

## Workshop "Voice-Enabled Services aus Anwender- und Benutzersicht: Ein State-of-the-Art-Bericht aus Forschung, Entwicklung und Praxis."

In den letzten 10 Jahren wurden beachtliche technische and methodische Fortschritte im Bereich sprachgesteuerter Systeme und Anwendungen gemacht: Dedizierte Entwicklungsumgebungen (Grammatikmodellierung, Dialoggestaltung), Standardisierung (SALT, Voice-XML), Sprachplattformen und Voice Application Development Environments, usw. sind mittlerweile vorhanden. Und: ASR-Technologie ist mittlerweile ausgereift, Rechenleistung (als limitierender Faktor der ASR-Qualität) ist sehr viel preiswerter geworden. Jetzt soll diese neue Sprachtechnologie „nur noch“ in die Praxis umgesetzt werden, d.h. es sollen Anwendungen ausfindig gemacht, entwickelt und vermarktet werden, die für Betreiber und Benutzer gleichermaßen sinnvoll und lohnend sind.

Dazu wird oft an bewährten oder gerade gefeierten Erfolgskonzepten wie z.B. Web Portals angeknüpft (→ Sprachportale). Auch werden alte, fast vergessene Konzepte wie multimodale Schnittstellen wieder aufgegriffen und unter dem Stichwort „Multi-channel“ als neuer Marketingansatz ausgearbeitet. Ein weiterer Ansatzpunkt ist die Einbettung von Sprachtechnik in bereits existierende Produkte wie bspw. Mobilfunkgeräte.

An innovativen Ansätzen und kreativen Vorstößen mangelt es wirklich nicht. Trotzdem lässt der endgültige Durchbruch im deutschsprachigen Raum noch auf sich warten! Haben wir die Zukunft der Sprachtechnologie zwar bereits erkannt, stecken jedoch mitten in derer Umsetzung? Sind wir also auf dem (einen?) richtigen Weg? Was fehlt noch, und wie weit sind wir wirklich von Erfolg versprechenden Ansätzen, Techniken und Diensten entfernt? Was müsste außerdem geschehen und wie viel Zeit und Aufwand wird uns das noch kosten? Wie sehen es die Akteure selbst? Ist die Sprachtechnik selbst Schuld daran, dass sie noch nicht anstandslos akzeptiert wird? Oder sind die Anwender, für die die Sprachsysteme entwickelt werden, noch zu ängstlich, zu zögerlich? Oder liegt der „Knackpunkt“ vielleicht an ganz anderer Stelle z.B. bei Vorbehalten und negativen Assoziationen bei Managern/Entscheidungssträgern?

In diesem Workshop möchten wir u. a. anhand konkreter Fallbeispiele aus Forschung, Entwicklung und Praxis dazu beitragen, zumindest einige Erfolgskriterien der sprachlichen Interaktion als Komponente ganz bestimmter Dienstleistungen zu identifizieren und zu definieren. Dabei spielen nicht nur ergonomische sondern sicherlich auch wirtschaftliche Kriterien eine wichtige Rolle. Sprachtechnologie kann – richtig durchdacht und sinnvoll eingesetzt – zu Kostensenkung und Serviceverbesserung gleichermaßen führen und trägt somit zur viel beschworenen Innovation in der Wirtschaft bei.

### Themencluster I: Status Quo der gegenwärtigen Systeme und Anwendungen

- Stimmen die zugrunde gelegten Metaphern aus dem Internetbereich, z.B. *Voice Portal*? Oder sollte man sich

eher an gesprochener Kommunikation zwischen Menschen oder auch an der Radio-Metapher orientieren? (Informations- vs. Transaktionssysteme).

- Welche Rolle spielt dabei die sog. *Natural Language Understanding*? Hilft die Kopplung von reinen *Voice-Enabled Services* mit alternativer multi-modaler Ein- und Ausgabe (Gestik, Mimik) bei der Schaffung benutzerfreundlicher Mensch-Maschine Schnittstellen?
- Beurteilung und Vorhersage der Qualität von Systemen und Diensten. Wie ist die Usability bei unimodalen, rein sprachgesteuerten, bzw. multimodalen Schnittstellen zu definieren, wie zu messen?
- Wo liegt momentan die Verantwortlichkeit für Sprachentwicklungsprojekten und nach welchen Erfolgskriterien werden die Sprachanwendungen entwickelt? Am Markt gibt es zu wenig erfahrene Dialogspezialisten! Wie kann nutzerzentriertes Entwickeln von Sprachanwendungen gefördert werden?
- Fokusbranchen und Nachzügler: wo spielt die „Musik“?

### Themencluster II: Wirtschaftlichkeit der gegenwärtigen Systeme und Anwendungen

- Welche sind die *unique selling points* von Sprachapplikationen? Wo bleiben die Erfolgsgeschichten aus Betreiber/Benutzer-Sicht?
- Wo liegen die (nicht nur wirtschaftlichen) Erfolgsfaktoren und Stolpersteine bei der Entwicklung und Einführung von Sprachdiensten? Was sagen uns die früheren sowie aktuellen Marktstudien? Was sagen uns die Nutzer über heutige Sprachanwendungen?
- Wie rechnet sich ein Sprachdienst? Konkrete Fallbeispiele mit Business-Case und ROI-Berechnungen schaffen ein besseres Verständnis für Sprachanwendungen aus ökonomischer Sicht.
- Das Spektrum der Anwendungen: Gibt es eine Vielfalt an Anwendungspotentiale oder bleiben Sprachdienste letztendlich eine Nische? Gibt es einen klaren Generationswechsel (Stichwort: next generation speech applications)?

### Themencluster III: Welche sind die aktuellen Entwicklungen und Herausforderungen?

- Marktstruktur (*Player*): Auf welchem Gebiet sind die momentanen Herausforderungen angesiedelt? Welche Rolle spielen in diesem Zusammenhang Voice-Application-Service-Provider (VASP), bei denen standardisierte Sprachdienste gehostet werden bzw. bei denen ein Unternehmen seinen Dienst ohne Infrastrukturinvestitionen betreiben kann?

- Technologie: Welche Marktentwicklungen stehen bevor: fallende ASR-SW Preise, Konsolidierung bei den Herstellern, Ausbildung von Spezialisten für VUI-Entwurf, usw.? Was muss von den Marktakteuren – Hersteller und Betreiber von Diensten – getan werden, damit aus der „Vision“ Realität wird: gezielte Marktkommunikation, gemeinsam aus Umsetzungsfehlern lernen, offene Systemarchitektur und Standardisierung, Optimierung von bestehenden reichweitenstarken Systemen u. ä.?
- Positionierung: Fortschritt vs. Akzeptanz: Brauchen wir neue Visionen, neue Leitbilder, neue Metaphern, neue Techniken, ... oder einfach nur mehr Mut und mehr Experimentierfreudigkeit – mit dem kalkulierten Risiko des Scheiterns?
- Dialogentwurf: hier besteht sehr großer Nachholbedarf! Im Dialogdesign liegt ein wesentlicher Schlüssel zum Erfolg. Sprachprojekte benötigen nicht nur IT-Profis sondern vielmehr Sprachdialogspezialisten!

### Zielgruppe:

Der Workshop soll nicht nur Forscher und Entwickler, sondern auch Betreiber und Benutzer sprachgesteuerter Diensten (welcher *Couleur* auch immer) ansprechen.

Im Mittelpunkt stehen die Qualitäts- und Nutzwertanforderungen an solchen Systemen und Anwendungen sowie die Frage, inwiefern die aktuellen technischen Möglichkeiten diesen Anforderungen aus Anwendersicht bereits gerecht werden.

Es sind also insbesondere auch Ergonomie-, Kommunikations- und Organisationsexperten, Markt- und Trendforscher sowie Entscheidungsträger aus der Wirtschaft angesprochen.

### Beiträge:

Akzeptiert werden deutschsprachige Beiträge von **maximal 5 Seiten Umfang**. Sie sind bis zum 30. April 2004 in digitaler Form als PDF-Datei an [p.h.vossen@gmx.de](mailto:p.h.vossen@gmx.de) zu senden. Eingereichte Beiträge werden von den Mitgliedern des Programmkomitees sorgfältig geprüft.

Akzeptierte Beiträge werden in der GI-Edition "**Lecture Notes in Informatics**" veröffentlicht. Beachten Sie daher bei der Verfassung die Autorenrichtlinien, nachzulesen unter <http://www.gi-ev.de/LNI/autorenrichtlinien/index.html>.

Zur Vermeidung späterer umfangreicher Korrekturen und zur Erleichterung unserer Arbeit, empfehlen wir allen Autoren, diese Richtlinien bereits bei der Ersteinreichung einzuhalten.

*Vorschlägen für Systemvorführungen oder Interactive Poster Sessions stehen wir aufgeschlossen gegenüber.*

### Termine:

- |                   |   |
|-------------------|---|
| 30.04.2004        | Einreichung von Workshop-Beiträgen                                  |
| 28.05.2004        | Benachrichtigung über Annahme bzw. Ablehnung von Workshop-Beiträgen |
| 30.06.2004        | Abgabe der fertigen Druckvorlagen für den Tagungsband               |
| <b>20.09.2004</b> | <b>Workshop von 14:00 Uhr bis 17:30 Uhr</b>                         |

### Organisation:

Paul Hubert Vossen, Voice & Visual  
Klingenstrasse 15  
D-70771 Leinfelden-Echterdingen  
0711.4409440 / 0163.7885599  
[p.h.vossen@gmx.de](mailto:p.h.vossen@gmx.de) / [voiceandvisual@gmx.info](mailto:voiceandvisual@gmx.info)

### Programmkomitee:

Dr. Elke **Kachelrieß**, Dialog Design & Usability, Berlin  
Dipl.-Kfm. Jens **Klemann**, Strateco GmbH, Bad Homburg  
Dr. Volker **Kraft**, VoiceObjects AG, Bergisch Gladbach  
Prof. Dr. Wolfgang **Minker**, Universität Ulm  
Dr.-Ing. Sebastian **Möller**, Ruhr-Universität Bochum  
Dipl.-Psych. Matthias **Peissner**, FhG-IAO, Stuttgart  
Dr. Peter Wetzler, Infoman AG, Stuttgart

### Weitere Infos:

Weitere Informationen zur GI-Jahrestagung *Informatik 2004* sind unter <http://www.informatik2004.de> zu finden