

Experiences from Verbmobil

Norbert Reithinger
DFKI GmbH
Stuhlsatzenhausweg 3
D-66123 Saarbrücken
bert@dfki.de



Content

- **Overview of Verbmobil project**
- **Scientific challenges and experiences**
- **Software technology challenges and experiences**
- **Management challenges and experiences**
- **Conclusion**



Verbmobil -The Project



Some information for those who haven't heard of Verbmobil recently

- **speaker independent speech-to-speech translation system for appointment scheduling and travel planning:**
 - German ↔ English (10 175 words German, 6871 words English)**
 - German ↔ Japanese (2566 words Japanese)**
- **69 modules, full configuration 3.5 GB**
- **23 participating institutions (in Verbmobil II)**
- **over 900 full workers and students involved**
- **project duration: 1993 - 2000**

□ scientific, software technology, and management challenges



Verbmobil - The Partners



A Remark about Project Duration

8 years is a long time, especially since the invention of Internet time

1993

- “You will need special hardware!”
- “1500 words speaker independent is impossible!”
- “Aren’t your goals unrealistic?”

2000

- “Does it run on my notebook?”
- “Only 10 000 words?”
- “Why can’t it also translate in the domains X, Y, and Z?”

but

it is a unique chance for

- **large scale, continuous research and development**
- **training people, collaborating, gaining experience**
- **collecting and annotating data**



Scientific Challenges

The goal

- **Progress in the areas of speech translation**

The situation 1993

- **Speech understanding and translation system existed**
- **Knowledge distributed in various scientific fields**
- **There were only few interactions**

The solution

- **Collaborate in a project**
- **Cross fertilization: check the theories of the others**



Scientific Results

There are over 600 refereed papers on the various aspects of and achievements in Verbmobil.

See also *W. Wahlster (ed.): Verbmobil: Foundations of Speech-to-Speech Translation, Springer Verlag, to appear July 2000* ... at any shop near your office :-)

Some highlights

- **Speaker independent speech recognition over various channels**
- **Language ID**
- **Unknown words**
- **Prosodic information (segmentation, stress etc.) used in various modules**
- **Repair of hesitations, repetitions**
- **Combination of parser analysis fragments**
- **Semantic representation: VIT**
- **Context and dialogue knowledge supports translation**
- **Efficient semantic transfer**
- **Content to speech generation**
- **Word concatenative speech synthesis**
- **Dialogue minutes and summaries**
- **Large data collection with annotation on various levels (e.g. tree-banks, dialogue acts)**
- **....**

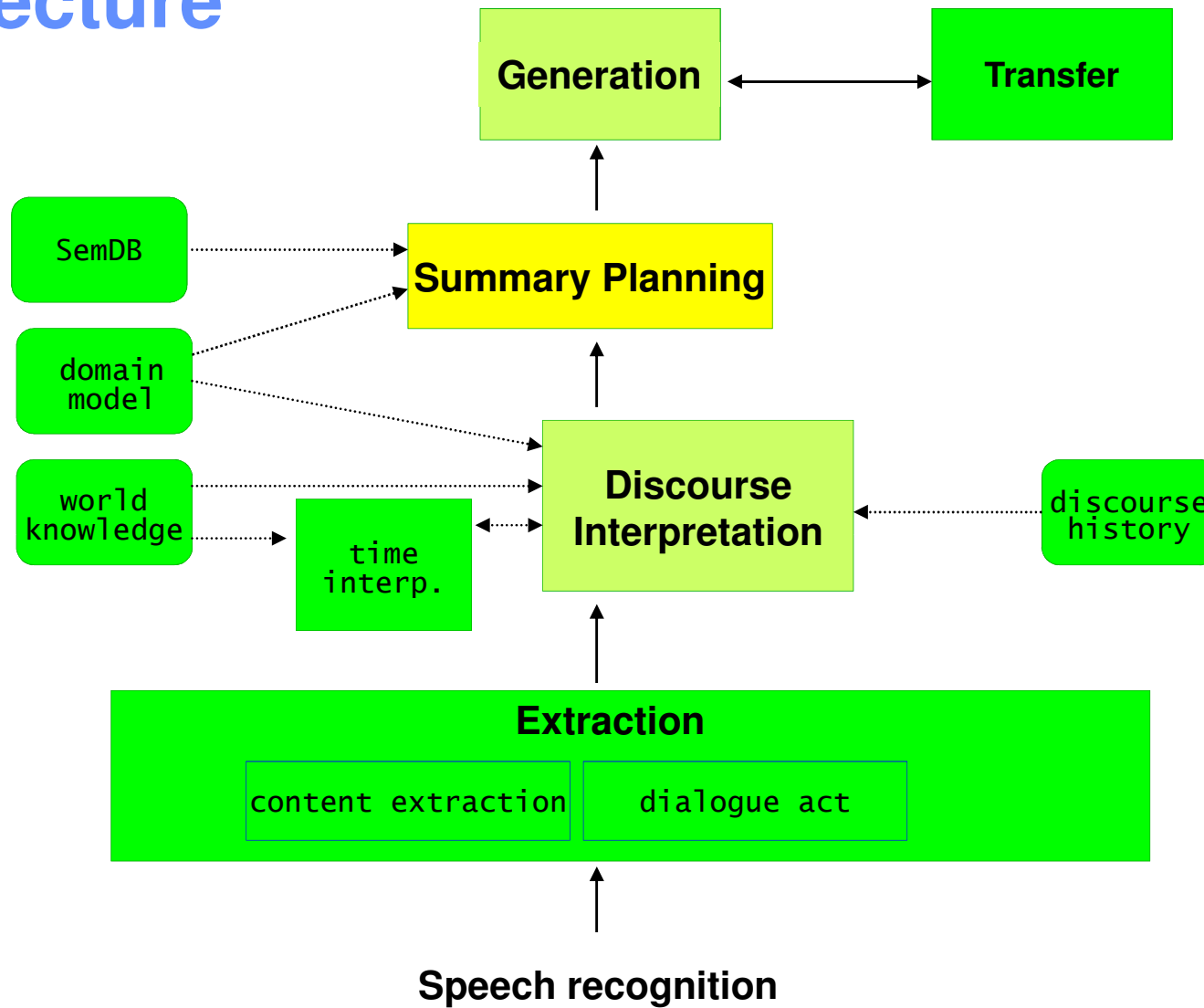


Collaboration for a New Functionality: Result Summaries

- **Provide the users with a summary of the topics that were agreed**
- **Two benefits**
 - have a piece of information to use in calendars etc.
 - control the translation
- **Approach: exploit already existing modules for**
 - content extraction
 - dialogue interpretation
 - planning the summary
 - generation
 - transfer



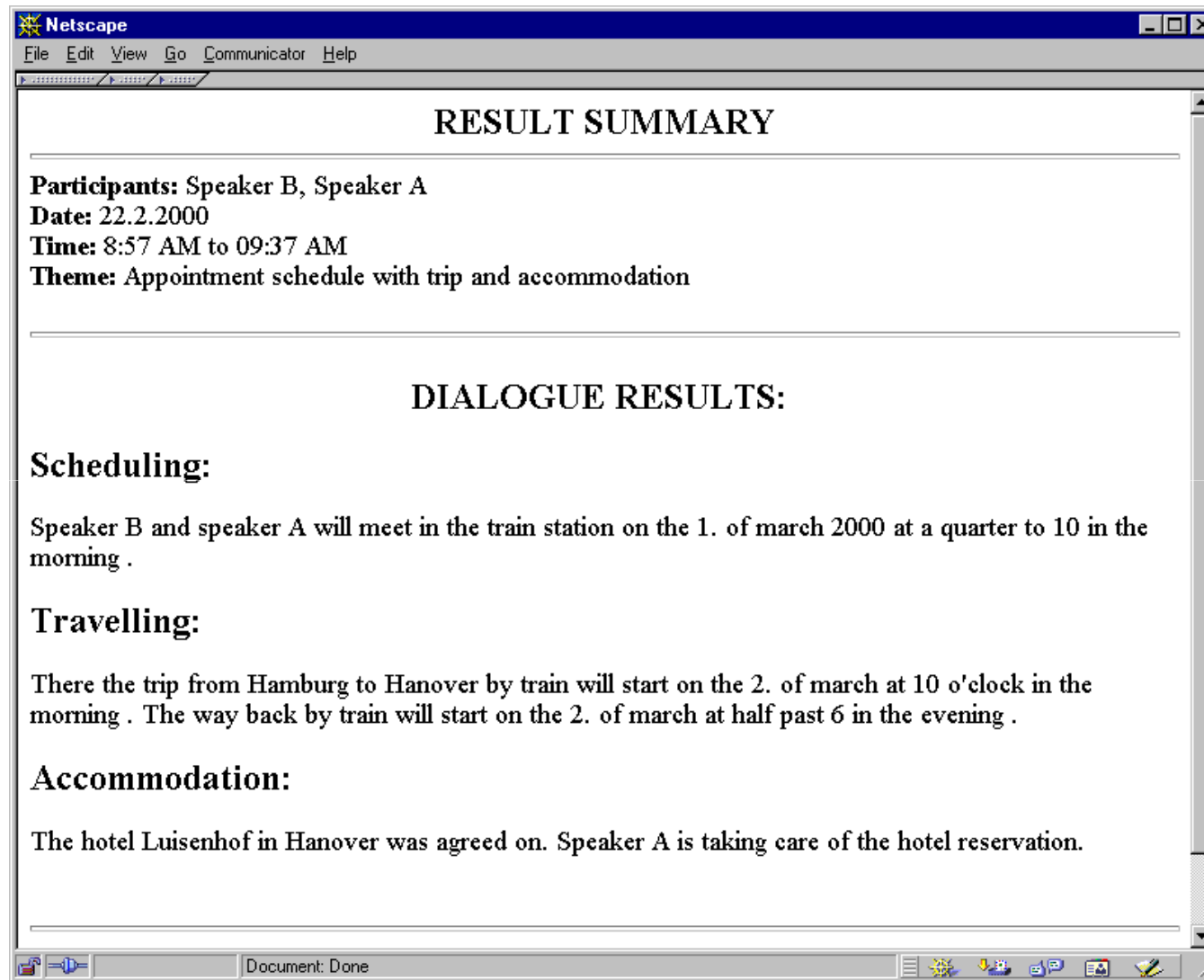
Architecture



- no changes
- few changes
- some changes
- new



Result Summary



Netscape
File Edit View Go Communicator Help

RESULT SUMMARY

Participants: Speaker B, Speaker A
Date: 22.2.2000
Time: 8:57 AM to 09:37 AM
Theme: Appointment schedule with trip and accommodation

DIALOGUE RESULTS:

Scheduling:

Speaker B and speaker A will meet in the train station on the 1. of march 2000 at a quarter to 10 in the morning .

Travelling:

There the trip from Hamburg to Hanover by train will start on the 2. of march at 10 o'clock in the morning . The way back by train will start on the 2. of march at half past 6 in the evening .

Accommodation:

The hotel Luisenhof in Hanover was agreed on. Speaker A is taking care of the hotel reservation.

Document: Done



Multiple Approaches

- **Mono-cultural approaches are dangerous**
 - humans vs. viruses ↓ diversity
 - Microsoft vs. ILOVEYOU and copycats ↓ alternative software solutions
- **Some sources of errors in a speech translation system**
 - external
 - spontaneous speech: not well formed, hesitations, repairs
 - bad acoustic conditions
 - human dialogue behavior
 - internal
 - knowledge gaps in modules
 - software errors
 - probabilistic processing

□ Use multiple engines, varying approaches on various stages of processing



Multiple Approaches in Verbmobil



- **Exclusive alternatives: three different German speech recognizers with various capabilities**
- **Competing approaches:**
 - three parsers: HPSG, Chunk, Statistical
 - five translation tracks: case-based, dialog-act based, statistical, substring-based, linguistic (deep) semantic translation
- **Needed: selection and combination of results from competing tracks**
 - parsers: combination of partial analyses in the semantic processing modules
 - translation: preselection module



Multi-Engine for Translation (D↑E)

- Large-Scale Web-based Evaluation: 25 345 Translations, 65 Evaluators
- Sentence Length 1 - 60 Words

<i>Translation Thread</i>	<i>Word Accuracy $\geq 50\%$ 5069 Turns</i>	<i>Word Accuracy $\geq 75\%$ 3267 Turns</i>	<i>Word Accuracy $\geq 80\%$ 2723 Turns</i>
Case-based Translation	37%	44%	46%
Statistical Translation	69%	79%	81%
Dialog-Act based Translation	40%	45%	46%
Semantic Transfer	40%	47%	49%
Substring-based Translation	65%	75%	79%
Automatic Selection	57% / 78% *	66% / 83% *	68% / 85% *
Manual Selection	88%	95%	97%

* After Training with Instance-based Learning Algorithm



The Selection Problem

Selection is a hairy business

- **confidence values are difficult to compare**
 - probabilistic vs. knowledge based approaches
 - no bird's eyes view possible
- **re-training necessary after changes in the engines**
- **training data must be produced**



Selection using Instance Based Learning

(Aha, D. W., Kibler, D., & Albert, M. K. (1991). Instance-based learning algorithms. Machine Learning, 6, 37-66.)

- **Software: MLC++ toolkit**
- **Test and training material 25 345 translations with ratings**
- **Classes**
 - STAT1 = statistical translation ok
 - STAT0 = statistical translation not ok
 - other tracks analogous \Rightarrow 10 classes
- **Used features: length, number of fragments, mean of confidence values, minimal confidence value, number of missing translations**
- **Cross validation on 10 disjoint test/training sets from the evaluation material**
- **Error: \approx 22% for word accuracy \geq 50%**



Experiences

- **Researchers, naturally, like their own approach most**
 - **Methods like statistical approaches spread in all fields**
 - **Cross-fertilization worked**
 - **New functionalities like dialog summaries through collaboration and reuse**
 - **The multi-engine approach is a Good Thing™**
 - **Collect data, and annotate it on various levels: indispensable for training, testing and evaluation**
- ❑ Build ONE integrated system, where everybody is involved**



Software Technology Challenges

The goal

- **Build an integrated system**

The situation

- **Researchers do research**
- **Using different programming languages**
- **Researchers don't want to be bothered with technical details**

The solution

- **Introducing: the System Group**
- **Maximal technical support for the researchers/developers**



Support from the System Group

Integration framework (testbed) with

- **common communication mechanism for all used programming languages (C, C++, Lisp, Prolog, Java, Fortran, Tcl/Tk)**
- **Narrow interface for all used programming languages**
- **Overall system control infrastructure**
- **Standards on various levels**
 - Installation
 - Compilation
 - Communication formats between modules
 - ...
- **Toolbox for recording, replaying, testing, inspecting data exchanged between modules, ...**



The Interface - Visualization and Debug Tool **Verbmobil**

The screenshot displays the Verbmobil software interface, which is used for visualization and debugging. The main window features a menu bar (File, Modules, Options, Debug, Actions, Repeat_Synthesis, Go, Stop) and a logo for 'bmb+f' and 'Verbmobil Verbundvorhaben'.

Overlaid on the main window are several sub-windows:

- VIM - Send Control Window:** Contains a 'Pool Selection Filter' with 'Content constraint' and 'Language constraint' options. It lists a 'Pool Selection' of recognized commands and hypotheses. Below this, it shows 'Turn Nr.', 'Reading', and 'Textual SID' fields.
- Dialogue Window:** Displays a dialogue act 'BVE' with a date of '23. May 2000, 2:36 pm' and a speaker 'Thompson, German'. It features a control panel with buttons for 'SCHEDULING', 'TRAVELLING', 'ACCOMMODATION', and 'ENTERTAINMENT'. A diagram shows a 'journey' node connected to a 'move' node, which is further connected to 'rail', 'Hannover', and 'date' nodes. A 'SCHEDULING' sub-window shows 'INIT' and 'SUGGEST' buttons.
- Control Panel:** Located on the right, it includes a 'Transfer' section with 'serv-101' and 'artup' fields, and a 'Plans' section with 'serv-102' and 'artup' fields.
- Dialogue History Window:** Shows a transcript of the dialogue:


```

[a.ge] schöner Guten Tag
[b.en] hello this is Thompson speaking
[b.en] hello hello Mr Schneider
[a.ge] ja es geht um das Geschäftstreffen in -hannover
[a.ge] das ist ja am zwanzigsten Januar um elf Uhr vormittags
[b.en] so we have to leave Munich at six o'clock
[a.ge] vielleicht fahren wir lieber den Tag davor
[a.ge] da gibt es einen Zug um zwei Uhr
            
```

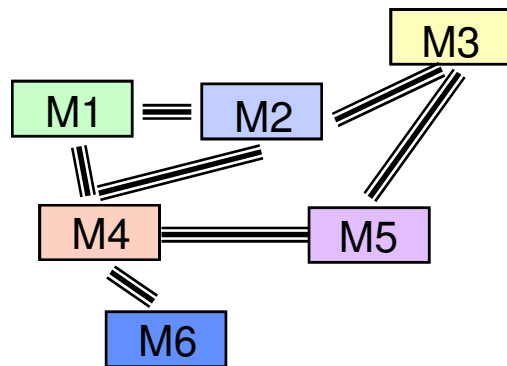
.... and much more



The Architecture

VerbMobil I

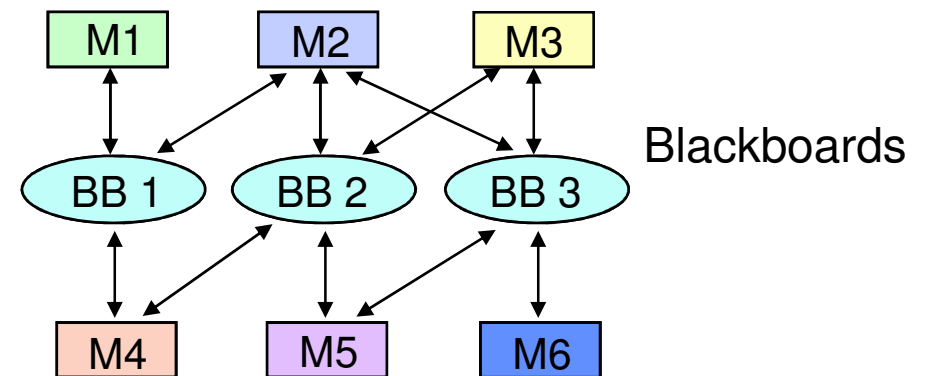
Multi-Agent Architecture



- Modules know all communication partners
- Direct communication between modules
- Reconfiguration difficult
- Software: ICE and ICE Master
- Basic Platform: PVM

VerbMobil II

Multi-Blackboard Architecture



- Modules know their I/O data pools
- No direct communication between modules
- **198 blackboards vs. 2380 direct comm. paths**
- Reconfiguration easy
- Several instances of one module/functionality
- Software: PCA and Module Manager
- Basic Platform: PVM

Experience

- **The System Group is a Good Thing™**
- **The multi blackboard architecture is a Good Thing™**
- **Crucial for the success of Verbmobil**
- **Software foundation for (almost) hassle free module development**

☐ Controlled distributed development possible



Management Challenges

The goal

- **Build an integrated system**

The situation

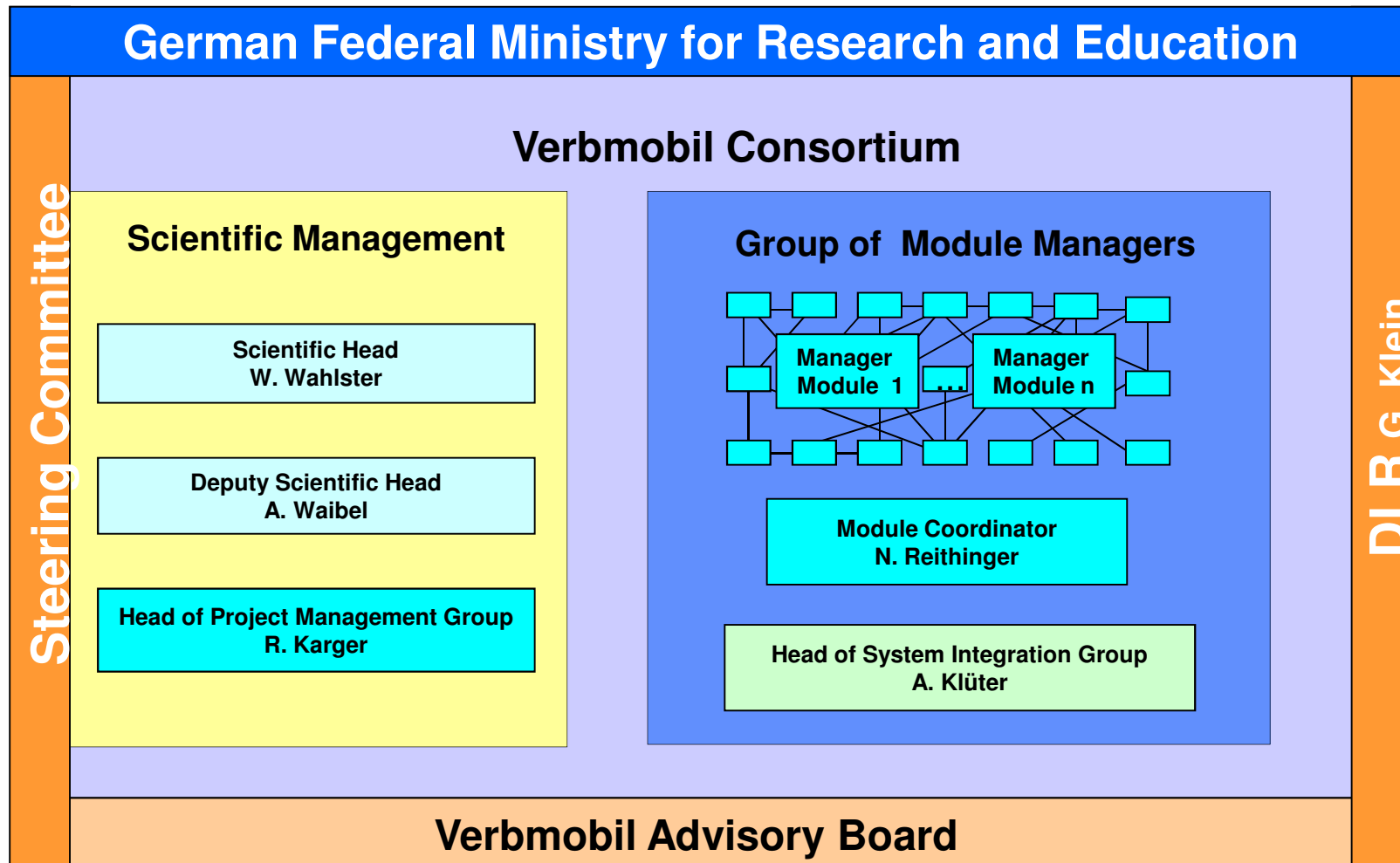
- **Partners distributed and pretty independent**
- **Great variation in project experience**
- **Adjustment of project plan and goals over time needed**

The solution

- **Define a flat management structure**
- **Create a group spirit**



Project Organization



Module Managers

- **Have technical hands on experience**
- **Responsible for one module, even if it is developed at different sites**
- **Volunteers (sort of ...)**
- **Meet regularly**
- **Define next milestones**
- **Define data and software integration plans**

Module coordinator coordinates the efforts and is the link to the scientific management



Example: Optimization Schedule 2000



- **21.02. Delivery of CeBit system**
- **21.02. - 30.04. Optimization phase**
 - **15.03. - 28.04. End-To-End evaluation with feedback to developers**
 - **27.03. - 07.04. Workshop Deep Processing**
- **09.05. Delivery Verbmobil System 1.0**
- **starting 09.05**
 - **speech recognizer evaluation**
 - **turn evaluation**



Experience

- **The group of module managers is a Good Thing™**
- **Common goals motivate**
- **Friendly peer pressure works most of the time**
- **Early problem detection and resolution in most cases**
- **Regular integration cycles focus and motivate**

□ Proactive consensus management (PCM)



Conclusion

- **Shared some experiences from a large distributed project**
- **Final Verbmobil Symposium July, 30., 2000, in Saarbrücken,
(<http://verbmobil.dfki.de>)**
- **Just before Coling 2000 (<http://www.coling.org>)**
- **Experiences will be used in SmartKom (<http://www.smartkom.org>) and other projects**



Thanks to ...



- **Wolfgang Wahlster (especially for a few slides :-)**
- **Jan Alexandersson, Michael Kipp, and Ralf Engel**
- **Reinhard Karger and Andreas Klüter**
- **all student assistants**

