden zufällig auf 4 Gruppen (zu je 11 Teilnehmern) verteilt. Es gab jeweils eine Testgruppe für die Visualisierung als Wurzelbaum bzw. als TreeMap, die auf derselben hierarchischen Struktur der vorgegebenen Ressourcen arbeiteten und identische Informationsdimensionen zur Verfügung hatten. Das prototypische System erlaubte es, den Ressourcenraum nach den analysierten Stichwörtern zu durchsuchen und Dateien nach Präferenzen (z. B. nur Lehrbücher) zu filtern.

Zudem gab es zwei Kontrollgruppen. Die erste Kontrollgruppe konnte auf die vorgegebenen Ressourcen mithilfe eines klassischen Dateimanagers zugreifen. Die Ressourcen waren nach ihrer Art (Wikipedia-Artikel, Lehrbuch etc.) in Ordner unterteilt. Die Teilnehmer dieser Gruppe konnten mithilfe einer Volltextsuche, welche alle Vorkommen der Suchwörter ungewichtet findet, die Ressourcensammlung durchsuchen. Der zweiten Kontrollgruppe stand das Web uneingeschränkt zur Verfügung. Um unterschiedliche Vorkenntnisse und Arbeitsweisen zu kompensieren, erhielten diese Teilnehmer eine Einführung in die Suche von wissenschaftlichen Publikationen bei Google Books und Scholar bzw. der Universitäts-Bibliothekssuchfunktion.

Aufgabe in der Studie war die Bearbeitung eines Multiple-Choice-Tests (im Folgenden als Haupttest bezeichnet), welcher 31 Fragen zu 7 Themenbereichen des Ressourcenraums beinhaltete. In jedem Bereich wiesen die Fragen unterschiedliche Schwierigkeitsgrade auf. Es wurde sichergestellt, dass sich die Antworten in unterschiedlichen Ressourcentypen finden ließen (z. B. sowohl in einem Wikipedia-Artikel als auch in einem Lehrbuch), als auch die Volltext- bzw. Stichwortsuche für jedes Signalwort in der Fragestellung ein Ergebnis liefert. Die Teilnehmer hatten 80 min. Zeit, um mit den Hilfsmitteln ihrer jeweiligen Testgruppe den Fragebogen zu bearbeiten. Neben dem Haupttest wurden ein Vorwissenstest und ein Nachwissenstest konzipiert, um den Wissensstand der Teilnehmer vor der Bearbeitung des Haupttests bzw. der Verwendung der zur Verfügung gestellten Ressourcen / des Internets und danach zu bestimmen. Beide Tests bestanden aus 7 Fragen aus dem Haupttest, die jedoch leicht umformuliert bzw. variiert wurden und ohne Hilfsmittel beantwortet werden mussten.

4.2 Hypothesen

Zur Untersuchung der Forschungsfrage wurden vor der Durchführung der Studie die folgenden drei Hypothesen aufgestellt:

H1: Relevante Ressourcen können durch eine strukturierte Visualisierung in Form eines Wurzelbaums oder einer TreeMap schneller und effizienter gefunden werden, so dass sich Nutzer mehr mit den für sie relevanten Inhalten auseinandersetzen und besser ein Verständnis für diese entwickeln können. Das führt letztlich in den Testgruppen „Wurzelbaum“ und „TreeMap“ zu einem größeren Unterschied zwischen Vor- und Nachtest als in den beiden Kontrollgruppen „Dateimanager“ und „Internet“.

H2: Die Visualisierung als Wurzelbaum schafft im Vergleich zur TreeMap einen besseren Überblick der hierarchischen Zusammenhänge und ermöglicht dadurch, einfacher relevante Ressourcen durch Navigieren in der Visualisierung zu finden. Daher müssen Nutzer in der „Wurzelbaum“-Testgruppe seltener die Suchfunktion nutzen als Nutzer in der „TreeMap“-Testgruppe.

H3: Steht keine Unterstützung in Form strukturierter Visualisierungen zur Verfügung, ist es für Nutzer aufwändiger, relevante Informationen zu finden. Die spiegelt sich in der Anzahl der verwendeten Ressourcen und der benötigten Zeit in den beiden Kontrollgruppen im Vergleich zu den beiden Testgruppen wider.

4.3 Ergebnisse

In Tabelle 1 findet sich eine Übersicht über die erzielten durchschnittlichen Ergebnisse der Test- und Kontrollgruppen für alle Fragebögen¹.

Test	Wurzelbaum	TreeMap	Dateimanager	Internet
Vorwissen	22,31 %	22,31 %	37,19 %	37,19 %
Haupt	23,86 %	41,62 %	35,8 %	55,4 %
Nachwissen	25,32 %	31,17 %	33,77 %	31,82 %
Vergleich zwischen Vor-/Nachwissen	p = 0,53 sd = 11,25 r = 0,18	p = 0,53 sd = 11,25 r = 0,18	p = 0,63 sd = 9,81 r = 0,13	p = 0,80 sd = 9,81 r = 0,07

Tabelle 1: Übersicht über die durchschnittlichen Ergebnisse des Vorwissens-, Haupt- und Nachwissenstests, sowie Vergleich der Vor- bzw. Nachwissenstests.

Eine erste Auffälligkeit zeigt sich im Vergleich des Haupttestergebnisses der beiden Testgruppen ($p = 0,003$; $sd = 15,22$; $r = 0,87$). Derselbe Unterschied ($p = 0,004$; $sd = 15,22$; $r = 0,85$) in diesen beiden Gruppen spiegelt sich ebenso im Vergleich der in der vorgegebenen Zeit bearbeiteten Aufgaben (siehe Tabelle 2) wider. Nur in der Kontrollgruppe „Internet“ waren Teilnehmer (4 von 11) vor den 80 min. mit der Beantwortung des Fragebogens fertig.

	Wurzelbaum	TreeMap	Dateimanager	Internet
bearbeitete Aufgaben	53,67 %	74,55 %	64,52 %	86,29 %

Tabelle 2: Übersicht über die durchschnittliche Anzahl an bearbeiteten Aufgaben im Haupttest.

¹ Signifikanz bei $p < 0,05$; Wilcoxon-Vorzeichen-Rang-Test innerhalb, Mann-Whitney-U-Test zwischen Gruppen.

Tabelle 3 vergleicht die durchschnittliche Anzahl an Interaktionen mit dem System. Im Vergleich der beiden Testgruppen zeigt sich eine höhere Interaktionsanzahl der Gruppe „TreeMap“ sowohl bei den Ressourcenaufrufen mit Dopplungen ($p = 0,39$; $sd = 15,22$; $r = 0,25$), der insgesamt genutzten unterschiedlichen Ressourcen ($p = 0,22$; $sd = 15,22$; $r = 0,36$) und Suchanfragen ($p = 0,29$; $sd = 15,22$; $r = 0,31$).

	Wurzelbaum	TreeMap	Dateimanager	Internet
Suchanfragen	21,45	28,09	37,09	61,72
Ressourcenaufufe	66,9	75,72	48	121,36
genutzte Ressourcen	27,09	32,63	28,18	68,63

Tabelle 3: Übersicht über die durchschnittliche Anzahl an getätigten Suchanfragen, Ressourcenaufufe gesamt (mit Dopplungen) und genutzten unterschiedlichen Ressourcen.

Zwar wurden in der Testgruppe „Wurzelbaum“ im Durchschnitt weniger Ressourcen genutzt, allerdings ist hier die durchschnittlich genutzte Zeit mit diesen im Vergleich ($p = 0,048$; $sd = 15,22$; $r = 0,61$) zur Testgruppe „TreeMap“ höher (siehe Tabelle 4).

genutzte Zeit mit	Wurzelbaum	TreeMap	Dateimanager	Internet
Ressourcen	70,29 %	61,83 %	70,76 %	91,4 %
System	29,71 %	38,17 %	29,24 %	8,6 %

Tabelle 4: Genutzte Zeit mit dem System (bei der Internet-Kontrollgruppe mit einer Suchmaschine) und der Betrachtung von Ressourcen.

4.4 Interpretation

Um die ursprüngliche Forschungsfrage zu klären, wurde die Hypothese (H1) aufgestellt, dass die Orientierung anhand einer visuellen Aufbereitung einer hierarchischen Struktur von Stichwörtern eine Verbesserung zur Erfassung der relevanten Ressourcen eines Ressourcenraumes ist. In den Ergebnissen ist zu sehen, dass es in den Testgruppen tatsächlich zu einer Zunahme zwischen dem Vor- und Nachwissenstest gekommen ist, während in den Kontrollgruppen sogar eine Abnahme festgestellt werden konnte. Dies spricht tendenziell für die Hypothese H1, da ein Anstieg auf ein gefestigteres mentales Modell über das Verständnis der Thematik hindeuten kann und eine geringere Informationsüberflutung stattfand.

Die geringere Anzahl an gestellten Suchanfragen in der „Wurzelbaum“-Testgruppe, als in der „TreeMap“-Testgruppe bestätigt tendenziell Hypothese H2. Allerdings zeigt sich im direkten Vergleich der Testgruppen, dass für die Teilnehmer mit der TreeMap einen besseren Zugang zu dem Ressourcenraum ermöglicht werden konnte als mit der Darstellung als Wurzelbaum. Die Nutzer der TreeMap kamen fast an die Ergebnisse der Internet-Benutzer heran, während die Nutzer der Visualisierung als Wurzelbaum hinsichtlich Ergebnis und bearbeiteten Aufgaben im Haupttest im Vergleich zu allen anderen Gruppen schlechter abschnitten. Das Ergebnis der Nutzer der TreeMap verstärkt sich dadurch, dass diese zwar 9% mehr Zeit mit dem System verbracht haben als Nutzer der anderen Testbedingung, aber scheinbar gezielter die relevanten Inhalte gefunden wurden.

Hypothese H3 hingegen kann nicht bestätigt werden. Zwar gibt es in der Kontrollgruppe „Internet“ eine deutlich höhere Anzahl an genutzten Ressourcen und Aufrufe, aber die Kontrollgruppe „Dateimanager“ nutzte weniger als die Testgruppen „Wurzelbaum“ und „TreeMap“. Ebenso ist der Unterschied der genutzten Zeit mit den Ressourcen zwischen der Kontrollgruppe „Dateimanager“ und der Testgruppe „Wurzelbaum“ nur minimal. Die Teilnehmer der beiden Testgruppen haben weniger Suchaktionen durchgeführt, dafür aber mehr Ressourcen genutzt. Die Kontrollgruppe „Dateimanager“ hat im Vergleich durchschnittlich ungefähr dieselbe Anzahl an Ressourcen genutzt, aber wesentlich häufiger die Suchfunktion. Durch die gewählten Visualisierungs- und Interaktionsmöglichkeiten konnten sich die Teilnehmer der Testgruppen damit möglicherweise mehr mit den konkreten Zusammenhängen der Ressourcen auseinandergesetzten, als sich nur durch die Suchergebnisse leiten zu lassen. Dies zeigt sich auch darin, dass es bei diesen Teilnehmern eine Zunahme zwischen Vor- und Nachwissenstest zu verzeichnen war. Dies ist ein weiteres Indiz, dass die Nutzer ein gefestigteres mentales Modell über das Verständnis der Thematik erlangt haben und keine bzw. eine geringere Informationsüberflutung stattfand.

5 Fazit und Ausblick

Die Ergebnisse der Evaluationsstudie unterstreichen die Problemstellung bei der Verwendung umfangreicher Sammlungen digitaler Ressourcen: Die Internet-Kontrollgruppe kam in nur 80 Minuten im Durchschnitt auf rund 69 genutzte Ressourcen. Dies zeigt, dass ein Nutzer bei einer heutzutage üblichen Informationsrecherche im Internet tatsächlich mit einer größeren Menge an Ressourcen konfrontiert ist, die er analysieren und strukturieren muss. Die Studienauswertung zeigt, dass diese Art der Faktensuche durch viele Suchanfragen und Ressourcenaufrufe zwar sehr viele aktive Interaktionen benötigt, aber gleichzeitig effizient ist: Die Internet-Kontrollgruppe hat das höchste Durchschnittstestergebnis und die meisten Aufgaben gelöst. Ebenso zeigt sich in der Kontrollgruppe „Dateimanager“, dass zwar die Testergebnisse relativ gut waren, aber auch hier eine intensivere Nutzung der Suchfunktion nötig war. Allerdings scheint diese Art der Recherche über eine Volltextsuche nicht zum Aufbau eines Verständnisses über die Thematik zu verhelfen. Bei beiden Kontrollgruppen gab es eine Abnahme zwischen dem Vorwissens- und Nachwissenstest. Eine mögliche Erklärung ist, dass es in diesen Gruppen zu einer Informationsüberflutung kam, durch die sich kein Verständnis über die Zusammenhänge der Inhalte der Ressourcen verfestigen konnte.

Das Konzept eines Ressourcenraums mit Ressourcen, welche sehr viele Informationsdimensionen besitzen, nicht direkt miteinander zu verbinden, sondern eine zusätzliche hierarchische Stichwort-Struktur aufzubauen wird, kann die Informationsüberflutung reduzieren. In der Visualisierung des Ressourcenraums wurden damit die vielen Informationen auf genau die relevanten reduziert, die für eine effiziente Orientierung notwendig sind. Damit kann insgesamt eine Forschungsfrage positiv beantwortet werden.

Die Abweichungen in den Ergebnissen der beiden Visualisierungen in der Studie wären durch eine granulare Untersuchung zu überprüfen. Neben der inhaltlichen Dimension könnte weiterhin untersucht werden, wie effizient die Orientierungsmechanismen sind, wenn weitere Dimensionen (Qualitätsbewertungen, Benutzerpräferenzen) eine Relevanz für die Auswahl der Ressourcen darstellen oder größere Ressourcenräume zur Verfügung stünden.

Literaturverzeichnis

- Aggarwal, C. C., & Zhai, C. (2012). A survey of text clustering algorithms. In *Mining Text Data* (pp. 77-128). Springer US.
- Alpert, S. R. (2005). Comprehensive mapping of knowledge and information resources: the case of Webster. In *Knowledge and information visualization* (pp. 220-237). Springer Berlin Heidelberg.
- Cañas, A. J., & Novak, J. D. (2014). Concept mapping using CmapTools to enhance meaningful learning. In *Knowledge cartography* (pp. 23-45). Springer London.
- Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann.
- Chen, N. S., Wei, C. W., & Chen, H. J. (2008). Mining e-Learning domain concept map from academic articles. *Computers & Education*, 50(3), 1009-1021.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval* (Vol. 1, No. 1, p. 496). Cambridge: Cambridge university press.
- Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of educational research*, 76(3), 413-448.
- Heer, J., Bostock, M., & Ogievetsky, V. (2010). A tour through the visualization zoo. *Commun. Acm*, 53(6), 59-67.
- Holten, D. (2006). Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *Transactions on Visualization and Computer Graphics*, 12(5), 741-748.
- Hsieh, T. C., & Wang, T. I. (2010). A mining-based approach on discovering courses pattern for constructing suitable learning path. *Expert systems with applications*, 37(6), 4156-4167.
- Huang, A. (2008, April). Similarity measures for text document clustering. In *Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008)*, Christchurch, New Zealand (pp. 49-56).
- Rajaraman, A., & Ullman, J. D. (2012). *Mining of massive datasets* (Vol. 1). Cambridge: Cambridge University Press.
- Spence, R. (2001). *Information visualization* (Vol. 1). New York: Addison-Wesley.
- Zeiliger, R., & Esnault, L. (2014). The constructivist mapping of internet information at work with Nestor. In *Knowledge cartography* (pp. 89-111). Springer London.