

A Hybrid Cognitive Architecture for Multimodal and Multilingual Human-Machine Interaction

Nana Schlage, Toni Thelen, Lukas Cramer, Edwin Naroska, Gudrun Stockmanns

Faculty of Engineering and Computer Science, Niederrhein University of Applied
Sciences (Germany)

Contact email: nana.schlage@hs-niederrhein.de



Nana Schlage

- **Research assistant at Niederrhein University of Applied Sciences in the ambient intelligence laboratory**

Research & Work

- **Human-Robot Interaction & assistive systems**
- **AI-based systems (LLMs, RAG, recommendation systems)**
- **Embedded systems & sensor fusion**

Challenges of HRI in Public Spaces

- **Heterogeneous users (language, age, abilities)**
- **Noisy, unpredictable environments**
- **Existing systems:**
 - **monolingual**
 - **rigid dialogue flows**
 - **limited adaptability**

Contribution

- **A modular, service-oriented system architecture for multimodal and multilingual HRI**
- **All computation runs locally on one host**
- **Robot = actuator only**
- **Designed for:**
 - **multimodal interaction (speech + touch)**
 - **multilingual users (German, English, Arabic, Turkish)**
 - **controlled & predictable behavior**
- **Technical validation via benchmarks**

System Overview

1. Input

- Speech Detection (VAD)
- Speech Processing (Whisper)

2. Understanding

- Intent Classification (BERT + GPT fallback)

3. Decision Making

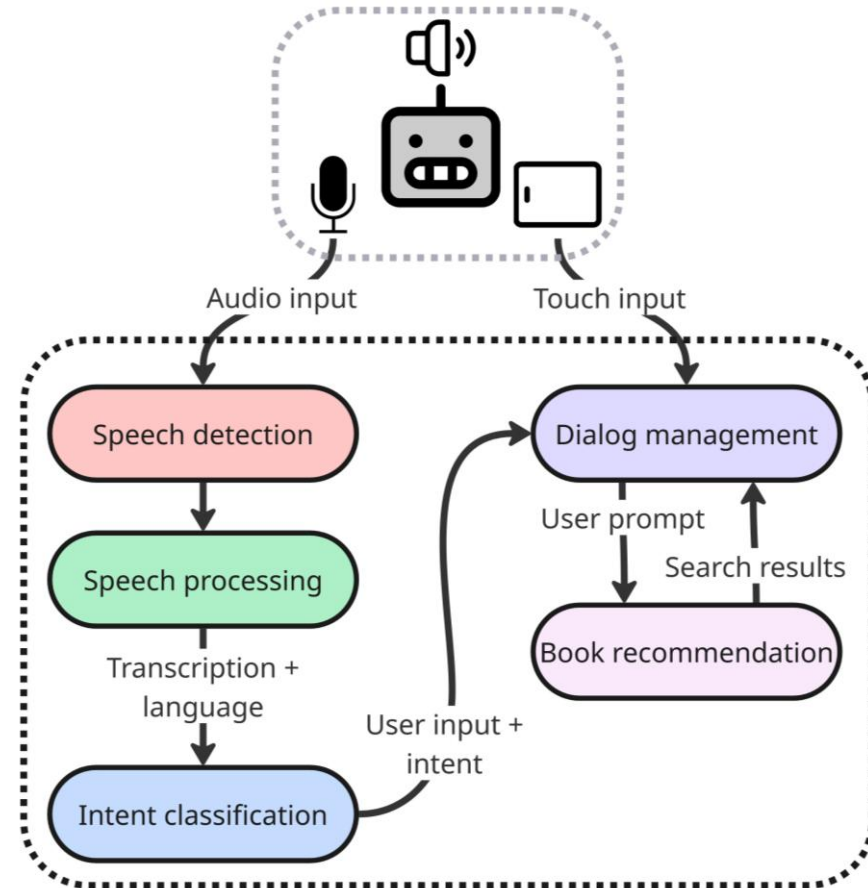
- Dialogue Manager (graph-based + LLM)

4. Features

- Book Recommendation (RAG + Agent)

5. Infrastructure

- Data Server (communication)
- GUI (monitoring/debugging)



Component-Level Evaluation - Translation

Model	Ø Latency [s]	Key Findings
nllb-200-3.3B	0.33	Unstable translations
Argos	0.66	Unstable translations
ChatGPT 4o	2.01	High quality, costly
Google	1.79	Good quality, free tier
MyMemory	1.65	High quality, free tier

Component-Level Evaluation – Intent Classification

Model	Ø Latency [s]	Error Rate [%]
Infloat E5	0.22	46.77
Infloat E5 (norm.)	0.06	41.67
Qwen3	0.15	31.99
Qwen3 (norm.)	0.09	46.77
BGE-M3	0.10	13.44
BERT	0.23	3.23

System Performance

	German	English	Turkish	Arabic	Total
Test Cases	247	254	253	251	Σ 1005
ØResponse time [s]	2.68	2.68	3.44	5.04	Ø3.46
ØTime book searches [s]	12.1	19.2	15.3	17.8	Ø15.7
Success rate [%]	98.8	97.2	96.4	96.4	Ø97.2
Intent error rate [%]	1.21	1.97	2.37	2.79	Ø2.09
Intent fallbacks	47	51	63	58	Σ 219

Conclusion & Future Work

- **Modular HRI system for multilingual, inclusive library access**
 - **Lowers barriers to public knowledge**
 - **Supports linguistic diversity (DE, EN, TR, AR)**

Future Work

- **Replace external services → fully local pipeline**
- **Improve intent classifier (fine-tuning)**
- **Extend to additional languages**
- **Enable adaptive, personalized interaction**

Hochschule Niederrhein
University of Applied Sciences



**Ingenieurwissenschaften
und Informatik**
Faculty of Engineering and
Computer Science