Talk and Look: Tools for Ambient Linguistic Knowledge

talk & look

Ivana Kruijff-Korbayová

Information Society
Technologies

tALK

The problem

- New technologies **should** make life easier.

- *But* ...

  our high-tech environment makes ever greater demands on people

- *We need* …

  natural conversational interaction instead of complex controls and operating instructions

- Natural communication between humans and complex devices. Focused on:

  - *content* – the user says *what* he/she wants.

  - *flexibility* – the user says it in *the way* he/she wants, instead of pressing buttons or using specific commands which the device "expects".

  - *adaptivity* – the system *adapts* to the knowledge and the ability of the user as well as to the context.

  - *learning* – the system uses dialogue strategies that it has *learnt* from experience, and continues to learn from its interactions with users.

- Generic design of dialogue systems – automatic reconfiguration of specific dialogue interfaces.

- Modality-independent representations – but modality-specific, tailored realisations.

- Dialogue systems which adapt to users and learn from conversation experiences.

- 6th Framework: IST / Multimodal Interfaces

- Budget: 5,7 Mio. € (4,4 Mio. € EU-funding)

- Duration: 1/1/2004 – 31/12/2006

- Consortium: 9 partners

- Coordinator: Saarland University

- Scientific coordinator: Edinburgh University

Background: state of the art

- Driven by the speech community, lacking expertise in high-level processing: reasoning, planning, KR.

- Conflates domain, task, and dialogue knowledge in a single representation:

  - 100s of domain-specific dialogue management rules (e.g., MIT and CMU dialogue systems).

  - Portability limited to reuse of dialogue components (e.g. "get date").

  - "Voice-buttons" and form-filling rather than flexible dialogue (e.g. Voice XML)

  - Limited to application-specific interaction.

- Little user adaptation or learning.

ISU theory: Information State Updates
for dialogue context management

TALK

Multimodal dialogue
management

Reinforcement Learning
of dialogue strategies
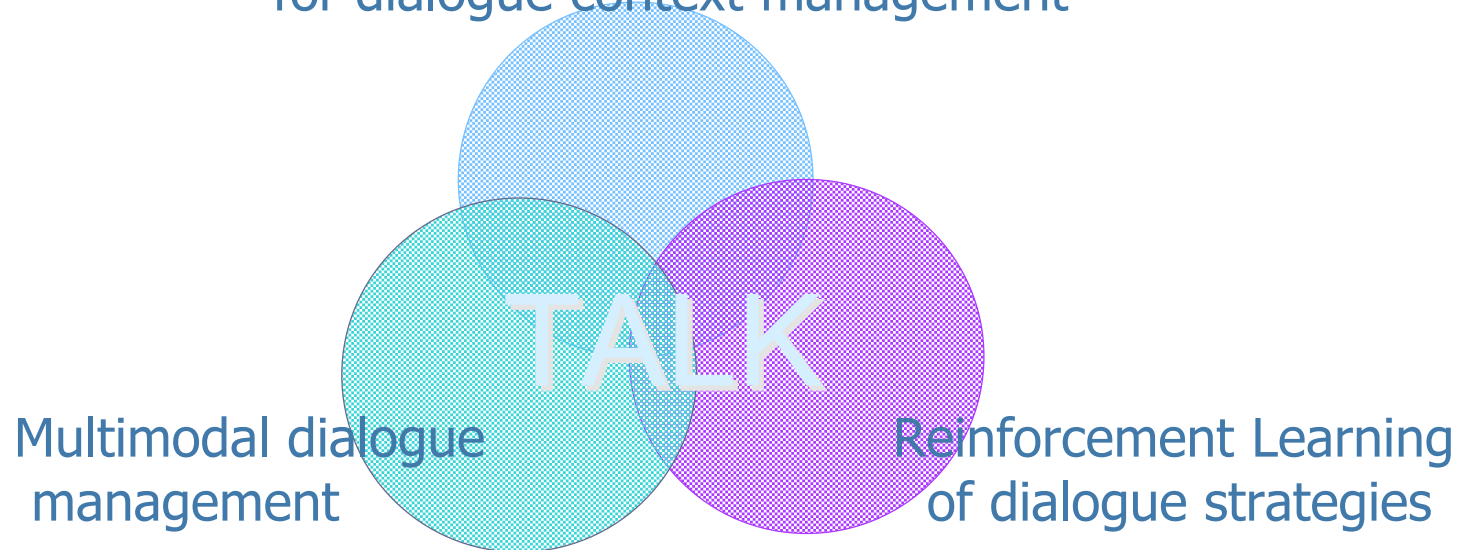
- Trindi/Siridus "information-state update" approach:

    - generic representation of dialogue states

    - flexible dialogue made possible by rich structure

    - But:

        - Task and domain information is encoded in dialogue plans

        - Output is limited to a single modality

        - No adaptivity or learning

- Potential for theoretical and practical advances:

    - The TALK research work packages ….

Four Research Themes

ISU theory: Information State Updates
for flexible dialogue management

Automatic generation and
reconfiguration
of multimodal interfaces

Unifying
Multimodality and multilinguality

Multimodal presentation
in the ISU approach

Learning and adaptivity
of dialogue strategies

Task domains: in-car, smart home
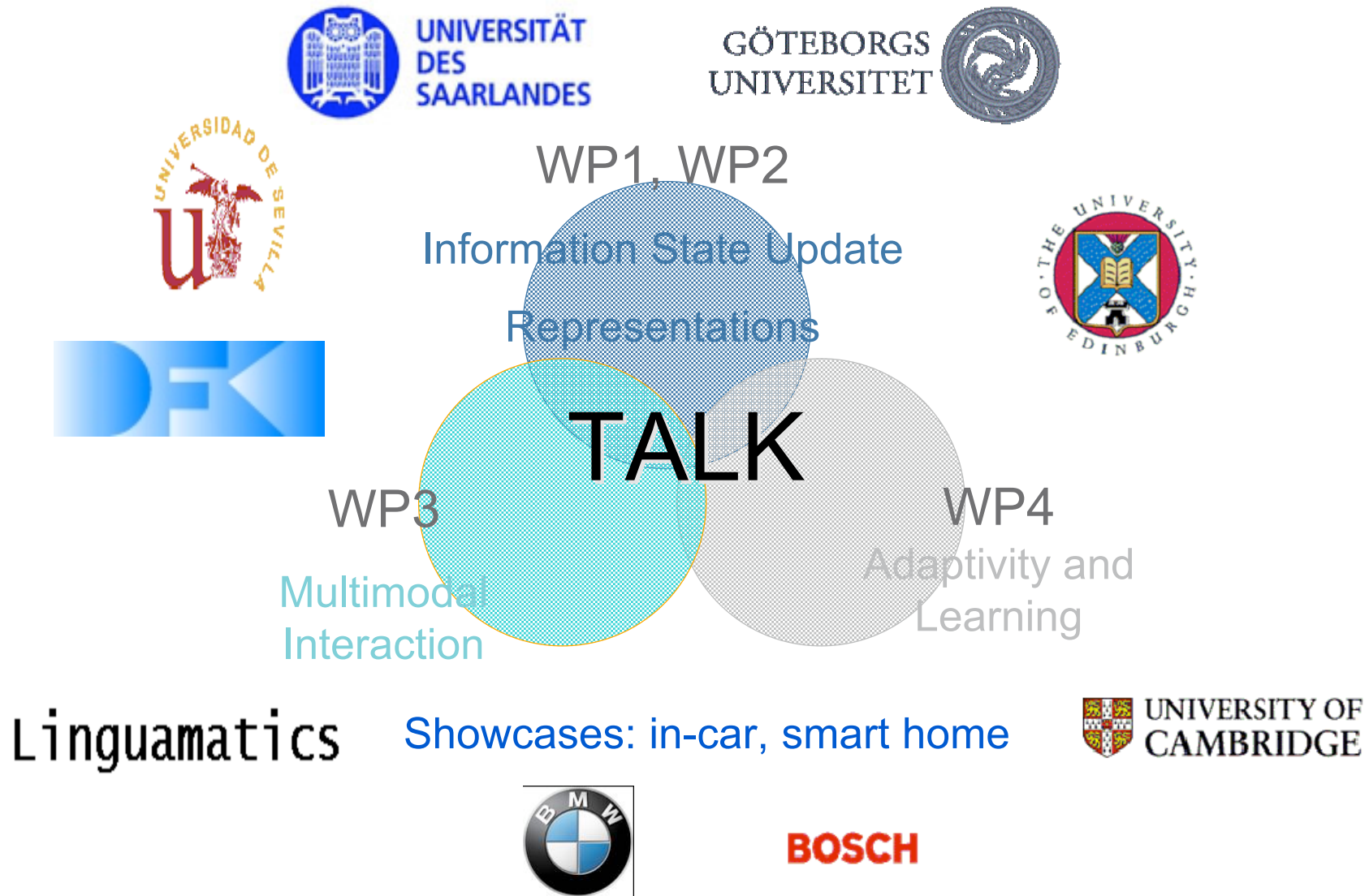
Unifying multimodality and multilinguality

- Develop an abstract representation of information which is modality independent.

- Develop criteria for presentation of information in different modalities.

- Extending statistical Language Models to robustly map multimodal inputs into internal representations in the ISU approach.

- Reconfiguration by "plugging in" task and domain descriptions.

- Can we reuse existing domain ontologies?

- Explore the suitability of different knowledge representations for generation of multimodal dialogue systems.

- Plug-and-play technology for devices and services.

- Explore the relationship between domain processes and dialogue processes.

Multimodal presentation in the ISU approach

- Generating user-tailored textual, tabular, or graphical presentations of information.

- Composite tailored and adaptive multimodality

- For each user, task and situation,

  - What information should be presented?

  - What are the best modes of presentation?

  - How to best realize it?

- What is the best abstract representation of information committed to during a dialogue?

Adaptivity and learning

- Adapting to different users – their knowledge and preferences.

- Multiple dialogue strategies available to the system, chosen depending on context.

- Reinforcement Learning applied to the problem of automatic strategy optimization.

- What representations are most suited to adaptivity and learning?

- What reward functions can be developed for learning about dialogue management?

UNIVERSITÄT DES SAARLANDES

GÖTEBORGS UNIVERSITET

UNIVERSIDAD DE SEVILLA

THE UNIVERSITY OF EDINBURGH

DFKI

WP1, WP2

Information State Update

Representations

TALK

WP3

WP4

Multimodal Interaction

Adaptivity and Learning

Linguamatics

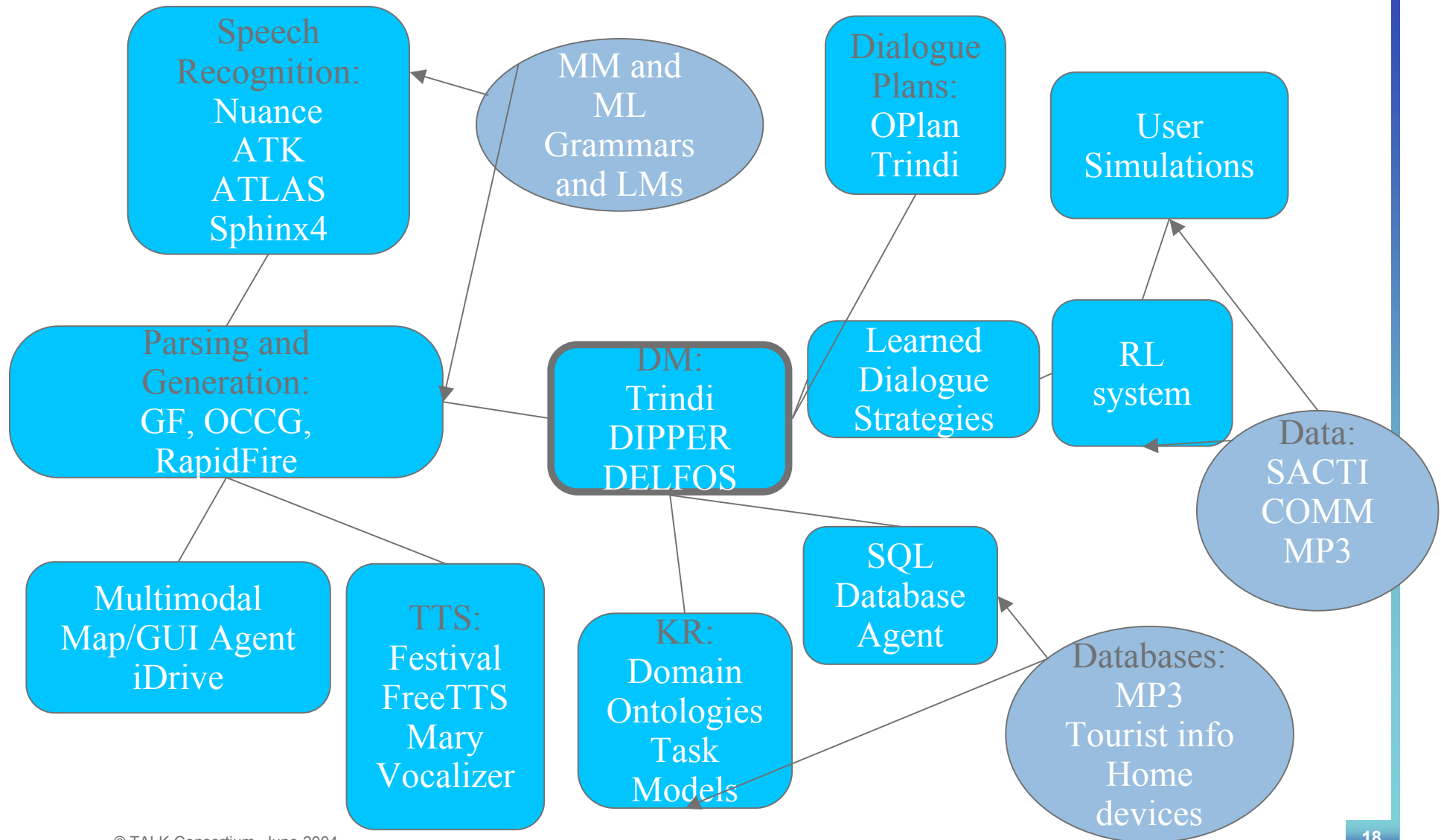Showcases: in-car, smart home

UNIVERSITY OF CAMBRIDGE

BMW

BOSCH

Software results

- "In-car" and "smart-home" application domains.

  - Multiple tasks, multiple modalities, multiple users.

- Not one single system:

  - varied applications provide challenges to the theory and are needed to test domain portability and reconfigurability.

Speech Recognition: Nuance ATK ATLAS Sphinx4

MM and ML Grammars and LMs

Dialogue Plans: OPlan Trindi

User Simulations

Parsing and Generation: GF, OCCG, RapidFire

DM: Trindi DIPPER DELFOS

Learned Dialogue Strategies

RL system

Data: SACTI COMM MP3

Multimodal Map/GUI Agent iDrive

TTS: Festival FreeTTS Mary Vocalizer

KR: Domain Ontologies Task Models

SQL Database Agent

Databases: MP3 Tourist info Home devices

- Methods for designing better – flexible and adaptive – dialogue systems that learn from interactions with users, based on ISU technology

- Methods for rapid and cost-effective deployment of new dialogue applications through reconfigurable dialogue systems:

  - Separating domain-specific information from generic communicative behaviour

  - Separating central aspects of dialogue structure from modality- and language-specific realisation

User categories

Industrial users

- ➢ Developers of dialogue applications who need better dialogue design and more efficient development techniques

- ➢ Information about industrial user requirements through feedback by industrial partners in the consortium

End users

- ➢ anybody with a car or a home

- ➢ Information about end user requirements through

  - ➢ Industrial partners

  - ➢ Market studies (e.g. JDPower & Assoc, Cisco Systems)

  - ➢ Project-internal wizard-of-oz experiments and usability tests

Subsidiary exploitable resuls

- Reusable software tools (e.g. TrindiKit, ATK, GF, RapidFire, Reinforcement Learner, User Simulator, Automatic ISU annotator)

- Reusable annotated data archives and databases of application-specific knowledge, e.g., corpora:

  - SACTI 1 and 2 (UCAM)

  - MP3 WoZ corpus (USAAR)

  - ISU-annotated COMMUNICATOR data (UEDIN)

- Design and methodology for conducting multimodal wizard-of-oz experiments

- Contributions to standards (W3C, ISO)

- Skills development: training PhD and Master students

Multimodal Presentation

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES

Multimodal Presentation

- Use the structured representations of shared knowledge and commitments accumulated in the information state for advanced multimodal presentation of system output to facilitate easy and efficient interaction, adapted to dialogue context, situation, user and available modalities

... Information State$_i$ → Turn$_i$ → Information State$_{i+1}$ ...

Multimodal turn planning& realization
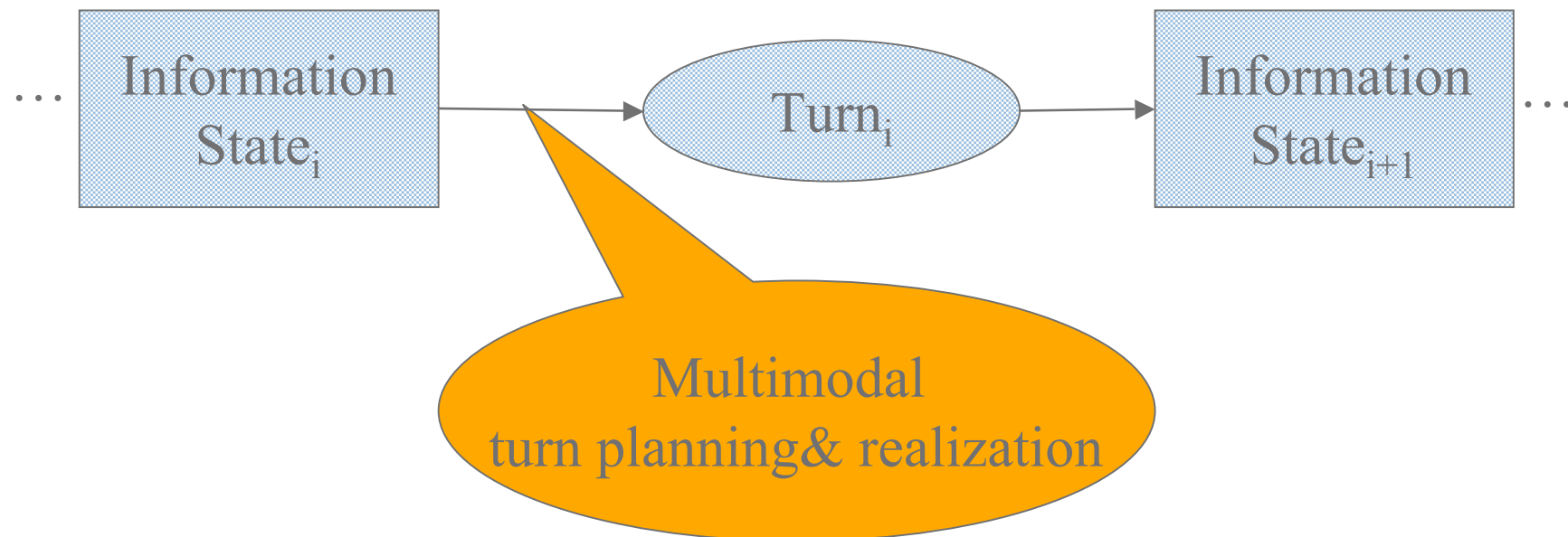
Goals:

- Gather pilot data on human multi-modal turn planning

- Collect wizard dialogue strategies

- Collect wizard media allocation decisions

- Collect wizard speech data

- Collect user data (speech signals and spontaneous speech)

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES

- MP3 domain

    - "in-car" with primary task Lane Change Task (LCT)

    - "in-home" domain without LCT

- Tasks for the subject:

    - Play a song from the album "New Adventures in Hi-Fi" by REM

    - Find a song with "believe" in the title and play it.

    - Make a playlist with 4 of your favorite songs

- Task for the wizard:

    - Help the user reach their goals
      (Deliberately vague!)

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES

A Walk Through the Final Turns

- Wizard:

  - Database search

  - Select "album presentation" (vs. songs or artists)

  - Select "list presentation" (vs. tables or textual summary)

  - "Ich zeige Ihnen die Liste an."
    I am displaying the list.

  - Audio is sent to typist

  - Text is sent to speech synthesis

- User: "Ok. Zeige mir bitte das Lied aus dem ausgewählten Album und spiel das vor."
  Ok. Please show me that song ("Believe") from the selected album and play it.

Wizard
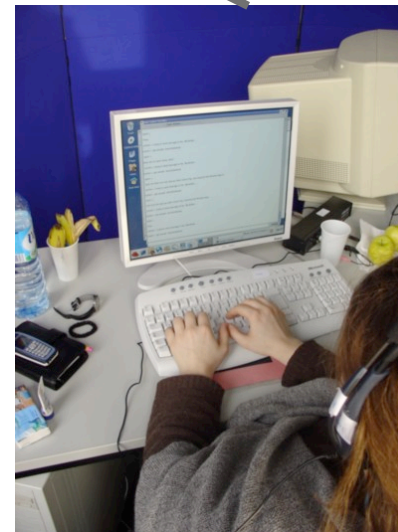
Subject

graphics

text

synthesized
audio data

audio data

audio data

Typist

Typist

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES

Wizard: "Ich zeige Ihnen die Liste an."
I am displaying the list.

User: "Ok. Zeige mir bitte das Lied aus dem
ausgewählten Album und spiel das vor."
*Ok. Please show me that song ("Believe") from the
selected album and play it.*

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES

- Primary task: driving

- Secondary task on second screen: MP3 player

Example(1) Wizard

- says: "Ich zeige Ihnen die Liste an." and clicks on the list presentation

I am displaying the list.

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES

> in Ordnung .

> .

> ja in Ordnung .

> .

File   Edit   Search   Playlists   Help

Albums   Playlists

1. Genre	2. Artist	3. Album	4. Title	5. Album or Title	6. Year	
			Believe		2004 –	✓

Search completed in 0.15s   🔍   🗑

Albums (15 matches)

	Genre	Artist	Album	Year	Tracks	Length	
1	Dance	Varios	20 Historicos Del Dance	2004	20	1:17:12	⊕
2	Eurodance	Dream	Eternal Dream Disk2	2002...	21	1:14:00	⊕
3	Rock & Roll	University	Make Believe (Side B)	2004	3	11:33	⊕
4	Easy Listening	Dusty Springfield	The Look Of Love (Disc 2)	2004	25	1:17:23	⊕
5	Pop	Falling You	Touch	2004	10	1:07:45	⊕

Display

Tracks (17 matches)

	Genre	Artist	Album	Year	Title	Length	Track	
1	Dance	Varios	20 Historicos Del Dance	2004	Ministers Dela Funk Feat. Jocelyn Brown – Believe	4:08	14	⊕
2	Eurodance	Dream	Eternal Dream Disk2	2002...	Believe In You "Original Mix"	2:17	7	⊕
3	Rock & Roll	University	Make Believe (Side B)	2004	Make Believe	4:41	3	⊕
4	Easy Listening	Dusty Springfield	The Look Of Love (Disc 2)	2004	In The Land Of Make Believe	2:31	22	⊕
5	Pop	Falling You	Touch	2004	Less Likely To Believe	8:06	4	⊕

Display

Artists (14 matches)

	Artist	Albums	Tracks	
1	Tandy, Sharon	1	26	
2	Dusty Springfield	1	25	
3	Red Star Belgrade	1	11	
4	Heidi Marie Vestrheim	1	11	
5	Carpenters	2	40	

Display

PresentationOptions

Wizard Screen | **User Screen**

Es wurden 15 Alben ab 2004 mit Songtitel "Believe" gefunden!

#	Album	Genre	Künstler	Jahr	Tracks	Selected
1	20 Historico…	Dance	Varios	2004	20	☐
2	Eternal Drea…	Eurodance	Dream	20020626	21	☐
3	Make Believ…	Rock & Roll	University	2004	3	☐
4	The Look Of…	Easy Listening	Dusty Spring…	2004	25	☐
5	Touch	Pop	Falling You	2004	10	☐
6	You Gotta B…	Soul	Tandy, Sharon	2004	26	☐
7	Heart Of Th…	Christian Rock	Jeremy Camp	2004	6	☐
8	The Best Hits	Alternative …	Nickelback	2004	21	☐
9	The Real Tr…	Rock	Red Star Bel…	2004	11	☐
10	Gold: 35Th …	Pop	Carpenters	2004	20	☐
11	Gold: 35Th …	Pop	Carpenters	2004	20	☐
12	Just The Bes…	Pop	Sampler	2004	20	☐
13	Signs And Fi…	Folk	Heidi Marie …	2004	11	☐
14	The Chillout …	Chillout	Ministry Of S…	2004	19	☐
15	Sardonic	Rock	Iron Rain	2004	15	☐

#	Album	Künstler	Jahr	Selected
1	20 Historicos Del …	Varios	2004	☐
2	Eternal Dream Disk2	Dream	20020626	☐
3	Make Believe (Side…	University	2004	☐
4	The Look Of Love (…	Dusty Springfield	2004	☐
5	Touch	Falling You	2004	☐
6	You Gotta Believe I…	Tandy, Sharon	2004	☐
7	Heart Of The Artist	Jeremy Camp	2004	☐
8	The Best Hits	Nickelback	2004	☐
9	The Real Traitors	Red Star Belgrade	2004	☐
10	Gold: 35Th Annive…	Carpenters	2004	☐
11	Gold: 35Th Annive…	Carpenters	2004	☐
12	Just The Best Vol. …	Sampler	2004	☐
13	Signs And Fiction	Heidi Marie Vestrh…	2004	☐
14	The Chillout Sessio…	Ministry Of Sound	2004	☐
15	Sardonic	Iron Rain	2004	☐

☐	20 Historicos Del Dance	1
☐	Eternal Dream Disk2	1
☐	Gold: 35Th Anniversary Edition (…	1
☐	Gold: 35Th Anniversary Edition (…	1
☐	Heart Of The Artist	1
☐	Just The Best Vol. 47 / Cd1	1
☐	Make Believe (Side B)	1
☐	Sardonic	1
☐	Signs And Fiction	1
☐	The Best Hits	1
☐	The Chillout Session – Summer C…	1
☐	The Look Of Love (Disc 2)	1
☐	The Real Traitors	1
☐	Touch	1
☐	You Gotta Believe It's…	1

Clear User Screen

Wizard Screen | **User Screen**

☐	20 Historicos Del Dance	1
☐	Eternal Dream Disk2	1
☐	Gold: 35Th Anniversary Edition (Disc 1)	1
☐	Gold: 35Th Anniversary Edition (Disc 2)	1
☐	Heart Of The Artist	1
☐	Just The Best Vol. 47 / Cd1	1
☐	Make Believe (Side B)	1
☐	Sardonic	1
☐	Signs And Fiction	1
☐	The Best Hits	1
☐	The Chillout Session – Summer Collection 2004 Cd2	1
☐	The Look Of Love (Disc 2)	1
☐	The Real Traitors	1
☐	Touch	1
☐	You Gotta Believe It's...	1

Example(2) WizardTypist

- types the wizard's spoken text

I am displaying the list.

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES

Albums

Artist	Album
Varios	20 Historicos Del Dance
Dream	Eternal Dream Disk2
University	Make Believe (Side B)

Wizard Typist Text-Box

(--------------> Panic Button <--------------)

typist >

```
<?xml version="1.0" encoding="UTF-8" ?>
<maryxml version="0.4" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns="http://mary.dfki.de/2002/MaryXML" xml:lang="de">
<p>
Dies ist nicht
richtig
geschrieben.
.
</p>
</maryxml>
```

system > trying to send message to a synthesizer module ...

s

e Latte Mix)

Dieser Satz hat ein	falschexx	Wort.	.	(commit)

(1. falschexx)   (new)

```
    xmlns="http://mary.dfki.de/2002/MaryXML" xml:lang="de">
<t>ich zeige Ihnen die Liste an</t>
<t>.</t>
</maryxml>
```

system > trying to send message to a synthesizer module ...

ich zeige Ihnen die Liste an   commit

Example(3) User

- Listens to wizard text synthesized by Mary and receives the selected list presentation

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES

☐	20 Historicos Del Dance	1
☐	Eternal Dream Disk2	1
☐	Gold: 35Th Anniversary Edition (Disc 1)	1
☐	Gold: 35Th Anniversary Edition (Disc 2)	1
☐	Heart Of The Artist	1
☐	Just The Best Vol. 47 / Cd1	1
☐	Make Believe (Side B)	1
☐	Sardonic	1
☐	Signs And Fiction	1
☐	The Best Hits	1
☐	The Chillout Session – Summer Collection 2004 Cd2	1
☐	The Look Of Love (Disc 2)	1
☐	The Real Traitors	1
☐	Touch	1
☐	You Gotta Believe It's…	1

Example(4) User

- Selects one album and says: "Ok. Zeige mir bitte das Lied aus dem aus gewählten Album und spiel das vor."

Ok. Please show me that song ("Believe") from the selected album and play it.

☐	20 Historicos Del Dance	1
☐	Eternal Dream Disk2	1
☐	Gold: 35Th Anniversary Edition (Disc 1)	1
☐	Gold: 35Th Anniversary Edition (Disc 2)	1
☑	Heart Of The Artist	1
☐	Just The Best Vol. 47 / Cd1	1
☐	Make Believe (Side B)	1
☐	Sardonic	1
☐	Signs And Fiction	1
☐	The Best Hits	1
☐	The Chillout Session – Summer Collection 2004 Cd2	1
☐	The Look Of Love (Disc 2)	1
☐	The Real Traitors	1
☐	Touch	1
☐	You Gotta Believe It's...	1

Wizard Screen **User Screen**

☐	20 Historicos Del Dance	1
☐	Eternal Dream Disk2	1
☐	Gold: 35Th Anniversary Edition (Disc 1)	1
☐	Gold: 35Th Anniversary Edition (Disc 2)	1
☑	Heart Of The Artist	1
☐	Just The Best Vol. 47 / Cd1	1
☐	Make Believe (Side B)	1
☐	Sardonic	1
☐	Signs And Fiction	1
☐	The Best Hits	1
☐	The Chillout Session – Summer Collection 2004 Cd2	1
☐	The Look Of Love (Disc 2)	1
☐	The Real Traitors	1
☐	Touch	1
☐	You Gotta Believe It's…	1

Example(5) UserTypist

- Types the user's spoken text

Ok. Please show me that song ("Believe") from the selected album and play it.

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES
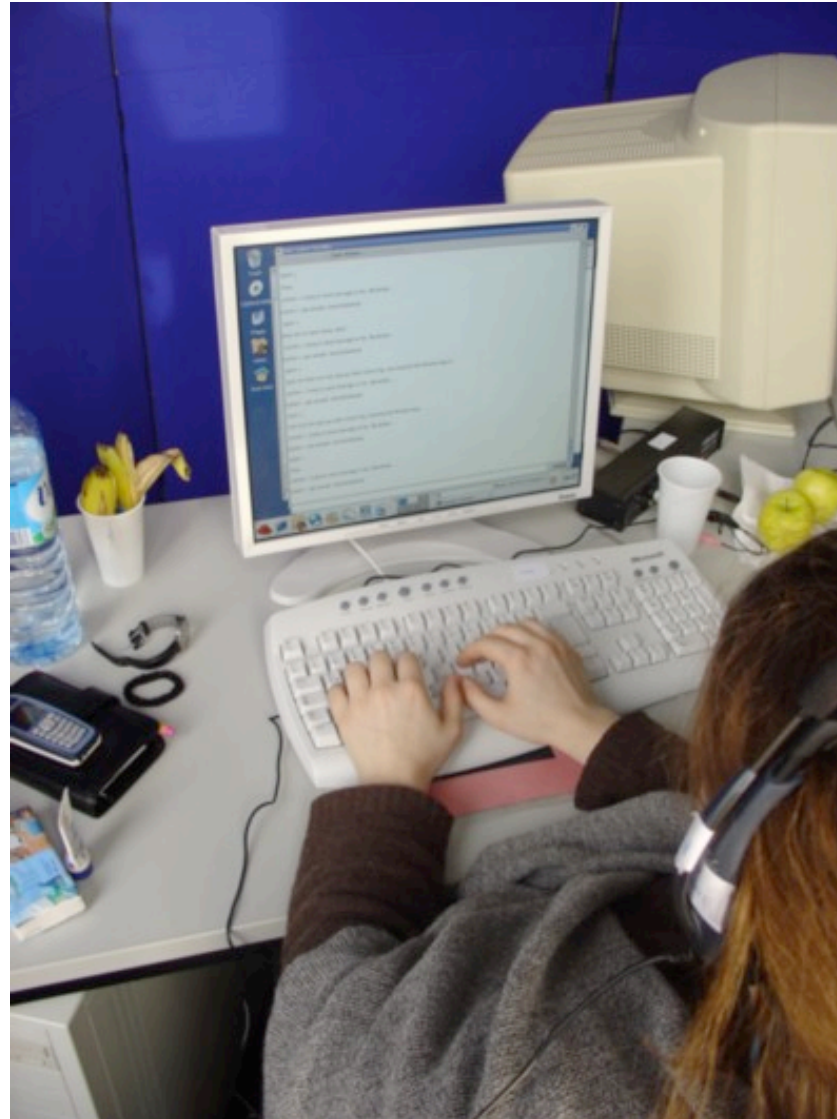
-------------- > _Panic Button_ <--------------

system > got answer: [wordsDeleted]

typist >

.

system > trying to send message to the WordKiller ...

system > got answer: [wordsDeleted]

typist >

OK .

system > trying to send message to the WordKiller ...

system > got answer: [wordsDeleted]

typist >

zeige mir bitte das Lied aus dem aus gewⁿhlten Album und spiel das vor .

system > trying to send message to the WordKiller ...

system > got answer: [wordsDeleted]

zeige mir bitte das Lied aus dem ausgewahlten Album und spiel das vor.

commit

Example(6) Wizard

- Gets a correspondingly updated TextBox Window

> Ja in Ordnung .

> .

> OK .

> zeige mir bitte das Lied aus dem aus gewählten Album und spiel das vor .

X-⊠ FreeDB Wizard: Database.export

File Edit Search Playlists Help

Albums **Playlists**

1. Genre	2. Artist	3. Album	4. Title	5. Album or Title	6. Year	
			Believe		2004 –	✓

Search completed in 0.15s

Albums (15 matches)

	Genre	Artist	Album	Year	Tracks	Length	
1	Dance	Varios	20 Historicos Del Dance	2004	20	1:17:12	⊕
2	Eurodance	Dream	Eternal Dream Disk2	2002...	21	1:14:00	⊕
3	Rock & Roll	University	Make Believe (Side B)	2004	3	11:33	⊕
4	Easy Listening	Dusty Springfield	The Look Of Love (Disc 2)	2004	25	1:17:23	⊕
5	Pop	Falling You	Touch	2004	10	1:07:45	⊕

Display

Tracks (17 matches)

	Genre	Artist	Album	Year	Title	Length	Track	
1	Dance	Varios	20 Historicos Del Dance	2004	Ministers Dela Funk Feat. Jocelyn Brown – Believe	4:08	14	⊕
2	Eurodance	Dream	Eternal Dream Disk2	2002...	Believe In You "Original Mix"	2:17	7	⊕
3	Rock & Roll	University	Make Believe (Side B)	2004	Make Believe	4:41	3	⊕
4	Easy Listening	Dusty Springfield	The Look Of Love (Disc 2)	2004	In The Land Of Make Believe	2:31	22	⊕
5	Pop	Falling You	Touch	2004	Less Likely To Believe	8:06	4	⊕

Display

Artists (14 matches)

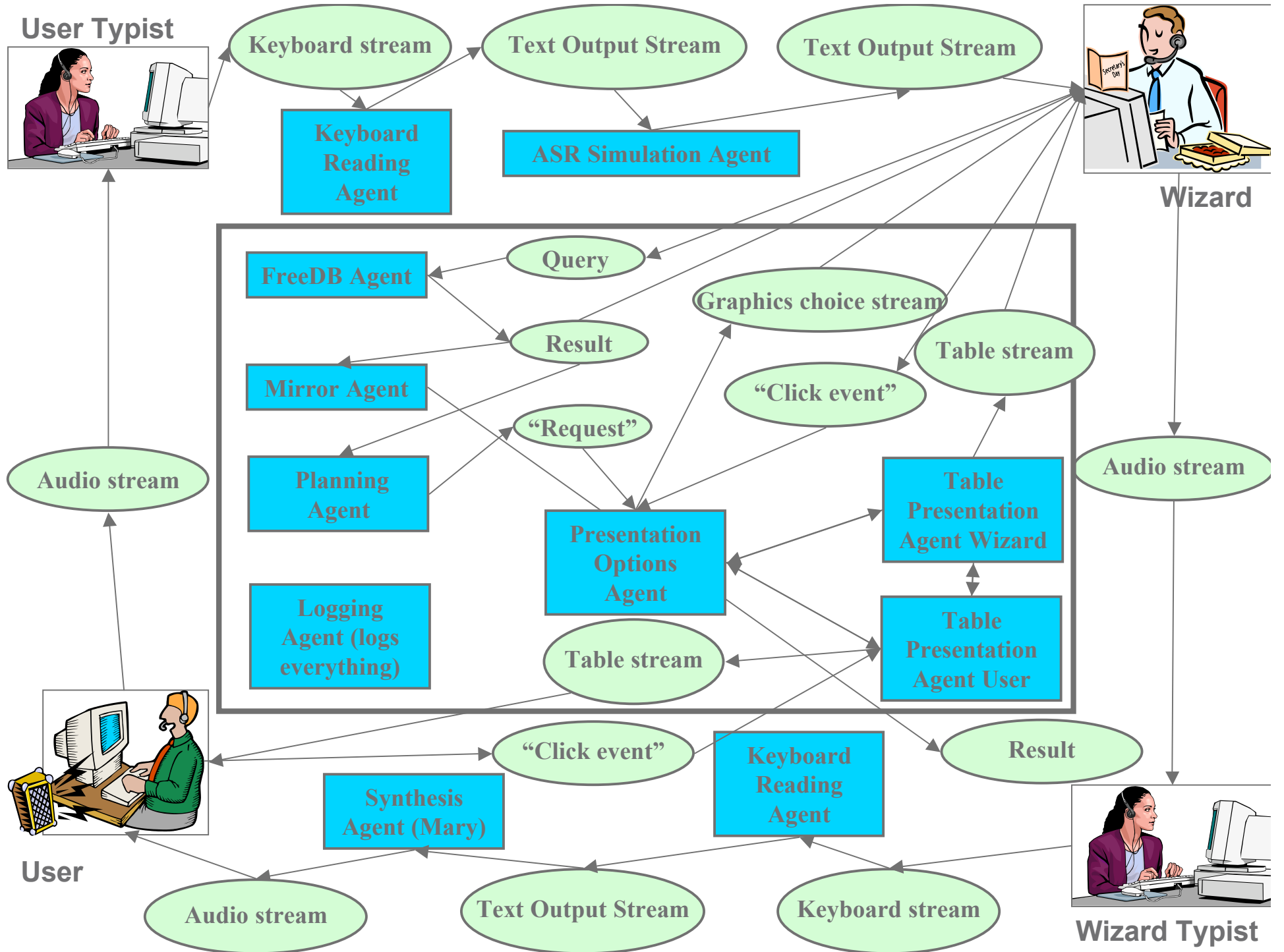|	Artist	Albums	Tracks	
1	Tandy, Sharon	1	26	
2	Dusty Springfield	1	25	
3	Red Star Belgrade	1	11	
4	Heidi Marie Vestrheim	1	11	
5	Carpenters	2	40	

Display

Wizard: "Ich zeige Ihnen die Liste an."
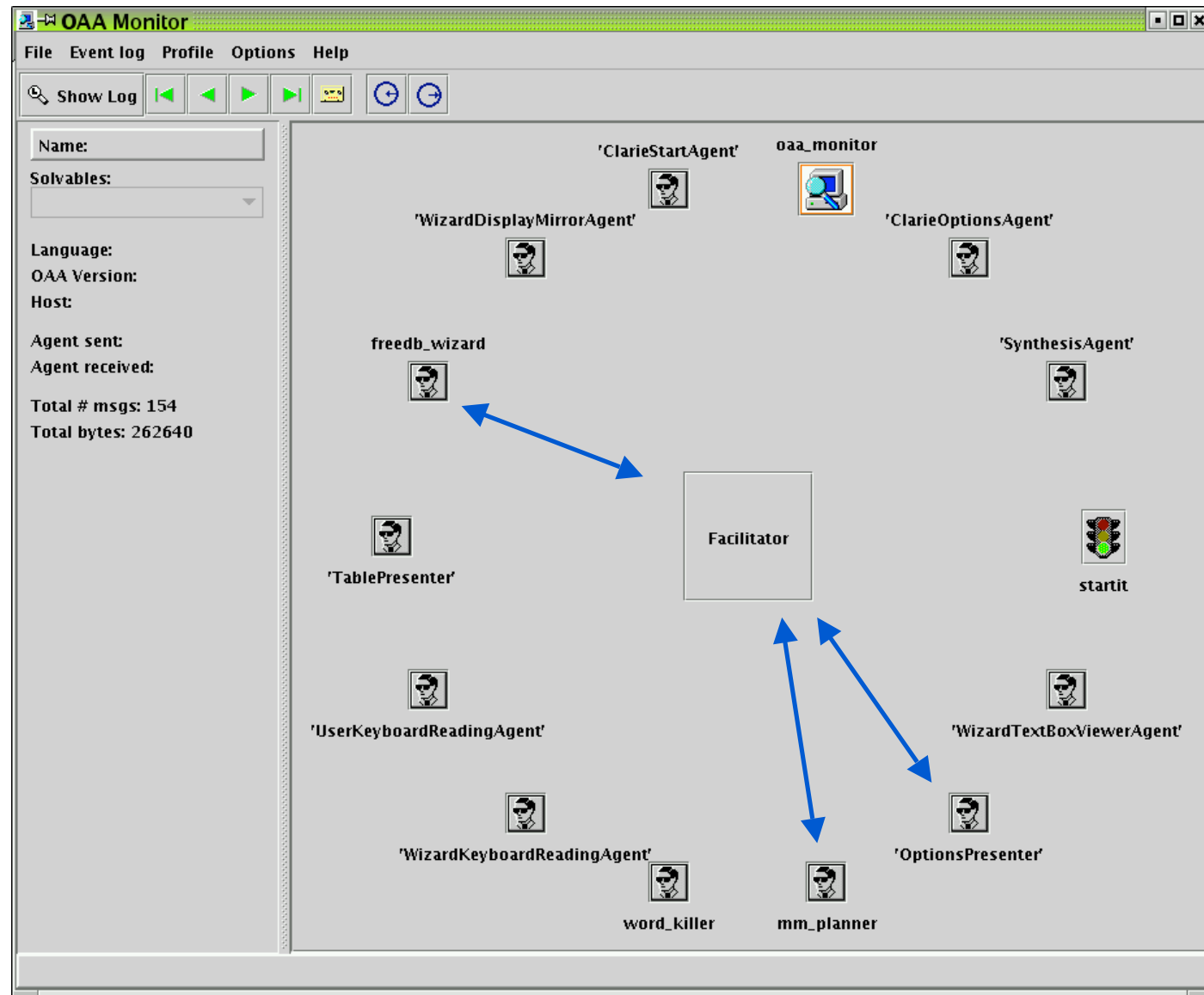*I am displaying the list.*

User: "Ok. Zeige mir bitte das Lied aus dem
ausgewählten Album und spiel das vor."
*Ok. Please show me that song ("Believe") from the
selected album and play it.*

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
SAARLANDES

Complex setup

- 5 people involved to run an experiment:

    - 1 experiment leader

    - 1 wizard

    - 1 subject

    - 2 typists

- System features:

    - 14 (via OAA) communicating components distributed over

    - 5 machines (3 windows, 2 linux)

    - Plus LCT on a separate machine

    $\Rightarrow$ robustness is a challenge

German Research Center for Artificial Intelligence GmbH

UNIVERSITÄT DES SAARLANDES

- 6 wizards , 24 subjects, each subject 1 session

  - 4 tasks, each 15 mins, 2 without and 2 with LCT

  - Tasks of 2 types (but varying specificity):

    - searching for a title in the database or in a playlist

    - building a playlist satisfying several constraints

- Video and audio recording and a log file, e.g.,

  - transcriptions of the spoken utterances

  - the wizard's database query and the number of found results

  - the type and form of the presentation screen

- cca 1800 turns, 17k words

German Research
Center for Artificial
Intelligence GmbH

UNIVERSITÄT
DES
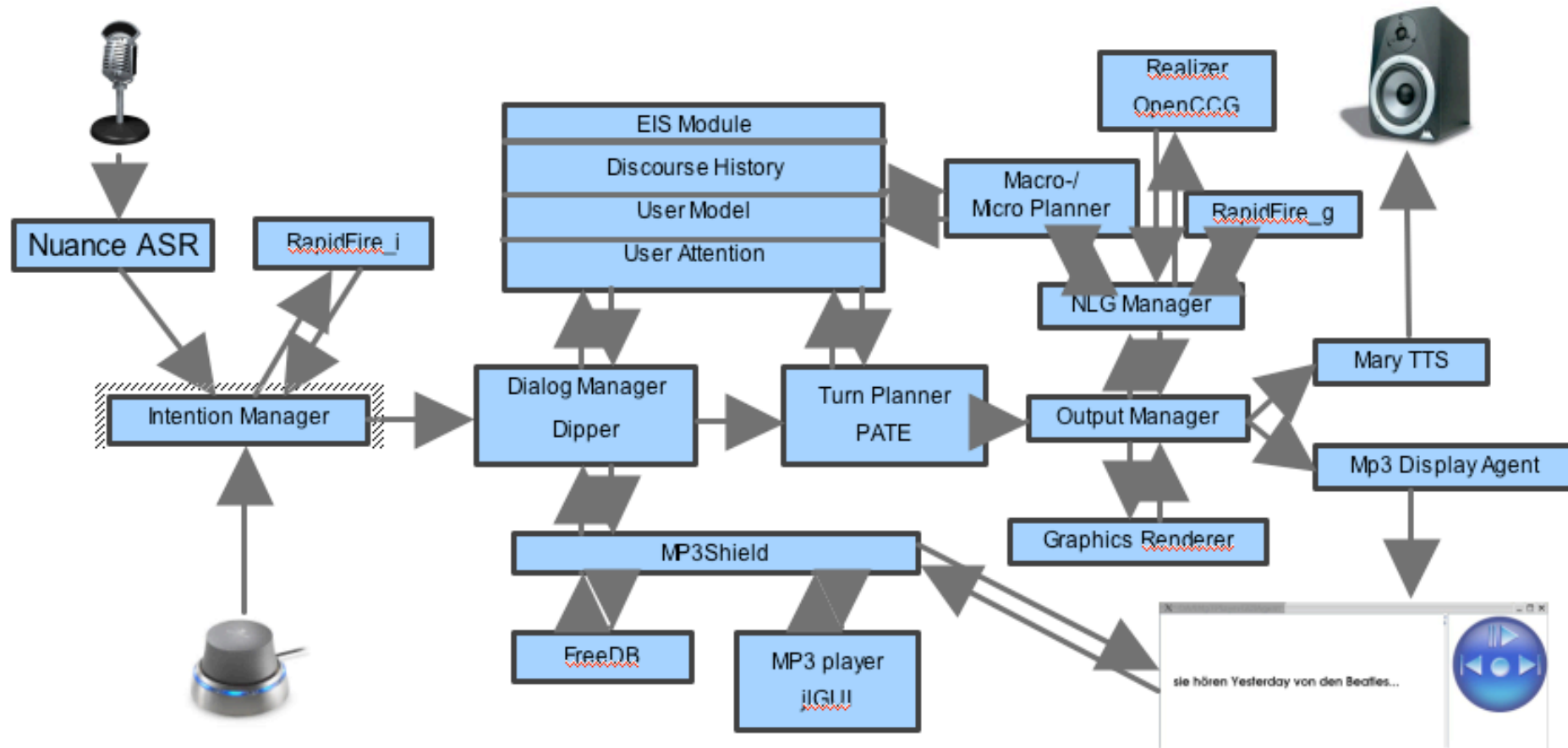SAARLANDES

Wizard interviews

- Overall understanding good, difficulties due to delays

- Text message output hardly used (0.04%)

- Differences in usefulness judgments across wizards for

    - Most detailed table (78.6%)

    - List of songs/albums + length (17.5%)

- Some wizards adapted amount to whether user was driving

- Multimodality:

    - Show + tell what showing (esp. when responding to questions, e.g. What did you find?)

    - Show only, when responding to request to show

User Interviews

- Mutual understanding good, easy error recovery

- Multimodal strategies helpful, but:

    - Wish more display feedback for disambiguation requests and grounding

    - Wish less display info and more verbal feedback when drive

- User satisfaction in terms of 5 aspects:

    - text-to-speech synthesis performance, task ease, user expertise, overall difficulty and future use

    - Result compare to average on COMMUNICATOR systems

    - No significant differences across wizards

§ Good task success and task satisfaction

Multimodal In-Car MP3 System (Baseline)

- Playback control

- Playback of songs, playlists, albums

- Queries for songs, playlists, albums

- Queries for artist, title

- Playlist manipulation: add/delete track

German Research Center for Artificial Intelligence GmbH

UNIVERSITÄT DES SAARLANDES

Thank You!

http://www.talk-project.org

Questions?