M. Herczeg & M. C. Kindsmüller (Hrsg.): Mensch & Computer 2008: Viel Mehr Interaktion.
München: Oldenbourg Verlag, 2008, S. 27-36

Ein Web-Service zur schnellen Erfassung graphisch repräsentierter Informationen durch visuell Beeinträchtigte

Gerhard Plaßmann, Martin Dilg

Fachhochschule Köln, Institut für Informatik, Labor für Medieninformatik

Zusammenfassung

Formatierungen gewichten die Informationsinhalte von *web-sites* und bieten den Rezipienten daher eine Unterstützung in der Interaktion (Selektion, nonlineare Rezeption, Navigation, Orientierung, etc.). Formatierungsinformationen sind visuell beeinträchtigten Rezipienten (etwa geburtsblinden Menschen) jedoch nicht a priori zugänglich. Das Ziel dieses Projektes war die Analyse von Struktur und Inhalt hypermedial-repräsentierter Informationsräume aus verschiedenen Perspektiven und die Implementierung des Paradigmas des visuellen „Überfliegens" graphisch repräsentierter Informationen zur Unterstützung von visuell beeinträchtigten Menschen (oder Maschinen) über einen als *web-service* realisierten Agenten. Zudem bietet der *web-service* visuell Beeinträchtigten eine Basis, auf der nonlineare Selektion und Rezeption in hypertext-basierten Informationsräumen stattfinden kann.

1 Einleitung

Die graphische Gestaltung einer *web-site* dient vor allem der Unterstützung der visuellen Kommunikation und stellt eine wichtige Basis für die Interaktion der Benutzer mit den Informations-Inhalten dar. Überlegungen zu den Kommunikationszielen, dem Design, der *usability* einer *web-site* etc. führen zu gedanklichen Vorstellungen ihrer Produzenten (*conceptual model*), und in Verbindung mit den Informations-Kodierungen (Auszeichnungs-, Skript- oder Programmiersprachen) zu dem Produktionsmodell für die *web-site*. Dieses Produktionsmodell enthält explizit und oftmals auch lediglich implizit viele für die Benutzer wichtige Informationen, die sie bei der Rezeption und Interaktion mit einer *web-site* unterstützen könn(t)en. Die visuellen Ausprägungen der kodierten Informationen spiegeln u. a. auch eine Gewichtung der Inhalte wider. Zusätzlich enthalten die Dokumente innerhalb einer

web-site Informationen über deren Struktur und könn(t)en die Benutzer dadurch beim Navigieren, Orientieren, bei Selektions- und Rezeptionsprozessen unterstützen.

Neben den primären visuellen Daten, den Texten und Bildern enthält das Produktionsmodell auch sekundäre visuelle Daten, die das Design und Layout eines Dokumentes ausmachen (z.B. *font weight, -size* und *color*). Die Qualität der Nutzung der primären Daten wird für einen Benutzer durch die Qualität der sekundären Daten mitbestimmt. Da sich das Produktionsmodell vor allem in den visuellen Ausprägungen der Dokumente widerspiegelt, kann es von Benutzern, die diese nicht wahrnehmen können, nicht genutzt werden (z.B. von Menschen mit Beeinträchtigung ihrer Sehfähigkeit oder von Maschinen).

Es existieren bereits Multi-Agenten basierte Ansätze (siehe z.B. Moukas & Maes 1998), die benutzer-adaptive und informations-filternde Systemkomponenten beinhalten. Andere Autoren (siehe Kobsa 1999) berücksichtigen visuell beeinträchtigte Benutzer durch multimodale *Interfaces* und zusätzliche Informationen (wie z.B. *Link*-Listen). Den Ansatz der Analyse von Strukturen der baumähnlich aufgebauten (X)HTML-Dokumente verfolgen etwa Walshe & McMullin 2004.

Für visuell nicht beeinträchtigte Benutzer könnte eine maschinelle Auswertung des Produktionsmodells auch eine Interaktionsunterstützung in Form eines als *web-services* realisierten Assistenzsystems bringen. Die Entwicklung eines Prototypen in Form eines *web-services* ist durch die Idee motiviert, Annahmen über das zugrunde liegende Produktionsmodell aus den einzelnen Dokumenten einer *web-site* zu treffen, und den Benutzern priorisierte Informationen am *Interface* für weitere Interaktionen zugänglich zu machen. Vornehmlich verfolgt die Entwicklung unseres Prototyps aber das Ziel, visuell beeinträchtigten Benutzern solche Informationen zugänglich(er) zu machen, wie sie visuell Unbeeinträchtigten in einer präattentiven Phasen der visuellen Perzeption zugänglich werden. Wie effizient visuell Unbeeinträchtigte in dieser Phase Informationen erfassen können lässt sich am Beispiel des „Werfens eines flüchtigen Blicks“ auf die erste Seite einer Tageszeitung verdeutlichen: Strukturen (z.B. *Layout*-Raster), Priorisierungen (Größen, Platzierung, ggf. Farbe, etc.) und Bedeutungen werden mental repräsentiert und dienen in der attentiven Phase dann als Ausgangspunkt für ggf. detaillierte Perzeptionsprozesse (etwa dem Lesen eines bestimmten Beitrages).

Die Implementierung dieses *web-services* stellt u. a. folgende Anforderungen, die in den Abschnitten zwei und drei dieses Beitrags behandelt werden:

- Entwicklung eines Modells, mit Hilfe dessen die aus den Dokumenten einer *web-site* gewonnenen Daten verwaltet werden können und das die Möglichkeit der Gewichtung der primären Daten durch die sekundären Daten bietet,
- Entwicklung einer Systemarchitektur, die eine effektive und effiziente Nutzung des entwickelten Modells ermöglicht.

2 Modellierung

Für die Analyse, der Verarbeitung und Aufbereitung der Inhaltsinformationen sollte ein geeignetes Modell entwickelt werden. Dazu versucht die Modellierung, die Perspektiven

- visueller menschlicher Informationsverarbeitung und der
- Bedeutungsrelevanz der Informationen einzunehmen und zu berücksichtigen.

Für die Perspektive der menschlichen Informationsverarbeitung visueller Reize sind die *Feature Integration Theory* (Treisman & Gelade 1980) und die Rahmenbedingungen menschlicher Gedächtnisleistung des *working memory* (Baddeley 2000) bzgl. Speicherkapazität (7 ± 2 *chunks*) modellbildend. Treisman und Gelade entwickelten 1980 ein Modell der visuellen Aufmerksamkeit, das zwei verschiedene Verarbeitungsstufen unterscheidet:

- zunächst werden in einer *präattentiven Phase*[1] alle Merkmale einer vorliegenden visuellen Szene analysiert und entsprechende *feature maps* erstellt (z.B. elementare Merkmale wie Farbe, Krümmung, Textur etc.) und
- in der sich daran anschließenden *attentiven Phase* werden diese Merkmale miteinander in der *master map of image* verknüpft (*binding*). Über *bottom-up* Verarbeitung können auf Basis der Elementar-Merkmale Objektwahrnehmungen entstehen oder und *top-down* Verarbeitung Objekte wieder erkannt werden. Die gerichtete visuelle Aufmerksamkeit ist als Metapher eines *spot-lights* auf der *master map of image* konzipiert.

Zudem sind aus Experimenten so genannte *PopOut-Effekte* bekannt, welche die visuelle Aufmerksamkeit auf sich ziehen können. Diese Effekte können entstehen, wenn ein Objekt (oder ein Objektverband) sich in (mindestens) einem *feature* der *feature map* signifikant von seinen Nachbarschafts-Objekten unterscheidet.

Neben der grundsätzlichen Frage der Bedeutungsrelevanz der Informationen erfordert auch die Auseinandersetzung mit der Speicherdauer im Kurzzeitgedächtnis die Auseinandersetzung mit Relevanzaspekten der Informationen; zudem macht die Entwicklung und Verbreitung von „*web 2.0*“ eine in diese Richtung migrierbare Modellierung interessant.

Für einen ersten Prototypen wird die Relevanz von Inhaltsinformationen über eine lediglich statistische Betrachtung durch die Analyse der Textelemente mittels TF*IDF Algorithmus (siehe Salton & Buckley 1988) modelliert. Für weitere Prototypen besteht hier sicherlich Handlungsbedarf bezüglich einer Verbesserung der Modellierung. Als weitere Perspektive wird die menschliche Gedächtnisleistung des *working memory* berücksichtigt.

[1] die Analyse in der präattentiven Phase verläuft automatisch und unbewusst und dauert ca. 160 ms

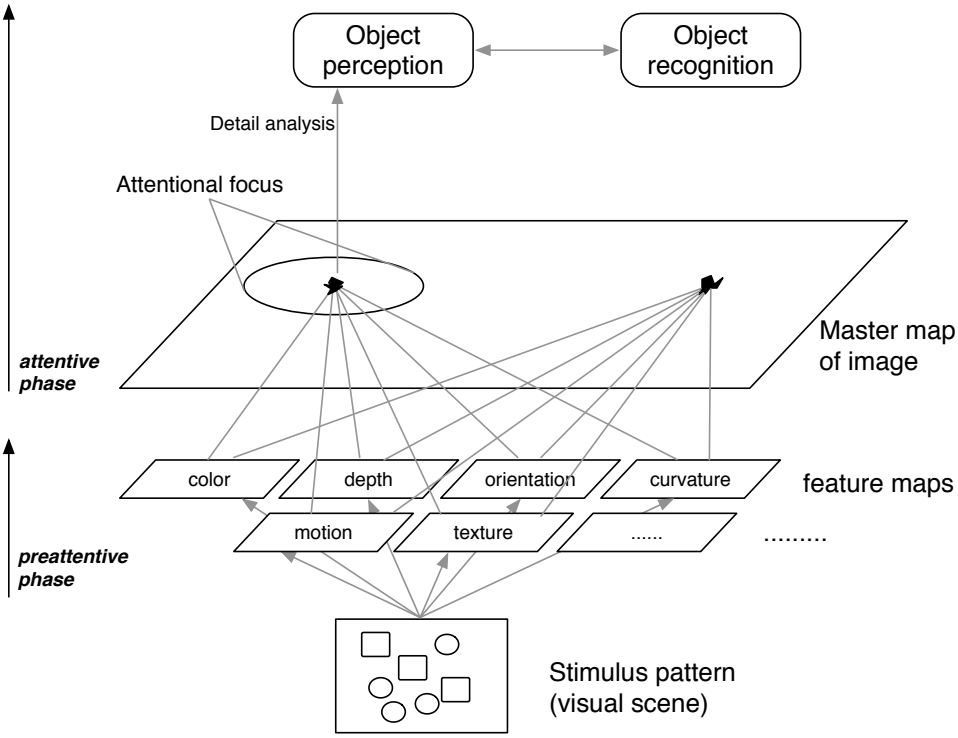

Abbildung 1: Modell der visuellen Aufmerksamkeit nach Treisman & Gelade (1980)

2.1 Modellierungsperspektiven

Die genannten Perspektiven münden in ein Analyse-Modell der Dokumente einer *web-site*; das Analyse-Modell ist in Abb. 2 skizziert (*DocumentElements* repräsentieren dabei Inhaltselemente wie z.B. Navigationselemente der *web-site*).

Die Abbildung 2 zeigt die Hierarchie und die Beziehungen der einzelnen für die Gewichtung genutzten Elemente; den Elementen sind die Attribute visueller (*off-set, color, fontweight, fontsize*) und organisatorischer Eigenschaften (*fqdn = fully qualified domain name, xpath, type*) der *web-site* Elemente zugeordnet.

2.2 Gewichtung

Das in diesem Abschnitt entwickelte Analyse-Modell bildet den Netzcharakter eines Hypertext-Systems durch die Elemente *Document* und *Hyperlink* ab. Jedes Dokument einer *web-site* und jeder *Hyperlink* zwischen einzelnen Dokumenten kann durch sie repräsentiert werden.

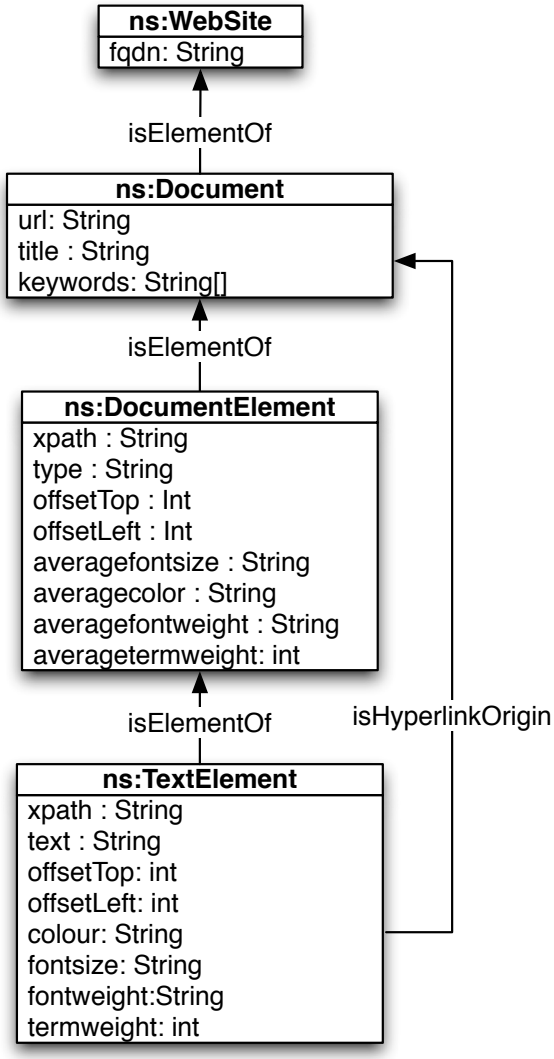

Abbildung 2: Übersicht über das Analyse-Modell der Dokumente einer web-site

Zusätzlich bildet das Analyse-Modell mit den Elementen *DocumentElement* und *TextElement* die Ausprägungen der einzelnen Dokumente ab. Ihre Attribute ermöglichen eine Gewichtung entsprechend visueller und statistischer Merkmale einzelner Elemente. Diese Merkmale stellen die sekundären Daten dar, mit denen die primären Daten gewichtet werden können. Dabei berücksichtigt das Modell sowohl die *Feature Integration Theory* nach Treisman & Gelade als auch die Speicherkapazität des Arbeitsgedächtnismodells nach Baddeley und die statistischen Bewertungen der Textelemente.

3 Architektur des Web-Service

Die zentrale Komponente der Architektur des *web-service* (Abb. 3) ist das Analyse-Modell. Es nimmt Daten auf, die von Komponenten (Extraktoren E1 – En) durch Analyse aus einer *web-site* und ihren Dokumenten gewonnen werden und dazu benutzt werden, die Elemente des Modells aufzubauen und mit Daten zu befüllen. Weitere Komponenten (M1 – Mn) übernehmen die Auswertung der im Modell enthalten Daten, deren Gewichtung und die Ausgabe gewichteter Daten an den Benutzer. Die Auswertung der Daten im Modell erfolgt dabei über Abfragen an das Modell, die eine Gewichtung der vom Modell zurückgegebenen Daten bewirkt.

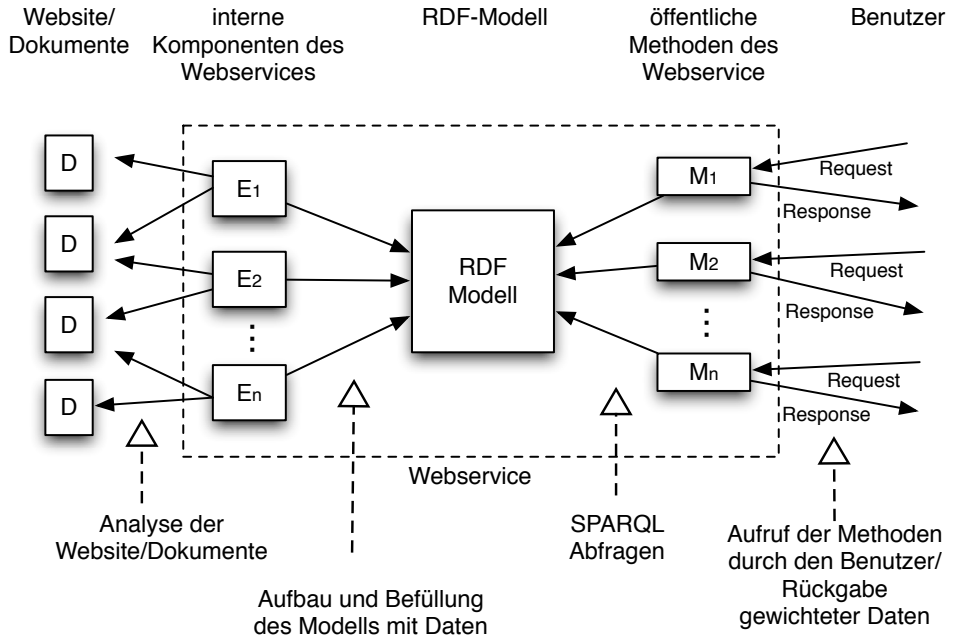

Abbildung 3: Komponenten der Systemarchitektur

Die vorgestellt Systemarchitektur wird als *web-service* unter Nutzung eines RDF-Modells und der SPARQL *Query-Language* umgesetzt. Abbildung 4 illustriert die eingesetzten Technologien und ihr Zusammenwirken. Ein Client fordert die Definition des öffentlichen Interfaces des *web-services* an und führt eine der darin definierten öffentlichen Methoden aus, indem er einen SOAP-*request* an den *web-service* schickt. Für die Implementierung des „semantischen Modells“ wurde RDF und für seine Abfrage SPARQL benutzt. Dazu wird das Jena-*Toolkit* verwendet, das eine Java-basierte Implementierung der genannten Technologien ermöglicht. Für die Implementierung des *web-services* wird die SOAP Implementierung von Apache Axis2 benutzt.

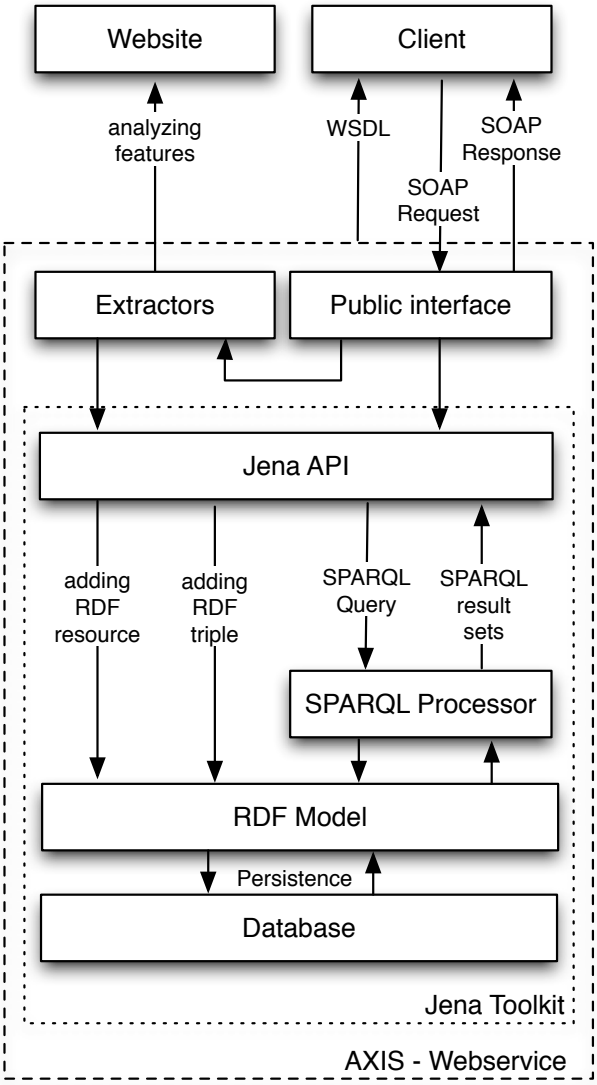

Abbildung 4: Technologie-Sicht auf den web-service

Die grundlegenden Funktionen der Extraktoren E1 – En sind das Extrahieren von Daten aus einer *web-site* und deren Dokumenten und das Integrieren dieser Daten in das RDF-Modell. Dies umfasst sowohl das Anlegen neuer *Resources* unter Verwendung des RDF Vokabulars, als auch das Zuweisen der Werte der *Properties* dieser *Resources*. Abbildung 3 veranschaulicht die einzelnen Extraktoren. Die Extraktoren sind Komponenten *des web-services*, die derzeit nicht öffentlich zugänglich sind.

Die Komponenten M1, M2 und M3 sind öffentliche Methoden des *web-service*. Die Komponente M1 bestimmt die Elemente einer *web-site*, die einen *PopOut-Effekt* erzeugen können. Dazu wird das DOM Modell, die *Tag*-Hierarchien und ggf. vorhandene CSS-Angaben herangezogen, heuristisch[2] bewertet und die *Container-Tags* der gefundenen Elemente bestimmt.

Die Komponente M2 durchläuft das RDF-Modell auswertend, indem es die Elemente mit dem höchsten statistischen Gewicht ermittelt. Entsprechend der Kapazität des Arbeitsgedächtnisses[3] wird eine geordnete Liste zurückgegeben.

Die Komponente M3 verwendet die Komponenten M1 und M2 und priorisiert die Ergebnis-Elemente entsprechend der graphischen Repräsentation der *web-site*. Kriterien sind hier links-rechts und oben-unten Prioritäten (Leserichtung in westlich orientierten Kulturen) sowie Prioritäten entsprechend der Navigations-Hierarchie in der *web-site*.

4 Evaluation

Die wesentlichen Komponenten des vorgestellten Systems wurden in einem Prototyp realisiert. Das Zusammenwirken der einzelnen Komponenten sollte überprüft und die Aussagekraft der Rückgabewerte der öffentlichen Funktionen auf Plausibilität getestet werden.

Dazu wurden fünf *Web*-Seiten unter dem URL http://www.tagesschau.de ausgewählt. An dem Test des *web-service* nahmen 23 Personen[4] teil. Zunächst wurden die Personen gebeten, die entsprechenden *Web*-Seiten zu „überfliegen“. Aufgrund der Länge der *Web*-Seiten wurde den Personen die Art und Weise des „*Scrollen*s“ durch die Seite frei gestellt (z.B. mittels Pfeiltasten, *mouse*). Zudem wurden die Personen darüber informiert, dass sie anschließend gebeten werden, die „Themen“ zu benennen, die für sie „von Interesse“ gewesen seien. Die mittlere Zeitdauer, welche die Personen für das „Überfliegen“ aufwendeten, betrug 19,8 Sekunden (Standardabweichung: 2,4 Sekunden). In 8 Fällen nannten die Personen 4 und in 3 Fällen 2 Themen, die sie interessierten. In allen anderen (104) Fällen nannten die Personen 3 Themen. Alle Nennungen erfolgten in Form von Substantiven (hier: Schlagworte aus dem Nachrichtenbereich). In 83 % aller Fälle erfasste der zu testende *web-service* die von den Personen genannten Themen und repräsentierte sie am *Interface*. In 3 % aller Fälle waren die Themen erfasst, aufgrund einer geringen Priorisierung jedoch nicht mehr am *Interface* repräsentiert[5]. In 14 % der Fälle lagen die Themen außerhalb des Hauptinhaltselements, so dass der derzeitige Prototyp des *web-service* diese nicht erfasst.

---

[2] die Heuristik orientiert sich an den Kriterien, die Anne Treisman für paralleles Suchen gefunden hat

[3] 7 ± 2 *chunks*

[4] die Altersspanne betrug 17-56 Jahre, 9 der Tester waren weiblichen und 14 waren männlichen Geschlechts, unterschiedliche Bildungsgrade waren vertreten

[5] derzeit ist die Ergebnisliste des *interface* auf maximal 9 Elemente (7 ± 2 *chunks*) beschränkt

Dieser erste Test zeigt, dass wichtige Phänomene der visuellen Perzeption wie etwa *PopOut*-Effekte von *Textonen* oder etwa die perzeptive Priorität von Überschriften oder von *bold face* gesetzten *strings* mit dieser Modellierung erfasst und beschrieben werden können. Das bedeutet, dass die unterschiedlichen Extraktoren visuell und perzeptiv relevante Merkmale der Dokumente extrahieren, das RDF-Modell entsprechend aufbauen und die vorgestellten SPARQL-*Statements* und öffentlichen Methoden zu nachvollziehbaren Ergebnissen führen. Die derzeitige Beschränkung der Analyse auf das Hauptinhaltselementes erweist sich als Schwäche. Zudem erscheint im Licht dieser Testergebnisse (im Mittel 3 genannte Themen pro *Web*-Seite) die Länge der Ergebnisliste mit 9 Elementen als zu lang. Für eine adäquate Reduzierung der Listenlänge wären jedoch dezidierte Informationen über Benutzerpräferenzen notwendig.

5 Fazit und Ausblick

Die Aufgabenstellung dieses Projektes war die Analyse von Struktur und Inhalt hypermedial-repräsentierter Informationsräume aus verschiedenen Perspektiven mit dem Ziel der Implementierung des Paradigmas des visuellen „Überfliegens“ graphisch repräsentierter Informationen und der Unterstützung visuell beeinträchtigten Menschen oder Maschinen über einen als *web-service* realisierten Agenten. Dazu sollten (der eigentlichen Rezeption der Inhalte vorgelagerte) Repräsentation am *user interface* angeboten werden, die eher geeignet sind, nonlineare Selektions- und Rezeptionsprozesse zu unterstützen (als dies etwa mit linearen *Link*-Listen bei derzeitigen Assistiven Technologien[6] der Fall ist). Diese Zielsetzung ist mit dem jetzt vorliegenden ersten Prototypen in Grundzügen erreicht; der Erfüllungsgrad kann jedoch noch deutlich gesteigert werden.

In der sich daran anschließenden Redesign-Phase soll die Modellierung der Analyse-Perspektiven weiter ausgeführt und die Anzahl und Funktionalität der öffentlichen Methoden erweitert sowie die Beschränkung auf das Hauptinhaltselement aufgehoben werden. Die Berücksichtigung und Modellierung benutzerbezogener Aspekte (Interessenslagen, Präferenzen, kognitive Stile etc.) wäre wünschenswert. Weitere für die Wahrnehmung sehr relevante Informationen könnten aus der Analyse bildbasierter Informationen (z.B. *color segementation, motion estimation*) ermittelt werden. Aufgrund der Laufzeiten derartiger Algorithmen wurde im ersten Prototypen (in dem ein möglichst latenzfreie Verarbeitung wünschenswert war) auf deren Verwendung verzichtet. Zudem wäre es sinnvoll, die aus der Leserichtung abgeleitete Priorisierung benutzer-adaptiv (z.B. kulturabhängig) zu modellieren.

[6] z.B. screen-reader oder Braille-Zeile

Literaturverzeichnis

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences. Vol. 4* (11), S. 417-423.

Salton, G. & Buckley, C. (1988). Term Weighting Approaches in Automatic Text Retrieval. *Information Processing & Management. Vol. 24*., No 5., S. 513-523.

Treisman, A. M. & Gelade, G. (1980). A feature-integration theory of attention. In: *Cognitive Psychology 12*(1), S. 97-136.

Kobsa, A. (1999). Adapting Web Information to Disabled and Elderly Users. *WebNet-99, Honolulu, HI*, S. 32-37.

Moukas, A. & Maes, P. (1998). Amalthaea: an evolving multi-agent information filtering and discovery systems for the WWW. *Autonomous agents and multi-agent systems. Vol. 1*, S. 59-88.

Walshe, E. & McMullin B. (2004)- Accessing web based documents through a tree structural interface. In: Klaus, J., Miesenberger, K., Zagler, W.L. & Burger, D. (Hrsg.): *Computers Helping People with Special Needs*, *9th International Conference, ICCHP 2004*, Heidelberg, Springer, S. 560-563.

Kontaktinformationen

Martin Dilg
Pixelpark AG
Friesenplatz 25
D-50672 Köln
Tel +49.221.951515-48
Email: martin.dilg@pixelpark.com

Gerhard Plassmann
Institut für Informatik
Labor für Medieninformatik
Fachhochschule Köln, Campus Gummersbach
Steinmüllerallee 1
D-51643 Gummersbach
Tel.: +49 2261-8196 0
E-Mail: plassmann@gm.fh-koeln.de