

SmartWeb

Ein multimodales Dialogsystem für das semantische Web

Wolfgang Wahlster

**Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, DFKI
Saarbrücken, Kaiserslautern, Bremen, Berlin**

Email: wahlster@dfki.de

Internet: www.dfki.de/~wahlster

1. Einleitung

SmartWeb wurde als kontextsensitives Dialogsystem konzipiert [3], das den mobilen Benutzer in verschiedenen Rollen überall und jederzeit unterstützt, so dass ein ubiquitäres Zugangssystem zum semantischen Web entsteht. Dabei wurde vom SmartWeb-Konsortium zunächst als Basisversion eine Variante für mobile Endgeräte der MDA-Serie (u.a. MDA Pro und Ameo) der Deutschen Telekom entwickelt, die dem Fußgänger unterwegs oder dem Sportfan im Stadium den multimodalen Zugang zu Webdiensten ermöglicht. Die seit 2002 von T-Mobile vermarktete Serie von Endgeräten zur „Mobilen Digitalen Assistenz“ (MDA) integriert die vertrauten Funktionen moderner Taschencomputer (PDA) unter dem Betriebssystem Windows Mobile mit einem breitbandigen Internetzugang über UMTS, HSDPA oder WLAN.

Zusammen mit den Partnern DaimlerChrysler, Siemens und dem FhG-Institut First entwickelten wir eine Einbauversion von SmartWeb für Autofahrer, die für die A-Klasse und die R-Klasse von Mercedes verfügbar ist. Als Weltneuheit wurde erstmals auch ein Sprachdialogsystem für Motorradfahrer entwickelt, die zusammen mit der BMW-Forschung als weitere Einbauversion von SmartWeb auf den BMW-Motorrädern K 1200 LT und R 1200 RT erprobt wurde. Über alle Systemversionen hinweg bietet SmartWeb seinem Benutzer einen symmetrischen multimodalen Zugriff [8], der auf der Eingabeseite neben gesprochener und geschriebener Sprache, Gestik, Haptik und Video kombiniert mit Sprache, Haptik, Graphik, Video und Ton auf der Ausgabeseite. Der Benutzer bekommt die Informationen also nicht nur durch Eintippen von Fragen auf der kleinen Tastatur des PDAs, sondern auch durch gesprochene Sprache z.B. kombiniert mit Stifteingaben auf dem Bildschirm.

Es wird also möglich, das Mobiltelefon in Alltagssprache zu 'fragen' und dieses wird mit Informationen aus dem Netz 'antworten' (vgl. Fig. 1). SmartWeb baut auf das Vorgängerprojekt SmartKom (1999 – 2003) auf, das erstmals das Prinzip der symmetrischen Multimodalität realisierte [7], allerdings im Gegensatz zu SmartWeb keinen freien Zugang zum offenen Web erlaubte, sondern ausschließlich in vorher festgelegten Diskursbereichen funktionierte.



Fig. 1: SmartWeb als mobile Antwortmaschine über die FIFA WM 2006

2. Von syntaktischen Suchmaschinen zu semantischen Antwortmaschinen

SmartWeb geht klar über traditionelle Suchmaschinen wie Google hinaus, indem es nicht nur die Eingabe von Schlüsselwörtern sondern komplexe natürlichsprachlich formulierte Anfragen zulässt und Suchergebnisse höherer Qualität liefert, die zudem an den aktuellen Kontext und die jeweilige Aufgabenstellung des mobilen Benutzers angepasst sind. In mobilen Anwendungssituationen will der Benutzer als Ausgabe keine Listen von Verweisen auf möglicherweise relevante Webdokumente, die er dann wiederum weiter durchsuchen muss, sondern eine konzise Antwort auf seine Anfrage. So will der Familienvater während einer Autofahrt, wenn das Kind auf dem Rücksitz plötzlich über starke Schmerzen klagt, auf seine Frage „Wo ist der nächste Kinderarzt?“ von SmartWeb nur die Antwort „Dr. Pfeifer in der Marienstrasse 23“, so dass er als Folgeauftrag SmartWeb bitten kann „Dann zeige mir die Route und führe mich dort hin“, um eine Navigationsunterstützung zum nächsten Kinderarzt zu erhalten.

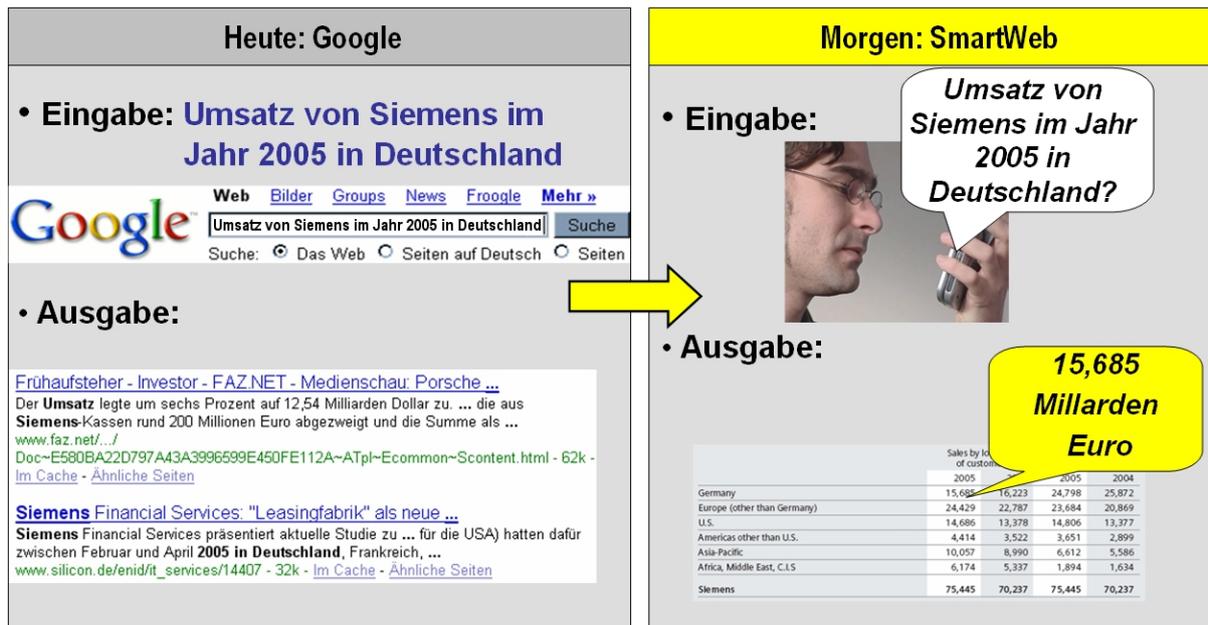


Fig. 2: Stichwortbasierte Suche versus Fragebeantwortung

SmartWeb zeigt damit den Weg auf, wie folgende Defizite bisheriger Suchmaschinen überwunden werden können:

- die Anfrage des Benutzers wird nicht inhaltlich verstanden, sondern nur als Zeichenkette betrachtet.
- jede Anfrage wird isoliert beantwortet, ohne den Inhalt vorangegangener Anfragen oder Suchergebnisse und damit den Dialogkontext zu berücksichtigen.
- Suchanfragen, die sich auf mehrstellige Relationen zwischen Begriffen beziehen und komplexe Restriktionen enthalten, führen zu unbefriedigenden oder sogar falschen Resultaten.
- „Versteckte Inhalte“, die im sog. „Deep Web“ nur über Portale, Webdienste oder in PDF-Dokumenten zugänglich sind, werden nicht gefunden.

Nutzer traditioneller Web-Suchmaschinen klagen wegen dieser Defizite oftmals über irrelevante Ergebnisse und einen zu umfangreichen Ergebnisraum. Sie verlassen die Suchmaschine ohne eine Antwort auf ihre Ursprungsfrage erhalten zu haben und versuchen weitere Recherchen mit anderen Suchmaschinen. SmartWeb bietet erstmals in einem integrierten Gesamtsystem alle technischen Voraussetzungen, um diese Defizite zu überwinden und so von traditionellen Suchmaschinen zu innovativen Antwortmaschinen zu kommen (vgl. Fig. 2).

Ein anschauliches Beispiel für die Leistungsfähigkeit von SmartWeb als Antwortmaschine entstand bei einem Blindtest durch den Leiter des Referats 524 des BMBF, Rainer Jansen, der Bernd Reuse als den Initiator des Projektes nach dessen Ausscheiden aus dem aktiven

Dienst im BMBF ablöste. Obwohl SmartWeb vielfach in der Suchdomäne „Sport“ und speziell „FIFA Fußballweltmeisterschaft 2006“ getestet war, stellte Herr Jansen eine Frage aus dem Kulturbereich (siehe Fig. 3), die komplexe Relationen (Name von Sänger/in in der Titelrolle von La Traviata) und zwei Restriktionen (bei den Salzburger Festspielen, in der Premiere) enthielt. Obwohl eine ähnliche Frage SmartWeb zuvor noch nie gestellt worden war, wurde zur Verblüffung auch der anwesenden Fachwissenschaftler diese schwierige Frage präzise und korrekt mit „Anna Netrebko“ beantwortet.

Fig. 3 zeigt auch ein Beispiel dafür, dass der SmartWeb-Nutzer eine bestimmte Modalität der Antwort anfordern kann, indem er am unteren Rand der Bildschirmmaske von SmartWeb auf eine der vier Ikonen für Bild, Videoclip, Text- oder Tondokument mit dem Stift zeigt. Im vorliegenden Fall wird dann zusätzlich zu der Standardantwort in textueller Form noch ein Bild der Künstlerin ausgegeben, das durch eine Bildsuche im Web gefunden wurde.

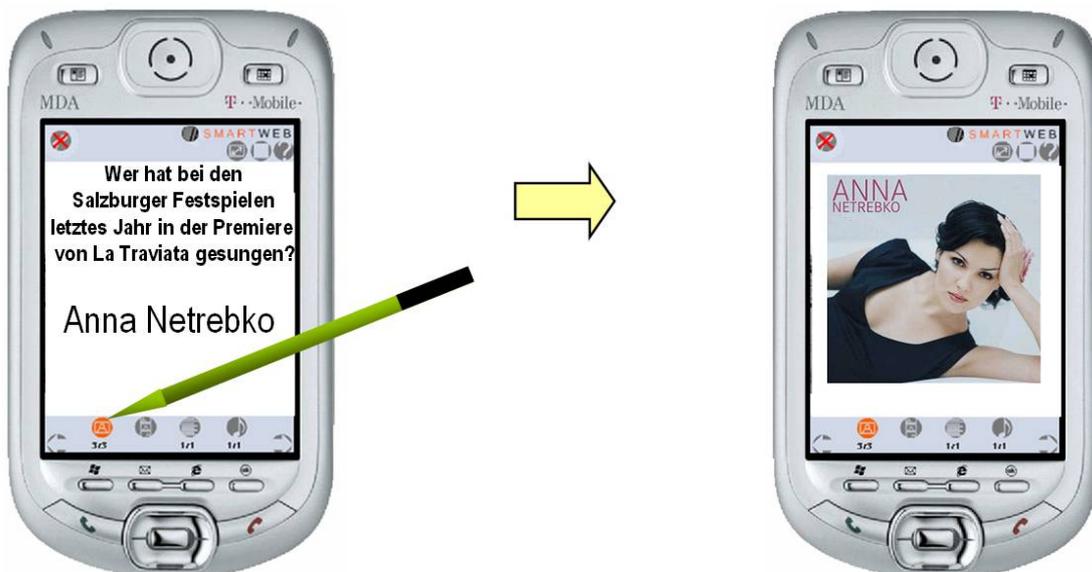


Fig. 3: Beantwortung einer komplexen Suchanfrage im offenen Web durch SmartWeb

3. Die Architektur und die ontologische Infrastruktur von SmartWeb

SmartWeb besteht aus drei großen Softwarebereichen [5]: den mobilen Subsystemen (für Mobiltelefone der MDA-Serie, Autos und Motorräder), der multimodalen Dialogverarbeitung und den semantischen Zugriffsdiensten (vgl. Fig. 4). Das Laufzeitsystem von SmartWeb umfasst 2,5 Gigabyte Software und initiale Daten. SmartWeb kann parallel eine Vielzahl von Benutzern unterstützen, weil das System nach dem Client-Server-Prinzip arbeitet und der Server gleichzeitig mehrere Dialoge mit verschiedenen mobilen Nutzern abwickeln kann. Dazu wird von der SmartWeb-Dialogplattform für jeden neuen eingehenden Anruf eine neue Instanz des Dialogservers instantiiert. Die Hauptverarbeitungsleistung von SmartWeb wird

nicht auf dem mobilen Endgerät, sondern auf dem über eine breitbandige Funkverbindung erreichbaren Server erbracht, auf dem eine leistungsfähige Spracherkennung mit einer semantischen Anfrageanalyse und einer Komponente zum Dialogmanagement kombiniert ist. Auf dem PDA-Client läuft eine Java-basierte Steuerungskomponente für alle multimodalen Ein- und Ausgaben sowie die graphische Bedienoberfläche von SmartWeb. Falls einmal die Verbindung zum Sprachdialog-Server abbricht, verfügt der PDA zur Überbrückung der Störung auch über eine einfache Sprachbedienung auf der Basis von VoiceXML.

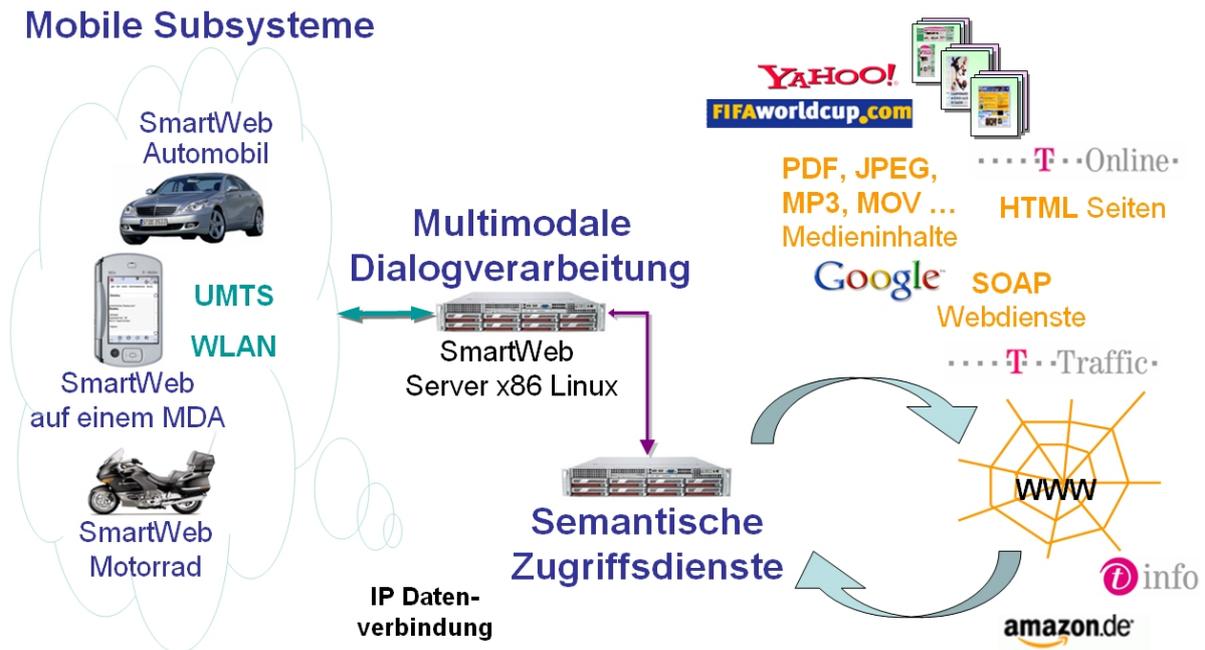


Fig. 4: Die drei großen Softwarebereiche von SmartWeb

Die Beantwortung von Anfragen an SmartWeb wird nicht nur über ein einziges Verfahren realisiert, sondern ein semantischer Mediator entscheidet, wie eine vom Dialogsystem analysierte Anfrage mithilfe der verschiedenen semantischen Zugriffsdienste von SmartWeb beantwortet werden soll. Neben der gezielten Informationsextraktion aus Passagen von Dokumenten im offenen Web, die zunächst mit klassischen Suchmaschinen wie Google gefunden werden, kann z.B. auch eine Komposition von Webdiensten oder ein Aufruf eines Zugriffsagenten auf ein Portal durch den semantischen Mediator ausgelöst werden. Wie in VerbMobil [9] wird in SmartWeb also ein Ansatz verfolgt, der mehrere Lösungswege parallel vorhält und je nach Aufgabenstellung den besten Ansatz zur Reaktion auf die Benutzereingabe verfolgt.

SmartWeb kann neun verschiedene Webdienste automatisch je nach Bedarf zu höherwertigen Diensten komponieren (u.a. Routenplanung, Wetterinformation, Veranstaltungskalender, Verkehrsinformation, Webcam-Suche). Dabei nutzt das System die semantische Beschrei-

bung dieser Webdienste mithilfe von Erweiterungen der standardisierten Dienstbeschreibungssprache OWL-S, die auf einer Beschreibungslogik mit effizienten Inferenzverfahren beruht. Die Komposition der Dienste erfolgt in SmartWeb über einen musterbasierten Planungsalgorithmus, der bei Unterspezifikation von Anfragen auch einen Klärungsdialog mit dem Benutzer initiieren kann.

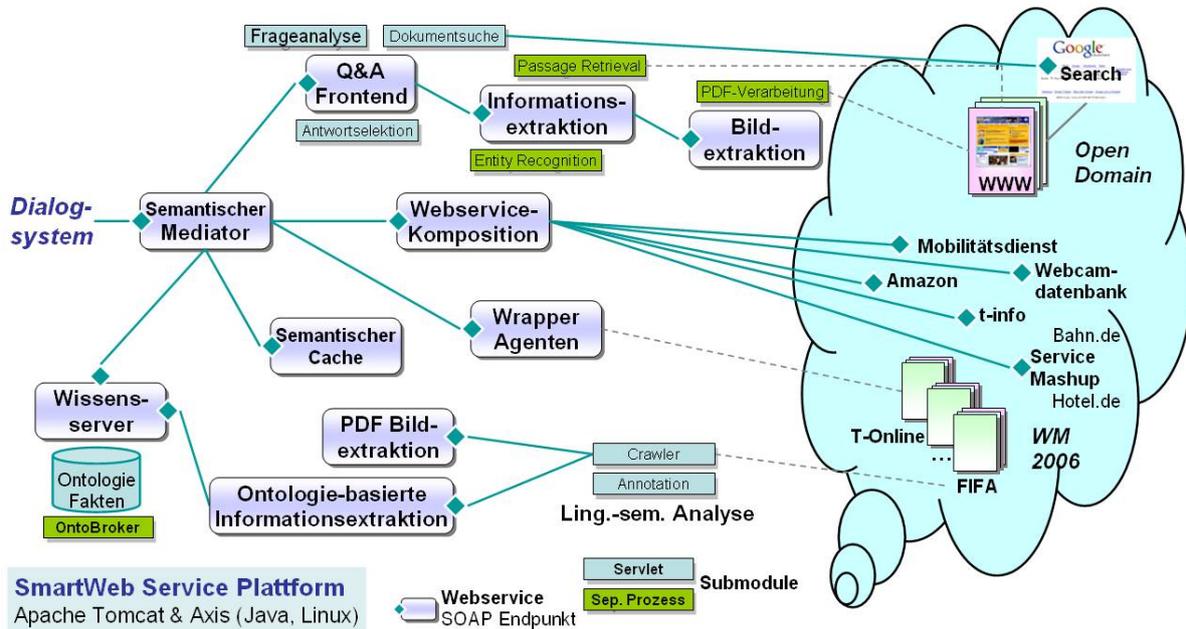


Fig. 5: Die Ansteuerung von Zugriffsdiensten durch den semantischen Mediator

Ein wichtiges Entwurfsprinzip von SmartWeb war die gezielte Benutzung von Ontologien in allen Verarbeitungskomponenten [6]. Die ontologische Infrastruktur SWIntO (**SmartWeb Integrated Ontology**) basiert auf einer Kombination der weit verbreiteten Universalontologien DOLCE und SUMO, die durch spezielle Diskurs- und Medienontologien ergänzt wurden [4]. Der W3C-Standard EMMA (**Extensible MultiModal Annotation Markup Language**) wurde in SmartWeb so ausgebaut, dass nicht nur die Semantik der multimodalen Benutzereingaben sondern auch der multimodalen Systemausgaben formal damit beschrieben werden kann, so dass alle Komponenten auf dem Server Information im EMMA-Format austauschen können. Da wir heute noch ganz am Anfang des semantischen Web als zukünftiges Internet stehen und nur ganz wenige Webinhalte schon ontologisch annotiert sind, wurde in SmartWeb auch intensiv an der automatischen Generierung von semantischen Annotationen für bestehende HTML-Seiten, Photos und PDF-Dokumente aus dem Web gearbeitet. Anhand eines Korpus zum Bereich „Fußball“ wurden 5000 HTML-Seiten, 300 PDF-Dokumente und 15.000 Photos mit sprachtechnologischen Methoden der Informationsextraktion [1] sowie wissensbasierten Layout- und Bildanalyseverfahren bearbeitet. Für die Anwendungsdomäne „FIFA Fußballweltmeisterschaft 2006“ konnten im finalen Prototypen von SmartWeb 5000 Photos anhand der Bildunterschriften mit Spielen und Personen und 3400 Textpassagen zu Spielereignissen

automatisch semantisch annotiert werden. Mit 110 komplett annotierten Mannschaftsphotos, 40 markierten Videosequenzen und 12 Audioreportagen ergeben sich rund 100.000 Ontologieinstanzen in einer umfangreichen Wissensbasis, die manuell erstellte mit automatisch extrahierten semantischen Annotationen aller Medienobjekte kombiniert.

4. Neue Formen der Multimodalität für mobile Dialogsysteme

In SmartWeb kann die Handykamera der MDA-Telephone genutzt werden, um das Eintippen oder Sprechen komplexer Objektnamen bei Anfragen an das System zu umgehen. Der Benutzer fotografiert einfach das Referenzobjekt selbst, ein Photo oder ein Modell davon und ein servergestütztes Objekterkennungssystem ordnet dann als Webdienst aufgrund charakteristischer Bildpunkte selbst bei geändertem Aufnahmewinkel und modifizierten Beleuchtungsparametern das Photo einem in einer Bilddatenbank abgelegten Referenzbild zu. Hier wird also erstmals eine Bildeingabe mit einer Spracheingabe kombiniert, so dass crossmodale Fragen wie in Fig. 6 möglich werden: Nachdem der Nutzer ein Modell des Brandenburger Tores mit seiner Handykamera photographiert hat, kann er einfach fragen: „Wie komme ich dahin?“

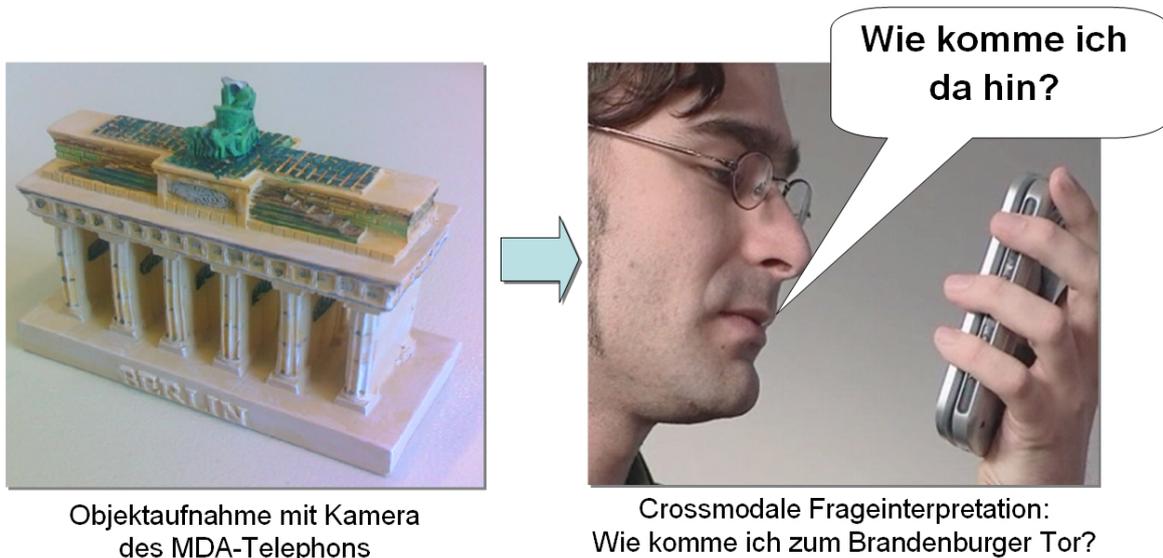


Fig. 6: Crossmodale Anfragen durch Kombination von Handykamera und Spracheingabe

Eine weitere Innovation von SmartWeb ist die Möglichkeit zur Beantwortung von crossmodalen Anfragen, die sich auf Bildmaterial beziehen, welches als Ergebnis einer vorangegangenen Anfrage von SmartWeb auf dem PDA präsentiert wurde. SmartWeb nutzt bei einer Folgefrage wie „Wer ist der dritte links oben?“ (vgl. Fig. 7) dazu die MPEG7-Repräsentation der Bildsegmente und interpretiert räumliche Relationen wie den Ausgangspunkt (links oben), um den referenzierten Spieler (der dritte) zu bestimmen. Der Spieler mit der Nummer 18 wird durch die semantische Annotation des Bildes als „Tim Borowski“ bestimmt, so dass

die Antwort von SmartWeb ausgegeben werden kann. Die Bilder sind dazu dekomponiert und die einzelnen Bildsegmente werden mit Instanzen aus der Medienontologie annotiert, die wiederum mit einer Sportontologie verknüpft ist, in der z.B. „Lehmann“ als Torwart klassifiziert ist.



Fig. 7: Beantwortung von Anfragen mit crossmodaler räumlicher Referenz auf Bildmaterial

Die Spracherkennung mit offenem Wortschatz war eine besondere Herausforderung in SmartWeb, da bei komplexen Navigationsaufgaben und der Informationssuche im Internet häufig Eigennamen vorkommen, die nicht in dem trainierten Sprachmodell enthalten sind. SmartWeb kann unbekannte Wörter detektieren und klassifizieren, da im Spracherkennung eine OOV-Erkennung (OOV = Out of Vocabulary) integriert wurde. Das Sprachsignal wird parallel auch von einem speziellen Erkennung für Wortuntereinheiten wie z.B. Silben analysiert, dessen Ergebnisse dann von einer Komponente zur Phonem-Graphem-Umwandlung weiterverarbeitet werden. Die Ergebnisse der beiden parallel laufenden Erkennung werden dann in einem gemeinsamen Worthypothesengraphen fusioniert. So wird z.B. in der Eingabe „Ich möchte nach Kleinsendelbach fahren“ der Ortsname „Kleinsendelbach“ als unbekanntes Wort detektiert, wonach als Wortuntereinheiten die Phonemkombinationen „klain“, „zEn“, „d@l“ und „bax“ erkannt und in die Graphemfolge „Kleinsendelbach“ verwandelt werden. Durch die hybride Erkennung von Wörtern und Silben konnte die Erfolgsrate in SmartWeb verdoppelt werden, da auch die Umwandlung von Laut- in Buchstabenfolgen nahezu fehlerfrei erfolgt.

Eine Weltneuheit bietet SmartWeb auch mit Sprachdialogen, in denen das System mehrere Internet-Dienste mit Fahrzeugdiensten automatisch verknüpft: Nachdem der Fahrer auf die lokationssensitive Frage „Wo gibt es hier Italiener?“ passende Restaurants auf der digitalen Karte angezeigt bekommt und eines ausgesucht hat, kann er den frei formulierten Dialog fortsetzen mit „Wie komme ich dahin und vorher muss ich noch tanken“. Er bekommt dann eine Routenplanung zum Restaurant, die auch noch einen Tankstopp einplant (vgl. Fig. 8).

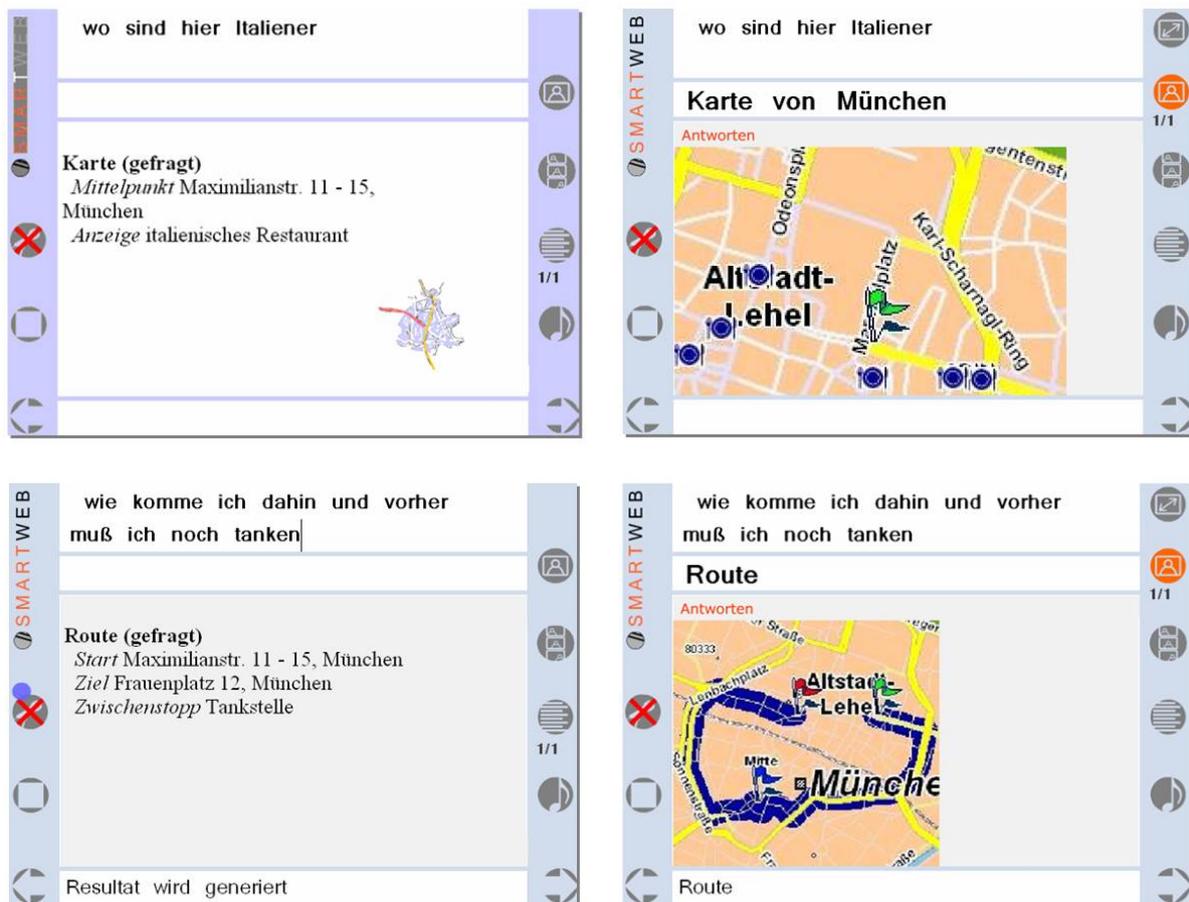


Fig. 8: Kohärenter Navigationsdialog durch die Kombination von mehreren Webdiensten

SmartWeb erlaubt erstmals auch eine multimodale Kommunikation für einen Motorradfahrer (vgl. Fig. 9) über ein im Helm (über Bluetooth mit der SmartWeb-Software verbunden) eingebautes Mikrophon und einen Lautsprecher, die selbst bei starken Fahrgeräuschen einen Sprachdialog ermöglichen [2]. Ein haptisches Bedienelement, das handschuhbedienbar im Lenker eingebaut ist und über Krafrückkopplung verfügt, ermöglicht zusätzlich die Selektion von Diensten aus dem graphischen Menü auf dem großen Cockpit-Bildschirm. SmartWeb warnt durch Nutzung von P2P-Kommunikation über WLAN zwischen Autos und dem Motorrad z.B. sofort vor der Gefahr eines Auffahrunfalls am Stauende hinter einer Kurve, die zu einer unmittelbaren Vollbremsung des Motorradfahrers führen muss, da auch das vorausfahrende Fahrzeug stark abbremst. Innovativ bei SmartWeb ist auch die Berücksichtigung

der kognitiven Belastung des Motorradfahrers (z.B. abhängig von der aktuellen Fahrgeschwindigkeit) und der Intensität des vorliegenden Gefahrenpotentials bei der Generierung multimodaler Warnhinweise. So kann eine Gefahr durch ein Warnsymbol auf dem Bildschirm des Motorrad-Cockpits angezeigt werden, durch ein akustisches Warnsignal oder durch Sprachsynthese, die im Motorradhelm ertönen. Das kontextsensitive Gefahrenwarnsystem benutzt eine Gefahrenontologie und nutzt neben Daten aus dem Motorrad wie Außentemperatur, Reifendruck, Geschwindigkeit und Reifenschlupf auch Meldungen von Sensoren aus Fahrzeugen im Umfeld des Motorrads, um regelbasiert zu berechnen, welche Kombination von Modalitäten in der derzeitigen Fahrsituation eine optimale Balance zwischen der Ablenkung des Fahrers und der Dringlichkeit der Warnung ergibt.

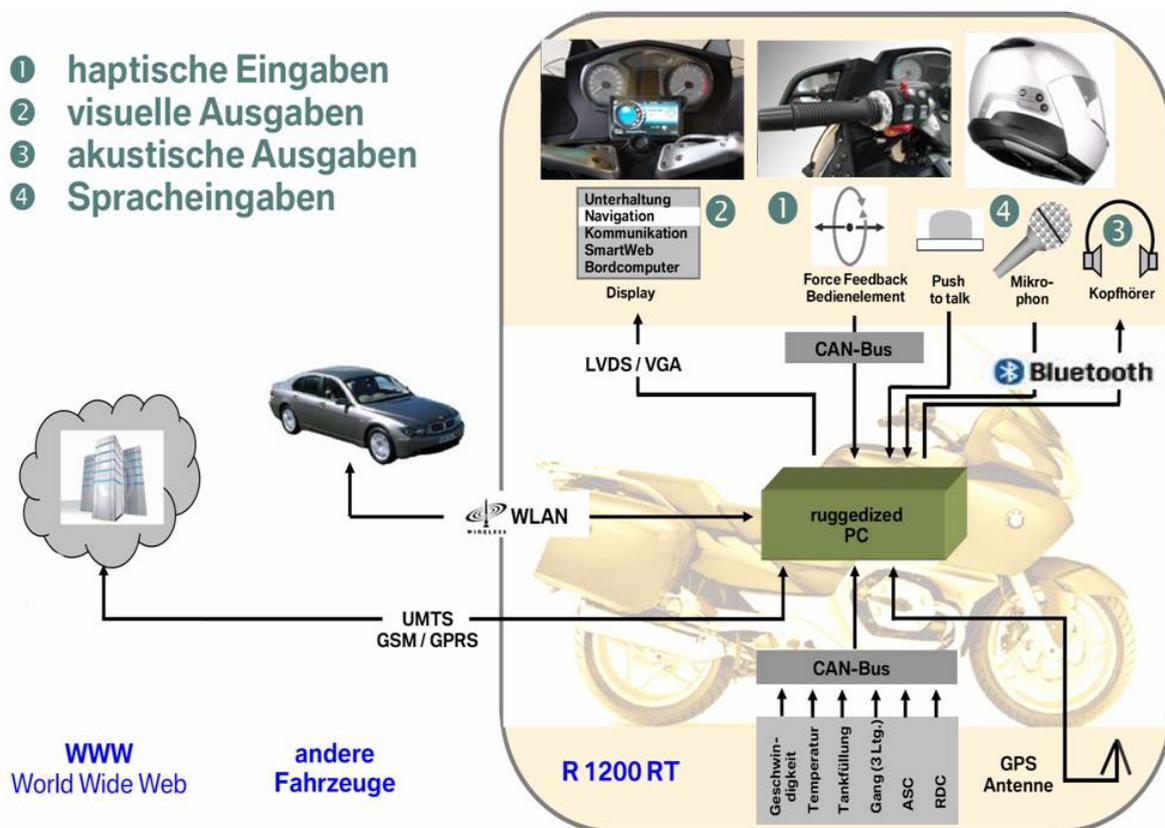


Fig. 9: Multimodale Dialogmöglichkeiten von SmartWeb für Motorradfahrer

5. Wissenschaftliche und Wirtschaftliche Bilanz des SmartWeb-Projektes

Mit 11 Patentanmeldungen, 8 Produktinnovationen und über 140 wissenschaftlichen Publikationen ist das SmartWeb-Projekt außerordentlich erfolgreich, weil es die gesamte Innovationskette von der Grundlagenforschung bis zur neuartigen Produktfunktion vollständig durchlaufen hat. Die Ergebnisse von SmartWeb sind in eine neue Telematikgeneration von Siemens VDO und DaimlerChrysler eingeflossen, bei denen Internet-Information mit Fahr-

zeugfunktionen kombiniert werden und aktuelle, externe Informationsdienste in das HMI des Fahrzeuges integriert werden ohne dass die Dialogschnittstelle verändert werden muss. Da mehr als 50 Prozent der schweren Unfälle mit mangelnder Information des Fahrers verknüpft sind, ist die in SmartWeb erstmals auf einem BMW-Motorrad demonstrierte multimodale Generierung von Warnmeldungen über lokal begrenzte Gefahren, z.B. Fahrbahnglätte oder Ölspuren, durch intelligente Car2X-Kommunikation zwischen Autos und Motorrädern ein wesentlicher Beitrag zu erhöhter Verkehrssicherheit. Unterstützt von den Deutsche Telekom Laboratories sind zahlreiche Transferaktivitäten von SmartWeb-Technologien in die T-Divisionen durchgeführt worden.



Fig. 10: Bundesministerin Dr. Annette Schavan beim Test von SmartWeb in der Mercedes R-Klasse mit der Frage „Was meldet „Die Zeit“ zum Thema Politik?“ auf der CeBIT 2007

Das vom Bundesministerium für Bildung und Forschung (BMBF) mit 13,7 Mio. Euro geförderte Projekt SmartWeb hat mit 16 Partnern aus Industrie und Forschung wesentliche wissenschaftliche und technische Grundlagen für das neue Leuchtturmprojekt *Theseus* gelegt, das vom Bundesministerium für Wirtschaft und Technologie (BMWi) zunächst mit 90 Mio. Euro ab 2007 für fünf Jahre gefördert wird. In *Theseus* sollen von 30 Konsortialpartnern die semantischen Technologien für innovative Anwendungen des zukünftigen Internet der Dienste erschlossen werden. Das in SmartWeb entwickelte multimodale Dialogsystem dient auch als Ausgangspunkt für ein weiteres Leuchtturmprojekt mit der Bezeichnung *SIM-TD* (Sichere Intelligente Mobilität – Testfeld Deutschland), das neben dem BMBF und BMWi auch vom Bundesverkehrsministerium ab Ende 2007 gefördert wird. Die Car-to-X Kommunikation soll

dabei unter realen alltäglichen Verkehrsbedingungen mit einer großen Fahrzeugflotte aller wichtigen deutschen Automobilhersteller erprobt und ausgebaut werden, um die Anzahl der Staus zu reduzieren, den Verkehrsfluss zu verbessern und die Sicherheit zu erhöhen.

6. Danksagung

Als Gesamtprojektleiter danke ich dem Bundesministerium für Bildung und Forschung (BMBF) für die Förderung von SmartWeb von 2004 – 2007 unter dem Förderkennzeichen 01 IMD 01 sowie allen beteiligten Industriepartnern für die zusätzlich eingesetzten Investitionsmittel. Ganz besonders möchte ich Herrn Bernd Reuse vom Referat 524 des BMBF danken für seine mutige und zukunftsweisende Entscheidung für SmartWeb und seine vorbildliche Unterstützung des Vorhabens. Sein Nachfolger im Amt, Herr Rainer Jansen, hat das Vorhaben konsequent weitergeführt und zu einem sehr erfolgreichen Abschluss im September 2007 verholfen. Ohne die hervorragenden Leistungen meiner Mitarbeiter Norbert Reithinger (Modulkoordination), Gerd Herzog (Systemgruppe) und Anselm Blocher (Projektmanagement) und ihrer Teams am DFKI wären die anspruchsvollen Ziele von SmartWeb nicht erreichbar gewesen. Insgesamt ist SmartWeb eine Teamleistung von über 100 Wissenschaftlern und Softwareingenieuren, die in einer verteilten Entwicklungsumgebung ein multimodales Dialogsystem geschaffen haben, das trotz seiner weltweit einzigartigen Leistungsfähigkeit eine große Robustheit unter Realzeitbedingungen aufweist. Daher möchte ich mich an dieser Stelle bei allen Mitarbeitern des SmartWeb-Konsortiums sehr herzlich für die exzellente Kooperation bedanken.

7. Literatur

- [1] Cramer, Irene, Leidner, Jochen, Klakow, Dietrich (2006): Building an Evaluation Corpus for German Question Answering by Harvesting Wikipedia. In: Proceedings of LREC 2006, International Conference on Language Resources and Evaluation, Mai 2006, Genova, Italy
- [2] Kaiser, Moritz, Mögele, Hans, Schiel, Florian (2006): Bikers Accessing the Web: The SmartWeb Motorbike Corpus. In: Proceedings of the LREC 2006, International Conference on Language Resources and Evaluation, Juni 2006, Genova, Italy
- [3] Maybury, Mark, Wahlster, Wolfgang (eds.) (1998): Readings in Intelligent User Interfaces. San Francisco: Morgan Kaufmann, 1998.
- [4] Oberle, Daniel, Ankolekar, Anuypriya, Hitzler, Pascal, Cimiano, Philipp, Sintek, Michael, Kiesel, Malte, Mougouie, Babak, Vembu, Shankar, Baumann, Stephan, Romanelli, Massimo, Buitelaar, Paul, Engel, Ralf, Sonntag, Daniel, Reithinger, Norbert, Loos, Berenike, Porzel, Robert, Zorn, Hans-Peter, Micelli, Vanessa, Schmidt, Christian, Weiten, Moritz, Burkhardt, Felix, Zhou, Jianshen (2007): DOLCE ergo SUMO: On Foundational and Domain Models in SWIntO (SmartWeb Integrated Ontology). Journal of Web Semantics, 3, 2007.

- [5] Reithinger, Norbert, Herzog, Gerd, Blocher, Anselm (2007): SmartWeb - Mobile Broadband Access to the Semantic Web. In: KI - Künstliche Intelligenz, 2/2007, p. 30-33.
- [6] Sonntag, Daniel, Engel, Ralf, Herzog, Gerd, Pfalzgraf, Alexander, Pflieger, Norbert Romanelli, Massimo, Reithinger, Norbert (2007): Smart Web Handheld - Multimodal Interaction with Ontological Knowledge Bases and Semantic Web Services. In: Huang, Thomas Nijholt, Anton, Pantic, Maja, Plentland, Alex (eds.): Artificial Intelligence for Human Computing. LNCS 4451, Berlin, Heidelberg, Springer, 2007, p. 272-295.
- [7] Wahlster, Wolfgang (ed.) (2006): SmartKom: Foundations of Multimodal Dialogue Systems. Cognitive Technologies Series, Heidelberg, Germany: Springer, 2006, 644 pp.
- [8] Wahlster, Wolfgang (2003): Towards Symmetric Multimodality: Fusion and Fission of Speech, Gesture, and Facial Expression. In: Günter, A., Kruse, R., Neumann, B. (eds.): KI 2003: Advances in Artificial Intelligence. Proceedings of the 26th German Conference on Artificial Intelligence, September 2003, Hamburg, Germany, Pages 1 - 18, Berlin, Heidelberg: Springer, LNAI 2821, 2003.
- [9] Wahlster, Wolfgang (2000) (ed.): Verbmobil: Foundations of Speech-to-Speech Translation. Berlin, Heidelberg, New York, Barcelona, Hong Kong, London, Milan, Paris, Singapore, Tokyo: July 2000, Springer.