

TP 2 - Verteilte Multimodale Dialogkomponente für das Mobile Endgerät

Norbert Reithinger, Daniel Sonntag, Ralf Engel,
Norbert Pfleger, Massimo Romanelli, David Gelbart*

DFKI GmbH

Intelligent User Interfaces (IUI)

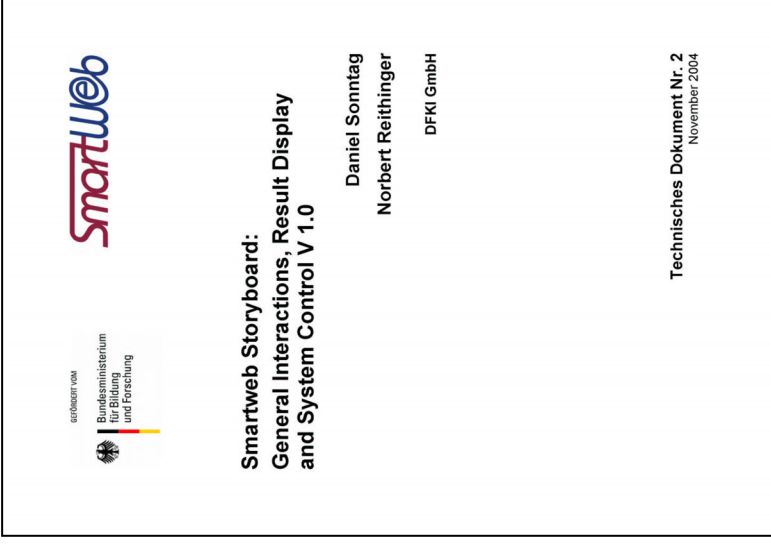
* ICSI

- TP 2 Verteilte Multimodale Dialogkomponente für das Mobile Endgerät
 - AP 2.1 Design der Multimodalen Benutzerschnittstelle und der Dialogverarbeitung
 - AP 2.2 Modalitätsspezifische Analyse, Integration und Fusion
 - AP 2.3 Multimodale Dialogverarbeitung
 - AP 2.4 Multimodale Dialogpräsentation
 - AP 2.5 Englisches System (ICSI)
- AP 1.4.3 Haptik-Interpretation
- AP 4.2.1.3 Diskursontologie
- AP 7.1.5 Verteilte Architektur der Dialogverarbeitung

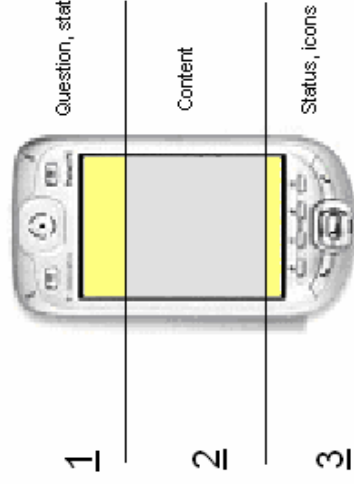
Design der Multimodalen Benutzerschnittstelle (AP 2.1)

Anforderungen (nicht alle bereits im Demonstrator umgesetzt)

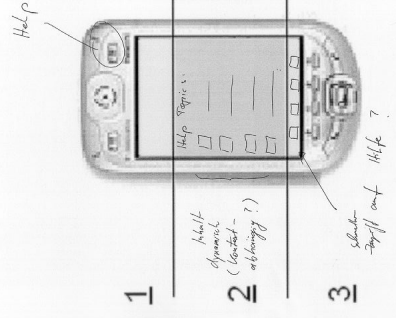
- Grundlegende User-Interface Prinzipien des Endgeräts beachten, z.B. Navigation und Tastenbenutzung (Einhandbedienbarkeit)
- Information des Benutzers über den Erkennungs- und Verarbeitungszustand
- Interaktiver Eingriff des Benutzers in die Verarbeitung, z.B. bei falsch erkannten Wörtern
- Berücksichtigung der Bildschirmgröße: nur die jeweils wichtigste Information wird dargestellt (simplicity by progressive disclosure)



- Bildschirm dreigeteilt
 - Benutzereingaben mit Korrekturmöglichkeiten
 - Inhaltsdarstellung
 - Status und Bedienelemente (den Knöpfen des MDA zugeordnet)



Initiale
Bildschirmaufteilung

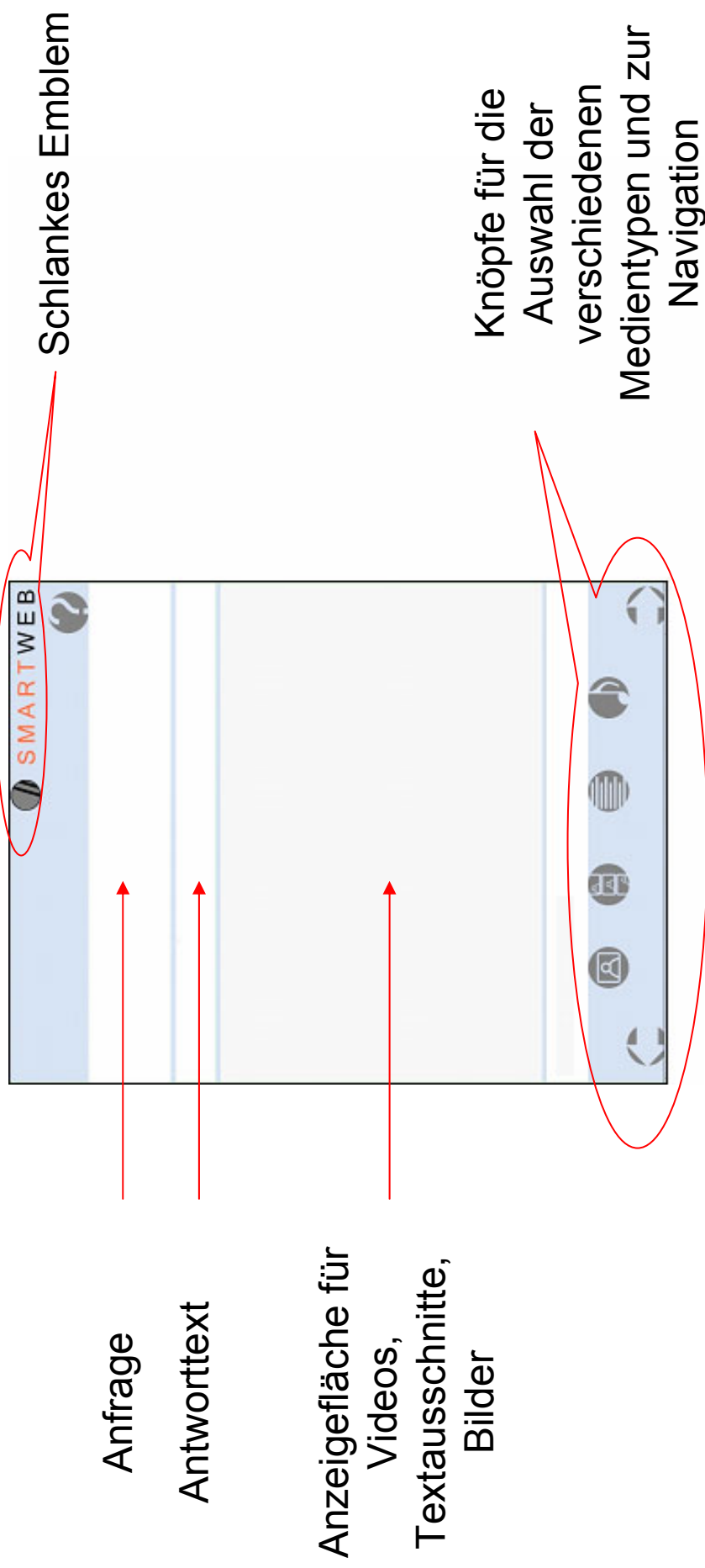


Interaktionsdurchläufe mit
Papierprototypen



Implementierte
Version

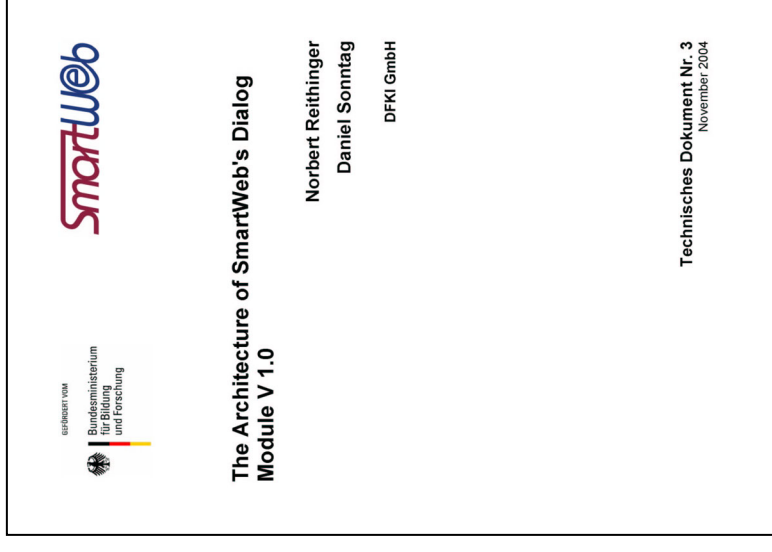
An Bildschirmgröße und Einsatz angepasstes kohärentes Design



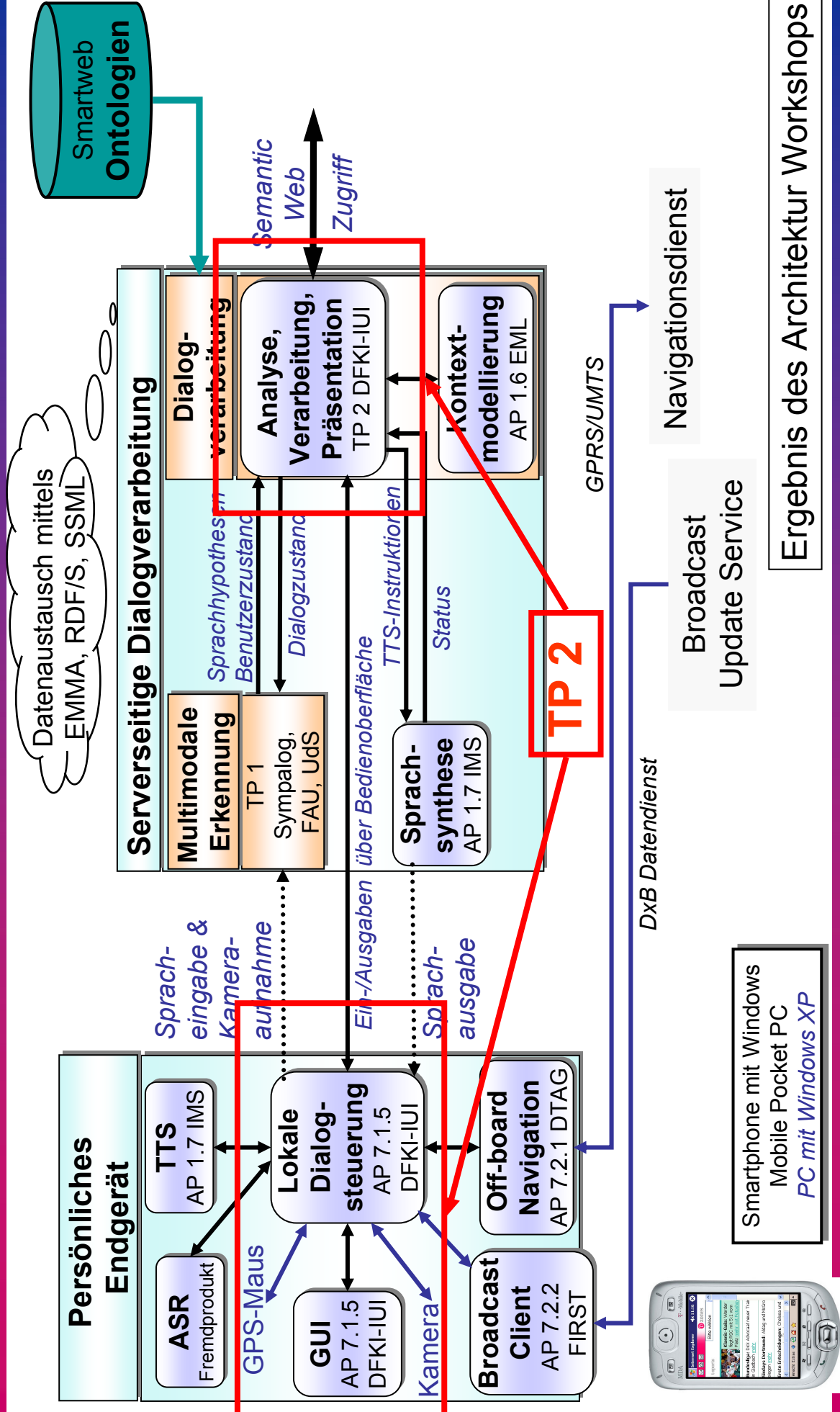
Realisiert in Macromedia Flash, lokale Verarbeitung in Actionscript

Architektur der Dialogverarbeitung (AP 7.1.5)

- Anforderungen:
 - Barge-In
 - Multiuser
 - Ständig unter Kontrolle des Benutzers
- Keine einheitliche Architekturplattform wie in SmartKom
 - SymCTM als Plattform für die serverseitige Integration und die Anbindung des Client
 - Eigene Integrationsplattform auf der Serverseite benötigt
 - Clientseitig Controller angebunden via SymClient

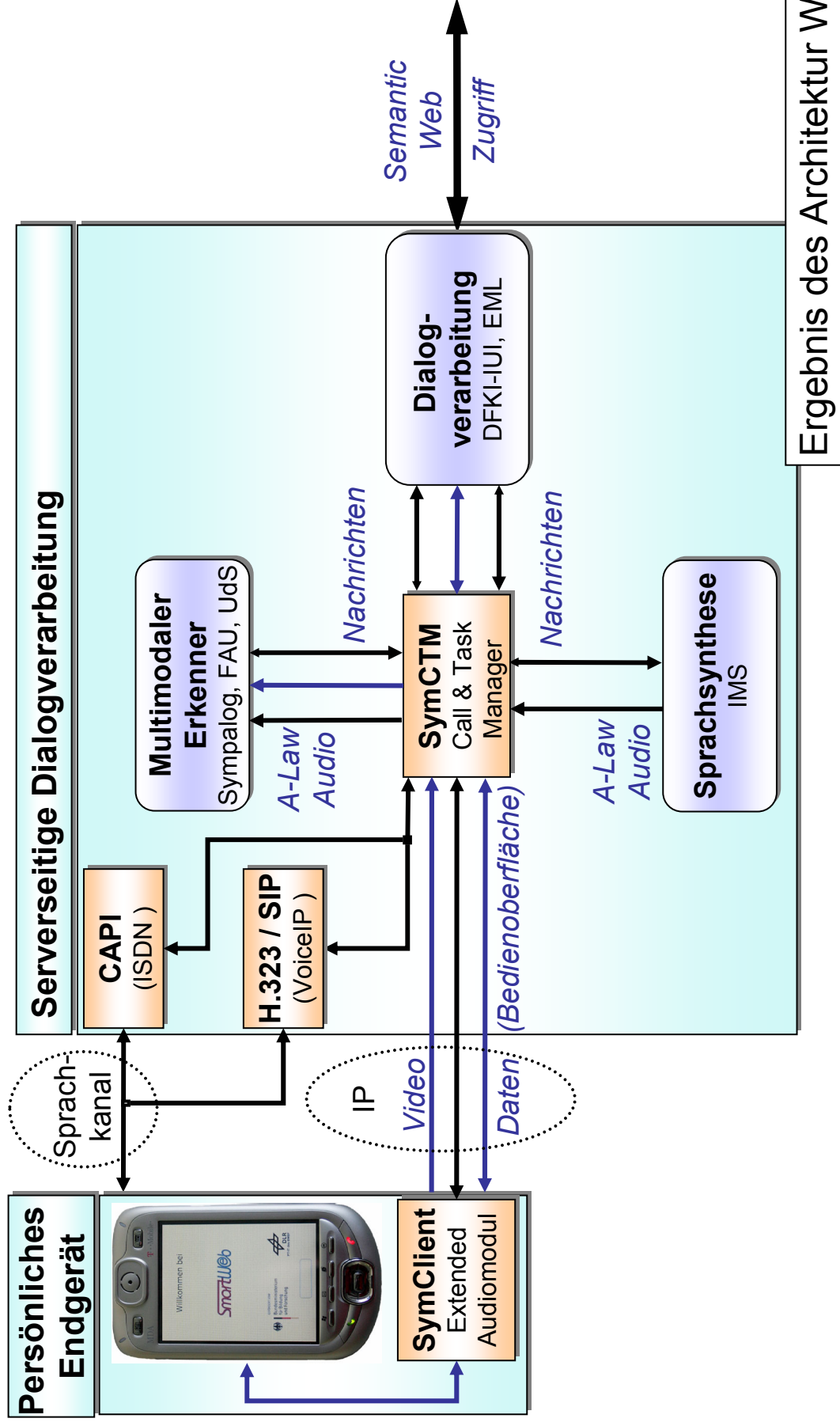


Verteilte multimodale Dialogverarbeitung



Einsatz der

Sympalog-Dialogsystemplattform



Der iHUB – Die Integrationsplattform der serverseitigen Dialogverarbeitung

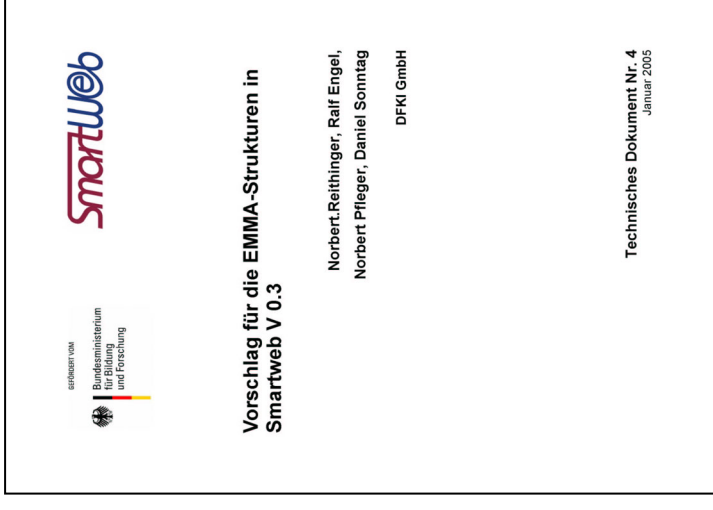
- Ziel
 - Einfache und robuste Integrationsplattform für eigenständige Verarbeitungsmodule (z.B. Sprachinterpretation)
 - Klare, ontologiebasierte Schnittstellen (Jena API)
 - Kontrolle des Nachrichtenversands durch Versionsnummern bei Barge-In/Reinterpretation
- Lösung: iHUB
 - Hub-and-Spoke Architektur
 - Multisessionfähig
 - Schneller Nachrichtenversand
 - Java-basiert

*N. Reithinger, D. Sonntag: An Integration Framework
for a Mobile Multimodal Dialogue System Accessing
the Semantic Web, Interspeech 2005*



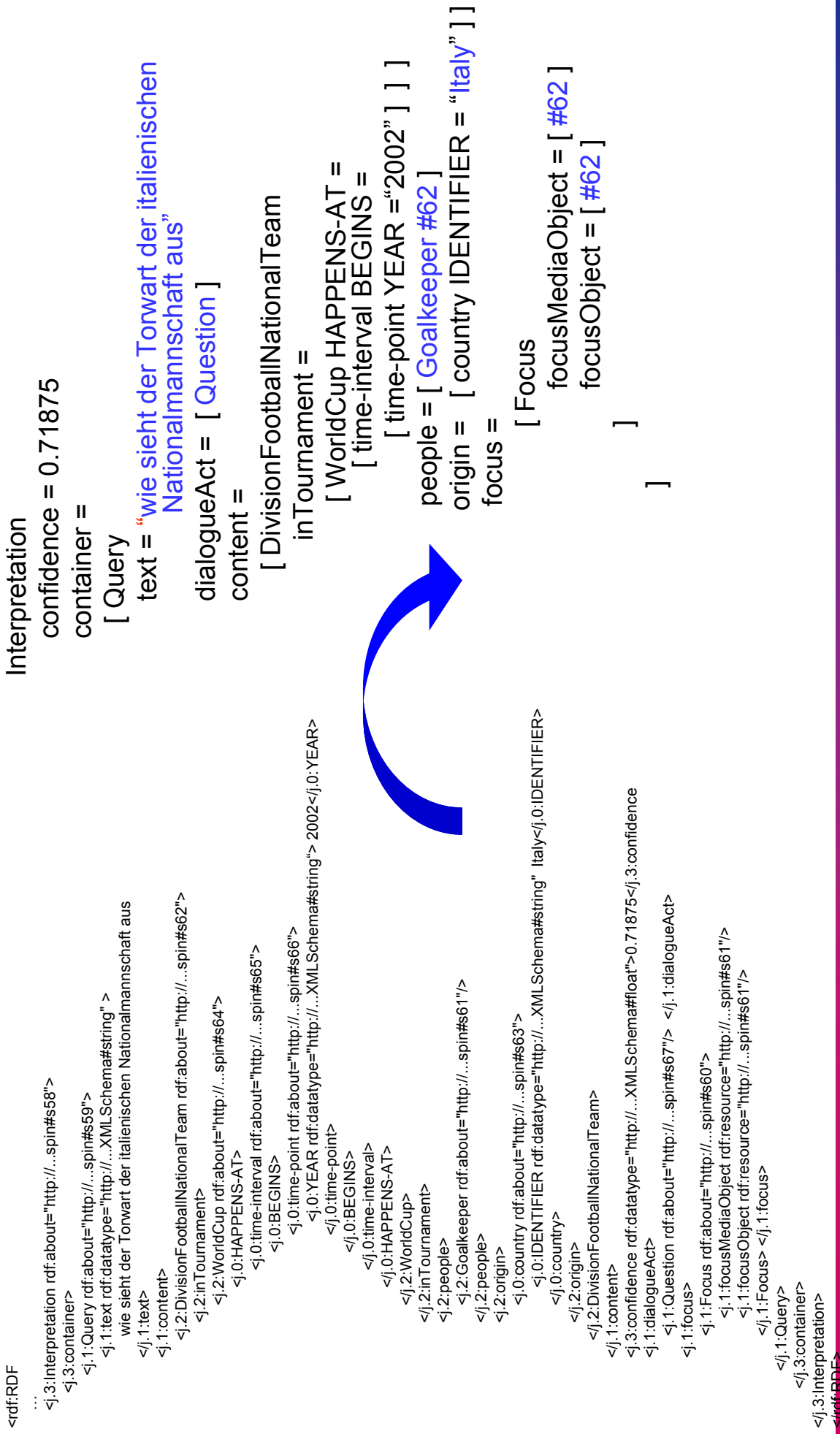
iHUB-GUI

- Datenaustausch basiert auf der W3C Empfehlung EMMA (Extensible MultiModal Annotation markup language):
 - Verwaltungsinformationen wie Zeitstempel, Prozessnummern etc
 - Generische Container für Inhaltsstrukturen
 - Wörtergitter für Ergebnisse des Spracherkenners
 - Speech Synthesis Markup Language (SSML) einbettbar
 - Offen für Erweiterungen
- Schnittstelle zum Semantic Mediator:
RDF/S-Repräsentation für EMMA in der Diskursontologie
- Erweiterung: Smartweb-Emma Namespace (SWEMMA) u.a. für **result** und **status** Tags



Sprachanalyse (AP 2.2)

- Eingabe:
 - Spracherkennung liefert n-beste Ketten im EMMA Wörtergitterformat (z.Zt. $n = 5$)
 - Erweiterung für Out-of-Vocabulary (OOV) Wörter in EMMA definiert
 - Format: XML-EMMA (extern vom Erkennen) bzw. RDF/S-EMMA (intern im iHUB)
- Ausgabe:
 - Abfrage-Pattern mit Frage-Fokus und Dialogakt
 - Format: RDF/S mit SmartWeb-Ontologien als Typsystem

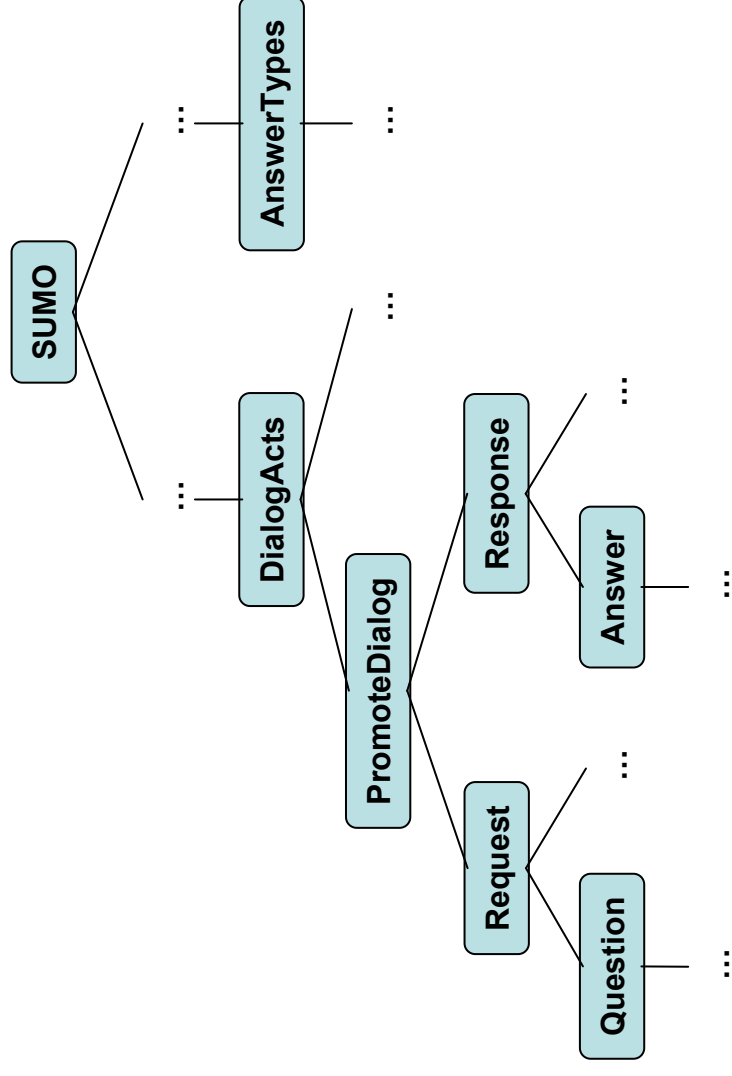


- Technologie
 - Semantischer template-basierter Parser
 - Weiterentwicklung des in verschiedenen Projekten bewährten Systems (SmartKom, EU-Projekte MIAMM und COMIC)
- Anpassungen
 - Einlesen der RDF/S – Ontologie als Typ-System (Jena)
 - Interface-Layer für Ein- / Ausgabe
- Zukünftige Arbeiten
 - Verarbeitung für offene Domänen (OOV)
 - (Semi-)automatische Erstellung der Wissensquellen

R.Engel, Robust and Efficient Semantic Parsing of Free Word Order Languages in Spoken Dialogue Systems, Interspeech 2005

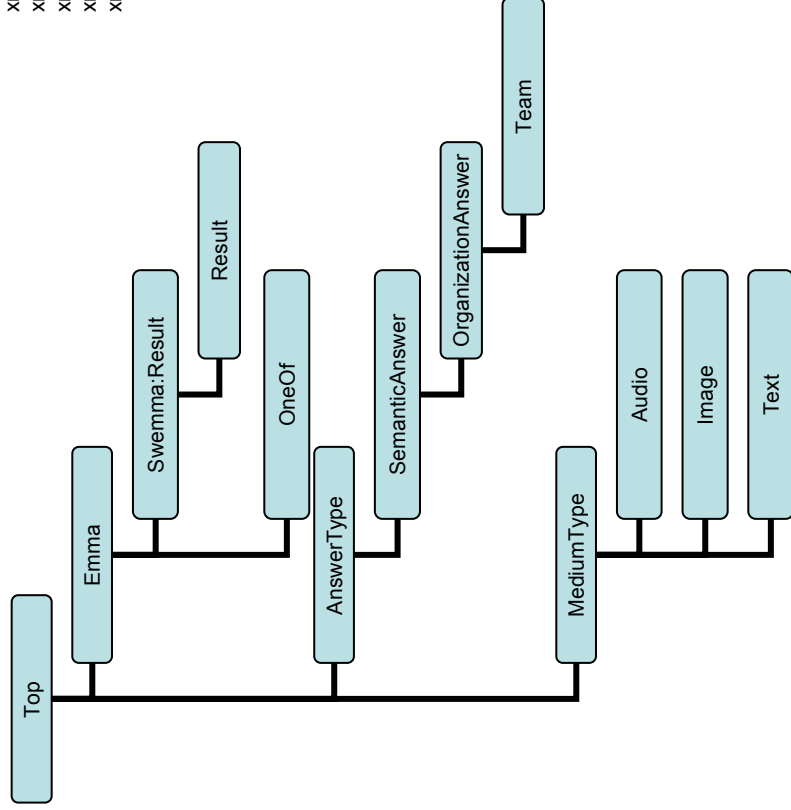
Diskursontologie (AP 4.2.1.3)

- Inhalt: diskurspezifische Konzepte, z.B.
 - Dialogakte
 - Q&A Anbindung
 - SWEMMA-Strukturen
- Definiert Input, Output und interne Strukturen des Dialogmoduls
- Abbildung der Resultate auf Medientypen



Beispiel: Wer war 1990 Weltmeister?

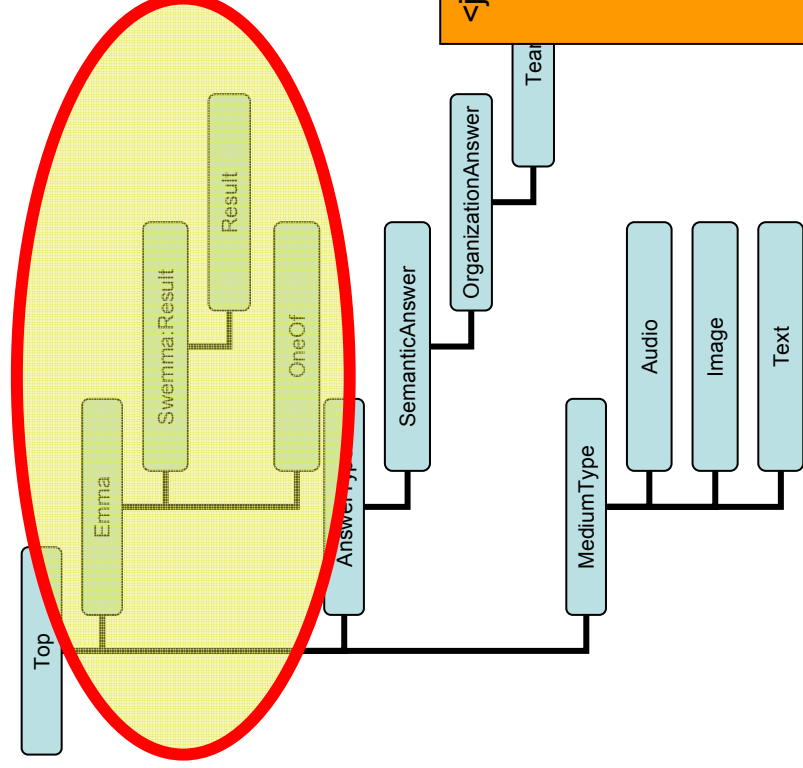
Typenhierarchie



```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:j.0="http://.../swemma#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:j.1="http://.../discourse#"
  xmlns:daml="http://www.daml.org/2001/03/daml+oil#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:j.2="http://.../emma#">
  <j.1:Text rdf:about="http://.../discourse#reapr_text_instance_Weltmeister1990">
  <j.1:sourceURI>http://fifa.worldcup.yahoo.com/06/de/lo/octeam/emb.html</j.1:sourceURI>
  <j.1:Content>DEUTSCHLAND</j.1:Content>
  </j.1:Text>
  <j.1:Result rdf:about="http://.../discourse#result_instance_Weltmeister1990">
  <j.1:answerTypes>
  <j.1:Team rdf:about="http://.../discourse#reapr_team_Weltmeister1990">
  <j.1:mediaTypes rdf:resource="http://.../discourse#reapr_text_instance_Weltmeister1990">
  </j.1:mediaTypes>
  <j.1:Image rdf:about="http://.../discourse#reapr_image_instance_Weltmeister1990">
  <j.1:sourceURI> http://www.zdf.de/ZDFde/img/28/0,1886,2425724,00.jpg</j.1:sourceURI>
  </j.1:Image>
  </j.1:mediaTypes>
  <j.1:Text rdf:about="http://.../discourse#reapr_text2_instance_Weltmeister1990">
  <j.1:Content>GERMANY, I want beer.</j.1:Content>
  </j.1:Text>
  </j.1:mediaTypes>
  </j.1:Team>
  </j.1:answerTypes>
  <j.1:Result>
  <j.2:Emma rdf:about="http://.../emma#emma_root_instance">
  <j.2:container>
  <j.0:Result rdf:about="http://.../swemma#reapr_swemma_result_instance_Weltmeister1990">
  <j.2:container>
  <j.2:OneOf rdf:about="http://.../emma#reapr_oneOf_instance_1">
  <j.2:container rdf:resource="http://.../discourse#result_instance_Weltmeister1990">
  </j.2:OneOf>
  </j.2:container>
  </j.0:Result>
  </j.2:container>
  <j.2:version>1.0</j.2:version>
  </j.2:Emma>
  </rdf:RDF>
  
```

Beispiel: Wer war 1990 Weltmeister?



```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:j.0="http://.../swemma#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:j.1="http://.../discourse#"
  xmlns:daml="http://www.daml.org/2001/03/daml+oil#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:j.2="http://.../emma#">
  <j.1:Text rdf:about="http://.../discourse#reapr_text_instance_Weltmeister1990">
  <j.1:sourceURI>http://fifa.worldcup.yahoo.com/06/de/o/octeam/emb.html</j.1:sourceURI>
  <j.1:Content>DEUTSCHLAND</j.1:Content>
  </j.1:Text>
  <j.1:Result rdf:about="http://.../discourse#result_instance_Weltmeister1990">
  <j.1:answerTypes>
  <j.1:Team rdf:about="http://.../discourse#reapr_team_Weltmeister1990">
  <j.1:mediaTypes rdf:resource="http://.../discourse#reapr_text_instance_Weltmeister1990">
  <j.1:mediaTypes>
  <j.1:Image rdf:about="http://.../discourse#reapr_image_instance_Weltmeister1990">
  <j.1:sourceURI> http://www.zdf.de/ZDFde/img/28/0,1886,2425724,00.jpg</j.1:sourceURI>
  </j.1:Image>
  
```

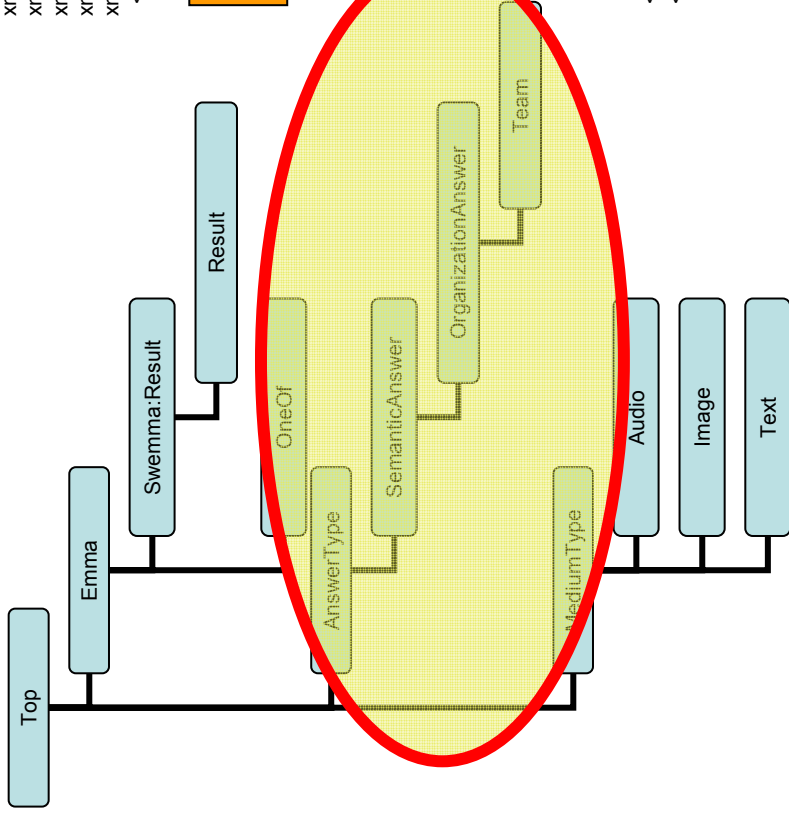
```

<j.2:Emma rdf:about="http://.../emma#emma_root_instance">
<j.2:container>
<j.0:Result rdf:about="http://.../swemma#reapr_swemma_result_instance_Wm1990">
<j.2:container>
<j.2:OneOf rdf:about="http://.../emma#reapr_oneOf_instance_1">
<j.2:container rdf:resource="http://.../discourse#result_instance_Wm1990"/>
<j.2:OneOf>
<j.2:container>
<j.0:Result>
<j.2:container>
<j.2:version>1.0</j.2:version>
</j.2:Emma>
  
```

```

</j.2:Emma>
</rdf:RDF>
  
```

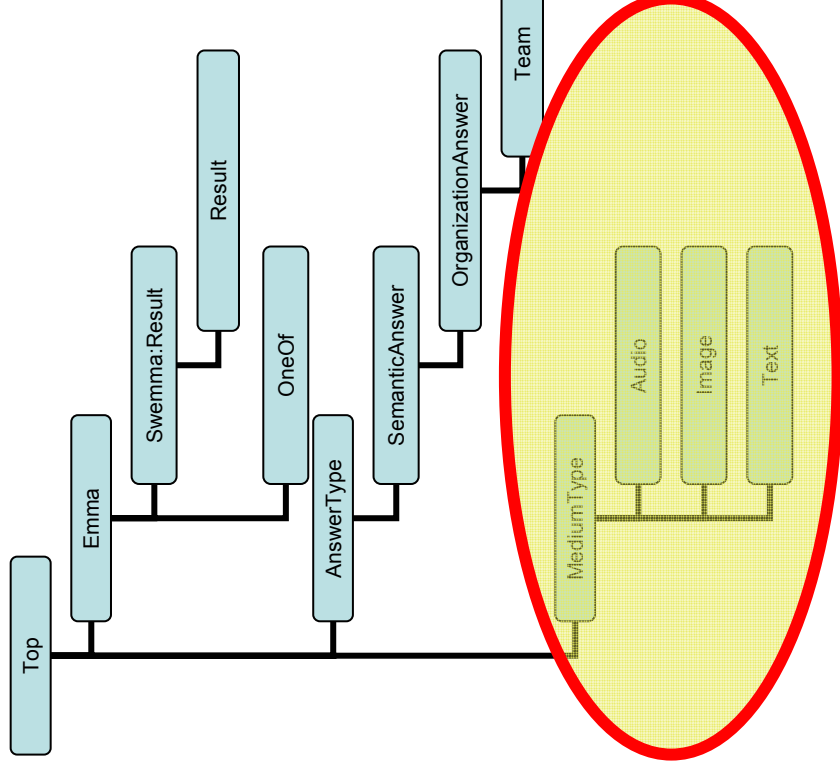
Beispiel: Wer war 1990 Weltmeister?



```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:j.0="http://.../swemma#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:j.1="http://.../discourse#"
  xmlns:daml="http://www.daml.org/2001/03/daml+oil#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:j.2="http://.../emma#">
  <j.1:Text rdf:about="http://.../discourse#reapr_text_instance_Weltmeister1990">
  <j.1:sourceURI>http://fifa.worldcup.yahoo.com/06/de/lo/octeam/emb.html</j.1:sourceURI>
  </j.1:Text>
  </j.1:Text>
  <j.1:Team rdf:about="http://.../discourse#reapr_team_wm1990">
  <j.1:Team rdf:about="http://.../discourse#reapr_team_wm1990">
  <j.1:mediaTypes rdf:resource="http://.../discourse#reapr_text_instance_Weltmeister1990">
  </j.1:mediaTypes>
  <j.1:Image rdf:about="http://.../discourse#reapr_image_instance_Weltmeister1990">
  <j.1:sourceURI> http://www.zdf.de/ZDFde/img/28/0,1886,2425724,00.jpg</j.1:sourceURI>
  </j.1:Image>
  </j.1:mediaTypes>
  <j.1:Text rdf:about="http://.../discourse#reapr_text2_instance_Weltmeister1990">
  <j.1:Content>GERMANY, I want beer.</j.1:Content>
  </j.1:Text>
  </j.1:mediaTypes>
  </j.1:Team>
  </j.1:answerTypes>
  </j.2:container>
  <j.0:Result rdf:about="http://.../swemma#reapr_swemma_result_instance_Weltmeister1990">
  </j.2:container>
  <j.2:OneOf rdf:about="http://.../emma#reapr_oneOf_instance_1">
  <j.2:container rdf:resource="http://.../discourse#result_instance_Weltmeister1990">
  </j.2:OneOf>
  </j.2:container>
  </j.0:Result>
  </j.2:container>
  <j.2:version>1.0</j.2:version>
  </j.2:Emma>
  </rdf:RDF>
  
```

Beispiel: Wer war 1990 Weltmeister?



```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:j.0="http://.../swemma#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:j.1="http://.../discourse#"
  xmlns:daml="http://www.daml.org/2001/03/daml+oil#"

```

```

<j.1:mediaTypes rdf:resource="http://.../discourse#reapr_text_instance_Wm1990"/>
<j.1:mediaTypes>
<j.1:Image rdf:about="http://.../discourse#reapr_image_instance_Wm1990">
<j.1:sourceURI> http://.../img/28/0,1886,244,00.jpg </j.1:sourceURI>
</j.1:Image>
</j.1:mediaTypes>
<j.1:mediaTypes>
<j.1:Text rdf:about="http://.../discourse#reapr_text2_instance_Wm1990">
<j.1:Content>DÉUTSCHLAND.</j.1:Content>
</j.1:Text>
</j.1:mediaTypes>

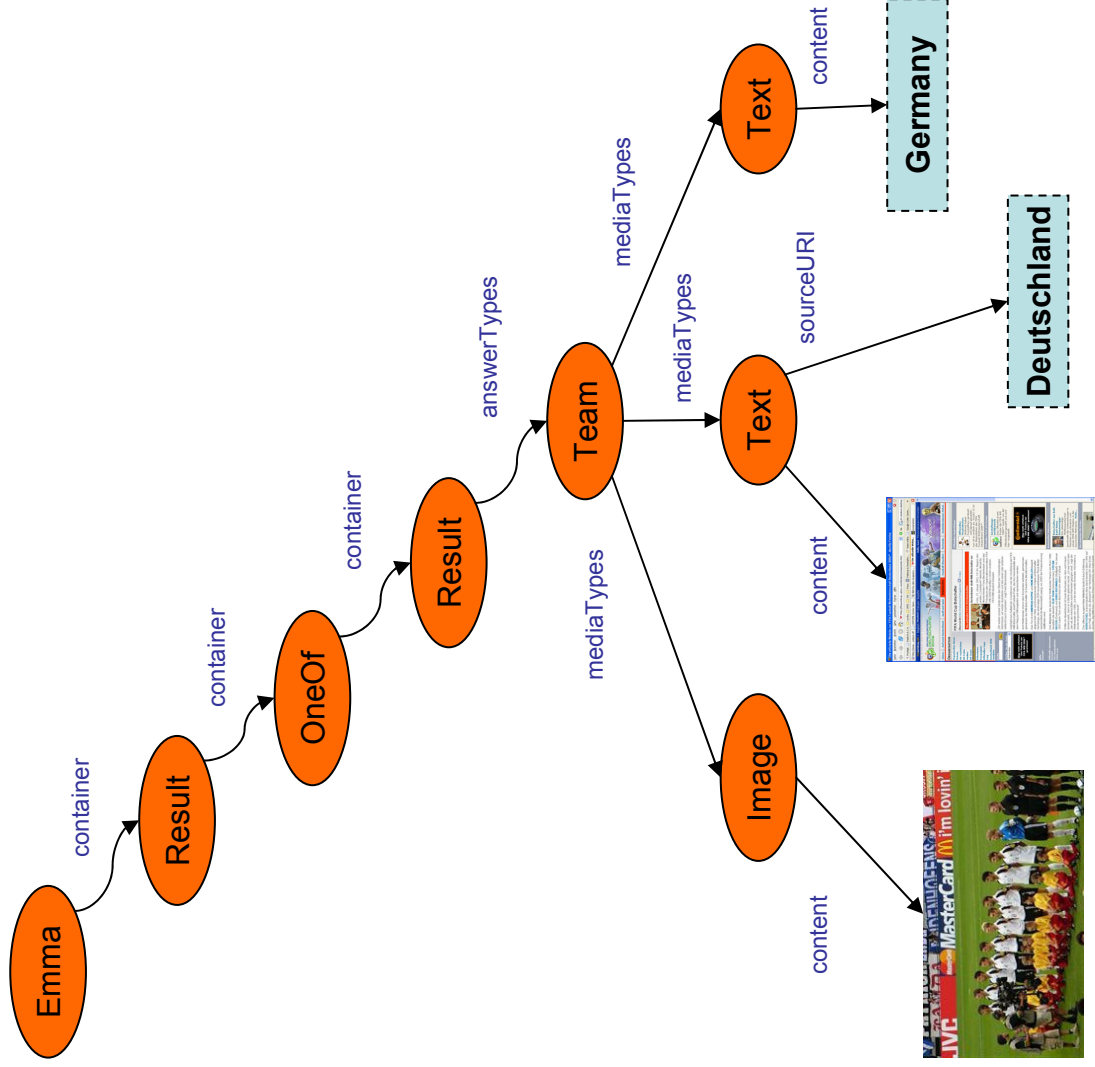
```

```

</j.1:mediaTypes>
</j.1:Team>
</j.1:answerTypes>
</j.1:Result>
<j.2:Emma rdf:about="http://.../emma#emma_root_instance">
<j.2:container>
<j.0:Result rdf:about="http://.../swemma#reapr_swemma_result_instance_Weltmeister1990">
</j.2:container>
<j.2:OneOf rdf:about="http://.../emma#reapr_oneOf_instance_1">
<j.2:container rdf:resource="http://.../discourse#result_instance_Weltmeister1990"/>
</j.2:OneOf>
</j.2:container>
</j.0:Result>
</j.2:container>
<j.2:version>1.0</j.2:version>
</j.2:Emma>
</rdf:RDF>

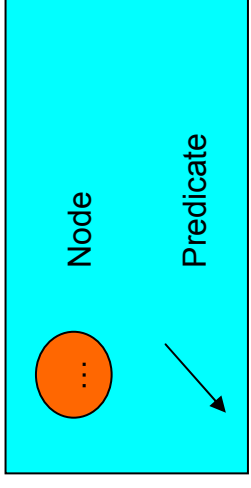
```

Beispiel als RDF-Graph



```

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:j.0="http://.../swemma#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:j.1="http://.../discourse#"
  xmlns:dami="http://www.dami.org/2001/03/dami-owl#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  >
  <j.1:Text rdf:about="http://.../discourse#reap_text_instance_Weltmeister1990">
  <j.1:sourceURI>http://ffrworldcup.yahoo.com/06/de/0octeam/emb.html<j.1:sourceURI>
  <j.1:Content>DEUTSCHLAND<j.1:Content>
  </j.1:Text>
  <j.1:Result rdf:about="http://.../discourse#result_instance_Weltmeister1990">
  <j.1:answerTypes>
  <j.1:Team rdf:about="http://.../discourse#reap_team_Weltmeister1990">
  <j.1:mediaTypes rdf:resource="http://.../discourse#reap_text_instance_Weltmeister1990">
  <j.1:mediaTypes>
  <j.1:Image rdf:about="http://.../discourse#reap_image_instance_Weltmeister1990">
  <j.1:sourceURI>http://www.zdf.de/fmgd/28/0,1886,2425724,00.jpg<j.1:sourceURI>
  </j.1:Image>
  </j.1:mediaTypes>
  <j.1:Text rdf:about="http://.../discourse#reap_text2_instance_Weltmeister1990">
  <j.1:Content>GERMANY, i want beer.<j.1:Content>
  </j.1:Text>
  </j.1:mediaTypes>
  </j.1:Team>
  </j.1:answerTypes>
  </j.1:Result>
  <j.2:Emma rdf:about="http://.../emma#emma_root_instance">
  <j.2:container>
  <j.0:Result rdf:about="http://.../swemma#reap_swemma_result_instance_Weltmeister1990">
  <j.2:container>
  <j.2:OneOf>
  <j.2:container>
  <j.0:Result>
  <j.2:container>
  <j.2:version>1.0<j.2:version>
  <j.2:Emma>
  </rdf:RDF>
  
```



Reaction And Presentation

(REAPR) bestimmt die Reaktionen des Systems:

- Legt fest, wann eine Benutzeräußerung als Anfrage akzeptiert und an den Semantic Mediator weitergeleitet wird
- Ermöglicht dem Benutzer ind laufende Dialogverarbeitung einzugreifen
- Zeitlichen Beschränkungen für Systemaktionen/Reaktionen!

Steuerungsregeln im iHUB:

```
(START_CALL,IHUB)-->(REAPR)
(NEW_TURN,REAPR)-->(REAPR)
(START_RECOGNITION,REAPR)-->(MMR)
(RECOGNITION_FINISHED,MMR)-->(REAPR)
(PROCESSING_ERROR,MMR)-->(REAPR)
(JENA,MMR)-->(SPIN,REAPR)
(JENA,SPIN)-->(REAPR)
(JENA,REAPR)-->(SEMANTIC_MEDIATOR)
(JENA,SEMANTIC_MEDIATOR)-->(REAPR)
(NO_RESULT,SEMANTIC_MEDIATOR)-->(REAPR)
(NO_RESULT,SPIN)-->(REAPR)
(NOTHING_SAID,SPIN)-->(REAPR)
(TEXT,REAPR)-->(NIPSGEN)
(TEXT,NIPSGEN)-->(SYN)
(COMMIT,REAPR)-->(*)
```

- Ergebnisse der Anfrage
 - Stammen aus unterschiedlichen Quellen
 - Konfidenzen vom Quellprozess und dem Semantic Mediator
- Derzeitige Auswahlkriterien
 - Konfidenz
 - Antworttype und Dialogakt
 - Medientyp
- Daten werden in gepacktem XML Format an den Client-Controller übertragen
- Darstellung voll parametrisierbar über LayoutXML
- Flash/Actionscript
 - Parst LayoutXML und Daten
 - Stellt Inhalte dar

Zusammenfassung der Arbeiten in den einzelnen Arbeitspaketen

Benutzerschnittstelle und der Dialogverarbeitung

- Ziel
 - Design der Benutzeroberfläche und der Interaktionsmethapher
- Ansatz
 - Szenariobasierte Entwicklung
- Stand der Arbeiten
 - Antragsgemäß abgeschlossen
 - Designdokument liegt vor
 - Oberfläche in der ersten Version implementiert
- Nächste Schritte
 - Inkrementelle Weiterentwicklung der realisierten Oberfläche im Projektverlauf

AP 2.2 Modalitätsspezifische

Analyse, Integration und Fusion

- Ziel
 - Realisierung der modalitätsspezifischen Interpretation
- Ansatz
 - Weiterentwicklung der Ansätze im Bereich multimodaler Analyse und Fusion am DFKI
- Stand der Arbeiten
 - Austauschformate definiert
 - Sprachinterpretation integriert im Demonstrator
 - Vorarbeiten für die Fusion geleistet
- Nächste Schritte
 - Analyse für offene Domänen (OOV-Verarbeitung)
 - On-/Off-Talk Unterscheidung berücksichtigen
 - Erweiterung der Multimodalen Eingaben (Stift)

AP 2.3 Multimodale Dialogverarbeitung

- Ziel
 - Entwicklung der Dialogsteuerung
- Ansatz
 - Dialogführung und Präsentation eng verzahnt
 - Multimodales Diskursmodell zur Analyse kontextspezifischer Phänomene
- Stand der Arbeiten
 - Version 0.1 der Dialogsteuerung implementiert
 - Gekapselte Kommunikation mit dem Semantic Mediator
 - Vorarbeiten für die Verarbeitung komplexer kontextbezogener Anfragen
- Nächste Schritte
 - Integration des Diskursmodells
 - Adaptives Dialogmanagement

- Ziel
 - Konzeption und Realisierung der Interaktionsmethapher für das mobile Endgerät
- Ansatz
 - Szenariobasierter Dialogentwurf
 - Portable Realisierung mit Java Controller und Flash-Präsentation
- Stand der Arbeiten
 - Realisiert im Demonstrator 0.1 (Basisversion)
- Nächste Schritte
 - Inkrementelle Erweiterung der Präsentation
 - Benutzerintervention integrieren
 - Anpassung an MDA IV



- English speech recognizer will use Sympalog's recognizer
 - Easy integration in the Smartweb system
 - Acoustic/language modeling, and OOV handling started but delayed slightly due to a manpower issue now resolved
 - English speech recognition will run in an English Milestone 1 system in August 2005 as planned
 - Tuning to raise recognition accuracy later in 2005
- Parser and text generation language conversion (ICSI/DFKI) starts after German system Milestone 1
- English speech synthesis (IMS) integrated by the end of June

ICSI personnel change

- Manny Rayner left our team due to other commitments
- Adam Janin joined our team

AP 1.4.3 Haptik-Interpretation

- Ziel
 - Anbindung der Motorradchnittstelle
- Ansatz
 - Anpassung und Erweiterung des Frameworks des mobilen Endgeräts
- Stand der Arbeiten
 - Erste Absprachen über die technischen Randbedingungen
- Nächste Schritte
 - Workshop DFKI – BMW im Juli 2005

AP 4.2.1.3 Diskursontologie

- Ziel
 - Erstellung und Verwaltung der ontologischen Strukturen für die Dialogverarbeitung
- Ansatz
 - Verwendung der SmartWeb-Standards zur Erstellung und Integration
- Stand der Arbeiten
 - Erste Version der Ontologie im System integriert
 - RDF/S-Formulierung von SWEMMA
- Nächste Schritte
 - Darstellung von zeitbezogenen Konzepten
 - Entwicklungsbegleitende Erweiterungen und Ergänzungen

AP 7.1.5 Verteilte Architektur der Dialogverarbeitung

- Ziel
 - Entwicklung der verteilten Architektur der Dialogverarbeitung und des Integrationsrahmens
- Ansatz
 - iHUB: Hub-and-Spoke Integrationsplattform für das Serversystem
 - Portable Java/Flash Lösung für das Endgerät
- Stand der Arbeiten
 - Version für den Demonstrator 0.1 realisiert
- Nächste Schritte
 - Konstante Weiterentwicklung der Architektur
 - Integration eines lokalen VoiceXML – zentrierten Dialogsystems auf den Endgeräten

Vielen Dank für Ihre Aufmerksamkeit!