Künstliche Intelligenz manuscript No.
(will be inserted by the editor)

The Cognitive Service Robotics Apartment

A Versatile Environment for Human-Machine Interaction Research

Sebastian Wrede · Christian Leichsenring · Patrick Holthaus · Thomas Hermann · Sven Wachsmuth

Received: date / Accepted: date

Abstract The emergence of cognitive interaction technology offering intuitive and personalized support for humans in daily routines is essential for the success of future smart environments. Social robotics and ambient assisted living are well-established, active research fields but in the real world the number of smart environments that support humans efficiently on a daily basis is still rather low. We argue that research on ambient intelligence and human-robot interaction needs to be conducted in a a strongly interdisciplinary process to facilitate seamless integration of assistance technologies into the users' daily lives. With the Cognitive Service Robotics Apartment (CSRA), we are developing a novel kind of laboratory following this interdisciplinary approach. It combines a smart home with ambient intelligence functionalities with a cognitive social robot with advanced manipulation capabilities to explore the all day use of cognitive interaction technology for human assistance. This lab in conjunction with our development approach opens up new lines of inquiry and allows to address new research questions in human-machine, -agent and -robot interaction.

Keywords Smart Environments · Social Robotics · Human-Machine Interaction · Ubiquitous Computing

S. Wrede, C. Leichsenring, P. Holthaus, T. Hermann, S. Wachsmuth and the CSRA Team
CITEC, Bielefeld University, Germany
Tel.: +49-521-106-5148
Fax: +49-521-106-0
E-mail: swrede@techfak.uni-bielefeld.de

1 Introduction

With microelectronics increasingly pervading our living spaces and homes we currently witness the transformation of homes into smart homes that are responsive to the presence of people and offer assistance, increase security, enhance comfort or avoid unnecessary energy consumption. In this context we established the Cognitive Service Robotics Apartment (CSRA) as a smart environment which(i) offers a densely equipped sensor- and actuator network embedding virtual agents and a mobile robot to facilitate personalized and situation-aware interaction integrated in a homogeneous service-oriented software architecture, (ii) enables and actively exploits 24/7 operation so that the system can be used, or learn by itself from many interaction episodes, and (iii) provides a basis for quantitative and qualitative research under controlled conditions with automatic recording and post-processing of experimental data.

In comparison to other approaches (cf. [10] for a recent survey) an important aspect of our work is to consider the role and interplay of embodied interaction devices with smart environments and their users. For this reason, the CSRA features an anthropomorphic mobile service robot. In addition to its assumed role as preferred agent for mediating interaction with the smart environment, the semi-autonomous robot exceeds the modes of assistance beyond what could be realized by a smart (even actuated) apartment alone. Furthermore, the apartment and robot are *weakly coupled* systems

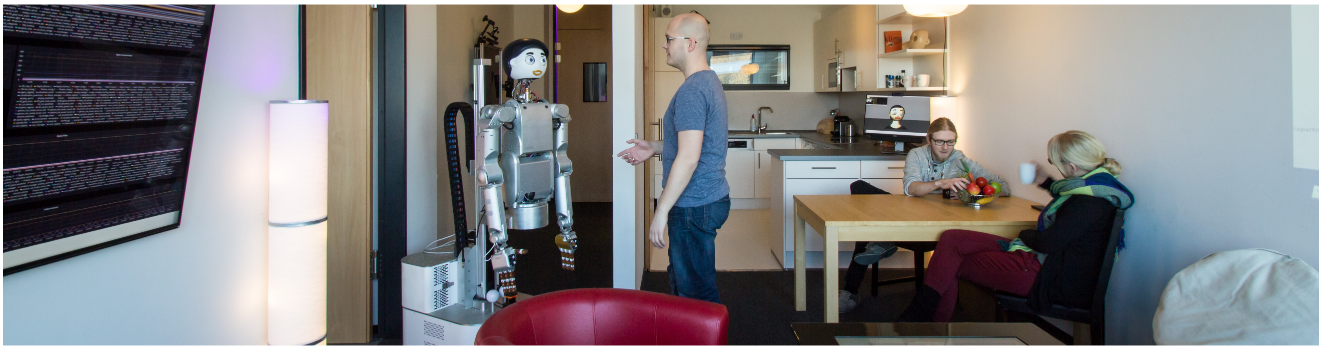

Fig. 1 Overview shot of the cognitive service robotics apartment and its inhabitant – the mobile robot *Floka*.

which can operate independently, yet can explore synergies by *cooperating with each other*.

Hence, the CSRA enables us to do research on the full scale from disembodied to embodied and even anthropomorphic human-machine interaction and how these might work together. Exemplary questions addressed inside the lab are: what kind of interfaces are best suited to enable specific functions of a smart home or mobile robot, how do users address these functions, what needs do actually arise at the user level, and how can approaches to satisfy them be taken up and used?

2 The Lab

The CSRA itself consists of three rooms with a total space of 60 m^2. One large room has kitchen, living room and hallway areas as shown in Fig. 2. Additionally, there are a bathroom and a multi-functional room, the gym, that can be changed according to changing research projects and further needs. Besides a developer access, there is a smart door connecting to a public corridor to allow the investigation of arrival and departure situations (see below). The area in front of this door and a control room where researchers operate experiments or develop new components are non-apartment parts of the lab where special privacy considerations apply. The size of the CSRA was chosen for reasons of ecological validity, i. e. service robots should deal with the limited space of a regular apartment for one or two adults.

The CSRA features a number of sensors. For visual tracking and recording, there are 12 RGB-D cameras in the ceiling used for person tracking, looking straight down to cover the whole apartment area approximately at waist level, 4 high-quality cameras for uncompressed video, and 4 lower-quality cameras that can stream H.264 video. The bathroom is exempt from any video surveillance for privacy reasons. As non-visual means for tracking and activity detection, there are 16 Fibaro motion sensors (using the Z-Wave protocol) and two areas outfitted with capacitive flooring. For audio recording and speech recognition, there are 12 cardioid microphones and 5 omnidirectional microphones installed.

Output-wise, the apartment itself can use a range of devices, including a Samsung Surface and a 4K TV seen in Fig. 1, a pan-tilt projector and two fixed projectors, 18 studio monitors plus 2 subwoofers and 50 Philips Hue LED light bulbs with configurable hue and intensity (using the ZigBee protocol), as well as smart drawer and cupboard handles that can show different light colors and patterns for targeted attention control. With the help of a KNX system, 66 plug sockets can be controlled and it can be sensed how much power is drawn from each socket. Furthermore, 56 wall buttons, 5 radiator and 2 temperature controllers, 2 shutters and 4 sun-blinds are integrated via KNX. The windows and doors of the apartment are equipped with wireless reed switches. There are also reed switches for the cupboard doors and drawers in the kitchen. The smart entrance door was equipped with a number of sensors as part of the KogniHome project (cf. Section 5), including an outward-facing camera, a radar sensor and pickup microphones. It also has a motor to open automatically, a remote-controlled lock, and an RGB frame, as well as an e-ink display on the outside.

There are two instances of a virtual agent displayed on two monitors, one in the hallway and one in the kitchen with webcams to track users with their eyes and recognize learned faces. The CRSA entails a mobile service robot as an embodied agent physically present in the apartment. The robot *Floka* is based on the Meka-Bot M1 and has an anthropomorphic upper body. It features multiple cameras, a laser range finder and mi-

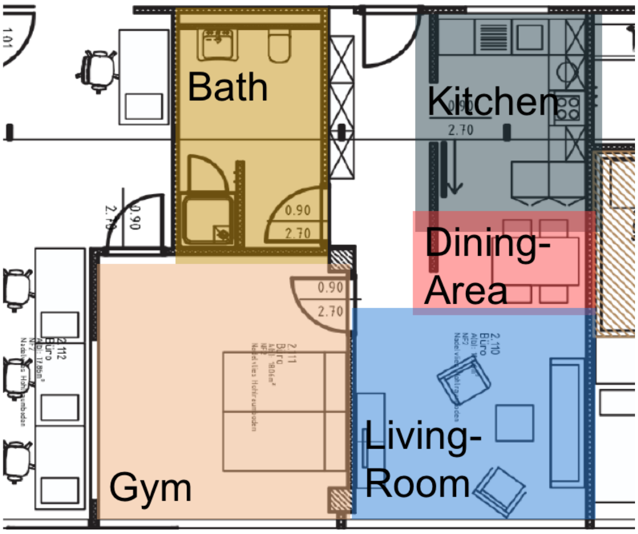

Fig. 2 CSRA room layout. Experiments are supervised from the operator room, which is located left to the living space.

crophones. The sensor head can be easily exchanged with a human-like robotic head that is able to show very expressive social cues. It was adapted from the robot head Flobi [12]. While it is realized as an independent autonomous system, it is also integrated with the apartment (cf. Section 3). The four-fingered hands feature compliant force controlled actuators. In total, the robot is equipped with 37 motor-powered joints. It has 7 per arm, 5 per hand, 2 in the head, 2 in the torso and 9 joints actuate the base including a z-lift.

Beyond KNX, we use two parallel gigabit Ethernet networks in the CSRA: one is used for IP-based connectivity of system devices and services while the other one is for transmission of introspection data and video streams that are recorded as ground truth in experimental studies. There is also a WLAN access point that allows wireless components to be integrated into the CSRA infrastructure, including the robot.

3 Software Architecture

The engineering of a smart environment system such as the CSRA with more than 300 sensor and actor devices and distributed computing nodes not only requires a carefully designed hardware setup. The continuous development and integration of a software architecture that provides the desired execution, i. e. responsiveness, robustness or usability, and evolution qualities such as extensibility or maintainability is even more challenging. In particular, if we want to achieve high availability and observability of the system state.

For these reasons and to cope with the intrinsic complexity of a large and evolving system, the CSRA software architecture provides a homogeneous interface to access the wide range of sensors, actuators and re-usable services required for human-machine interaction. The component-based architecture promotes separation of concerns and supports multiple pogramming languages (C++, Java, Python). It follows the “5C” [3] model for robotics software as a major design principle at all abstraction levels. While *computation* is a domain-specific concern realized in individual libraries, *configuration* already requires special consideration in a smart environment such as the CSRA. We utilize specific cross-cutting services, so-called registries, that manage the state, location and identity of every single sensor or actor unit in the environment and which i. e. supports spatial reasoning on data signals in the apartment. For *communication*, the event-driven communication patterns of a robotics middleware [14] are used that allow flexible coupling of devices and services with support for partitioning of the network and selective routing to prevent latency and bandwidth bottlenecks. The middleware allows to dynamically add or remove components to the system at runtime and supports synchronized recording of the raw interaction data for corpora and later analysis. The provided level of introspection and system observability is critical for conducting research in this environment. To facilitate efficient storage and retrieval of semantically annotated data of embodied interactions in our smart environment [9], a prototype of a hybrid data storage architecture is developed based on current graph and document-oriented database technology.

Another important concern is the *coordination* (and arbitration) of components to support multiple parallel interactions in the smart environment. CSRA services can dynamically allocate resources such as sensors or actuators or require exclusive access to other services such as face identification while a central component administrates the scheduling of these requests. Furthermore, modular hierarchical state machines modeled in a domain-specific extension of the State Chart XML language executed in an RSB-extended interpreter allow the model-based implementation of interactive behavior.

The services of the CSRA that we realized so far on the aforementioned basis can be *composed* in subarchitectures along two dimensions: abstraction level

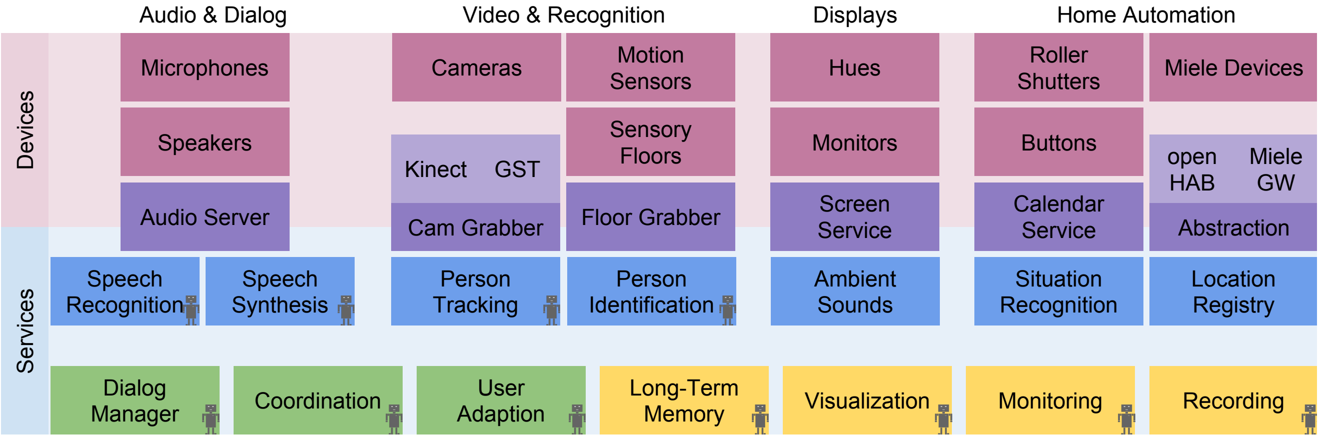

Fig. 3 Functional system overview. Vertical *stacks* describe functional areas of the architecture. Abstraction layers are depicted horizontally. Hardware is colored in lavender, sensor/actuator abstraction in purple, and functional components (partially available at the robot) in blue. Behavioral components (green) and cross-cutting concerns (yellow) are related to many stacks.

and functionality (cf. Fig. 3). Some components handle raw data, e. g. from specific hardware interfaces such as cameras, while others offer several abstractions or other virtual services. More abstract are functional components, such as the person tracking system. In the apartment context, all of these components are considered *base services* as they are always accessible and do not involve any behavioral routines. In contrast, interactive software, such as the dialog flow, can be re-configured or disabled depending on the current needs, for example a demonstration situation, or a user study. On the dimension of functionality, base services can be categorized into four different *stacks*: audio & dialog, video & recognition, ambient displays, and home automation.

The latter stack provides an interface to *home automation* hardware inside the apartment. Buttons can be configured to control the interaction, e. g. trigger scripted interaction segments. Motion sensors and a calendar integration allow for an estimation of the current situation, so that behaviors can be tailored to certain events or activities. In addition, a hierarchical location registry allows to address each device by its position inside the apartment, so that location-specific actions can be carried out. With the help of an *ambient display* stack, visual and auditive content can be provided to the apartment's users. All screens can be remotely switched and configured to display various content including websites and images. Furthermore, auditory devices and ambient lights can elicit personal music and illumination or other ambient configurations. A further stack provides *person recognition & personalization*. In order to provide positions of visitors inside the apartment, the tracking system incorporates data from the depth cameras mounted at the ceiling. User presence in each room is given with the help of motion detectors. Persons can also be learned and identified visually when interacting with the robot and the apartment's virtual avatars. A central component then provides access to names of people that have been recognized lately. For advanced *auditory processing & dialog* purposes, microphones and speakers are addressable through a central audio server that enables audio rendering and recording either globally or at specific locations. Speech recognition, synthesis and a dialog manager [4] are available at all agent instances. While the mobile robot *Floka* is equipped with special components that are needed for localization, manipulation, and movement, it also shares major parts of the apartment's architecture, as indicated in Fig. 3. For instance, certain functional components, such as speech processing, are also available on the robot and shared with the apartment.

On top of the functional stacks, behaviors of varying complexity can be realized. The current portfolio of *CSRA apps* consists of(i) *short routines* that often involve only one location or component, e. g. when asking for information on the current power consumption, (ii) *scripted interactions* which are especially suitable for demonstration purposes: e. g. on recognizing a new person entering the apartment, a special demo situation could be triggered and the virtual agent would offer to remember a person's face, and (iii) *automatic behaviors* such as recording data from all components or switching the light upon detecting user activity. Behavioral components and cross-cutting concerns also often consider the robot, so that e. g. routines involving parts of the apartment can be triggered at the robot and vice versa.

4 Operation

The operation of our smart environment entails several commitments and best practices for developers and users. As motivated in the introduction, the complete system is targeted to be continuously active and available. This facilitates extensive testing by developers and visitors who provide valuable feedback while being able to use the environment in their spare time. Parts of the system internals are thereby available for investigation on displays and as ambient cues. A sonification of system events provides information about component communication and colored lights display the status of the recording process.

Only for short periods of time that are scheduled regularly, the system is shut down partially. During these intervals maintenance work and cleaning occurs in a surveillance-free manner. Besides protecting the privacy of non-specialist employees, also rights of project staff and visitors are taken care of. With the help of agreement forms that everybody has to sign before entering the facilities, usage of personalized data can be restricted individually and deleted on-demand. Such requests are handled by the administration of the institution, where the anonymized forms are stored. In addition, efforts regarding data security are taken, in that the system is accessible only from inside a specially secured network.

In such a heterogeneous environment, development is needs to happen in a coordinated fashion. Therefore, periodically reoccurring code sprints are organized in order to collaboratively work on singular aspects and functionalities. Development, integration, and deployment are thereby heavily supported by the CITEC toolkit process [11] and a comprehensive testing framework that verifies functionality on a regular basis.

5 Research Projects

The CSRA provides a research environment for a long-term goal of supportive technology that deeply understands the daily routines and extraordinary events in a domestic environment and always offers the right kind of interaction with humans at the right point in time considering the full range of physical, virtual, social, or functional interactions. This research goal goes beyond a single project and beyond a single discipline. The CSRA as a lab infrastructure helps to bootstrap the research cycle of this process by providing a unique, rich, and technically integrated environment that allows testing of new components in complete scenarios, conducting new interaction studies (human-human, human-robot, human-machine) with parallel recording, or modeling and evaluating new concepts for, e. g. social interaction over longer periods of time.

Each project and study reported in the following also had an impact, vice versa, i. e. it enriched the research environment itself by adding behaviors, components, interfaces, or corpora:(i) The project 'Cognitive Service Robot as an Ambient Host'[1] explores new interaction strategies for addressing apartment functionalities [1, 5], memorizing them for later use, and adapting them with regard to social factors [13]. It explicitly considers the integration of the robotic platform and the smart environment and compares different degrees of embodiment. Different corpora have already been recorded by the project [7, 8, 13]. (ii) The 'KogniHome'[2] project takes specific use cases (cooking, entering and leaving the home, fitness exercises at home) and their interconnections for developing assistive functions for diverse user groups (children, adults, elderly) based on personalized models or the learning of automation rules [7]. (iii) The project 'Interactional Coordination and Incrementality in HRI' analyzes and models strategies to keep humans engaged during interaction utilizing robot platforms such as Floka or NAO [6]. The perceptual services of the CSRA are exploited to improve the performance of the engagement tracking approach. (iv) The CSRA yields an ideal environment for the team preparation in RoboCup@Home [2]. Here the robot must show an autonomous behavior based on its own sensors and actors. Thus, the smart environment can be used to automatically record ground truth data for test runs.

The established toolchains for CSRA infrastructure allow to easily import data into typical annotation tools like ELAN [8]. In order to get an idea, why an interaction with a human user was successful, why it failed, why there was a misunderstanding, what level of grounding was achieved, how the interaction was experienced, a

[1] https://www.cit-ec.de/en/content/csra
[2] https://www.kogni-home.de/

multitude of methods from different disciplines (conversational linguistics, psycho-linguistics, social psychology, informatics, etc.) is necessary for analyzing the data.

6 Synopsis

The presented Cognitive Service Robotics Apartment facilitates interdisciplinary research on human-machine interaction in the context of ambient intelligence, smart environments, as well as service robotics and allows to efficiently explore new approaches in longitudinal use in the 24/7 operation of the apartment. Within this environment we have integrated the MekaBot Floka – which is an extension of the Meka M1 robot with a newly adapted Flobi robot head. The virtual agent and Floka allow us to conduct research on the role of virtual and embodied agents and how these might work together with human users in smart environments. In this context, a number of research projects started in 2016 that utilize the apartment in different studies or as an experimental platform and itself contributed re-usable services to its architecture. These initial results suggest that the lab and its infrastructure is effective and usable for research teams from various disciplines.

During our work with and within our smart environment many more research questions, technological and organizational challenges show up, such as the self-organization of devices and services, the exploration of cooperative behavior of humans, robots and the environment as well as how to plan, distribute and schedule multiple interactive behaviors in multi-party situations. One of our immediate next steps will be to increase our own usage of the apartment for every-day situations such as meetings, demos, development, and social interaction in order to achieve a high-level of 24/7 usage of the smart environment. Realizing apps that are actually used by us regularly or daily gives us much more interaction episodes over months than we could obtain by conducting artificial studies and thus allow longitudinal studies observing our own changes in behavior.

Acknowledgements The authors would like to thank the development team (Birte Carlmeyer, Eduard Frese, Michael Goerlich, Michael Götting, Norman Köster, Florian Lier, Sebastian Meyer zu Borgsen, Jan Moringen, Marian Pohling, Viktor Richter, Simon Schulz, Johannes Wienke, René Zorn) for their hard work on the development of re-usable software, system integration and administration. This work was funded as part of the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277), Bielefeld University.

References

[1] J. Bernotat, B. Schiffhauer, F. Eyssel, et al. "Welcome to the future – How naïve users intuitively address an intelligent robotics apartment." In: *Int. Conf. on Social Robotics.* 2016.

[2] S. Meyer zu Borgsen, T. Korthals, and S. Wachsmuth. "ToBI – Team of Bielefeld The Human-Robot Interaction System for RoboCup@Home 2016". In: *Team Description Paper.* 2016.

[3] H. Bruyninckx, M. Klotzbücher, N. Hochgeschwender, et al. "The BRICS Component Model". In: *Annual Symp. on Applied Computing.* ACM, 2013, p. 1758.

[4] B. Carlmeyer, D. Schlangen, and B. Wrede. "Towards Closed Feedback Loops in HRI: Integrating InproTK and PaMini". In: *Workshop on Multimodal, Multi-Party, Real-World Human-Robot Interaction.* ACM, 2014, pp. 1–6.

[5] B. Carlmeyer, D. Schlangen, and B. Wrede. ""Look at Me!" — Self-Interruptions as Attention Booster?" In: *Int. Conf. on Human-Agent Interaction.* 2016.

[6] T. Dankert, M. Goerlich, S. Wrede, et al. "Engagement Detection During Deictic References In Human Robot Interaction." In: *Int. Conf. on Social Robotics.* 2016.

[7] K.-F. Engelmann, P. Holthaus, S. Wrede, et al. "An Interaction-Centric Dataset for Learning Automation Rules in Smart Homes". In: *Int. Conf. on Language Resources and Evaluation.* 2016.

[8] P. Holthaus, C. Leichsenring, J. Bernotat, et al. "How to Address Smart Homes with a Social Robot? A Multimodal Corpus of User Interactions with an Intelligent Environment". In: *Int. Conf. on Language Resources and Evaluation.* 2016.

[9] N. Köster, S. Wrede, and P. Cimiano. "An Ontology of Human Machine Interaction Data in Smart Environments". In: *Intelligent Systems Conf.* IEEE, 2016.

[10] R. Li, B. Lu, and K. D. McDonald-Maier. "Cognitive Assisted Living Ambient System: a Survey". In: *Digital Communications and Networks* 1.4 (2015), pp. 229 – 252.

[11] F. Lier, J. Wienke, A. Nordmann, et al. "The Cognitive Interaction Toolkit – Improving Reproducibility of Robotic Systems Experiments". In: *Simulation, Modeling, and Programming for Autonomous Robots.* 2014, pp. 400–411.

[12] I. Lütkebohle, F. Hegel, S. Schulz, et al. "The Bielefeld Anthropomorphic Robot Head "Flobi"". In: *Int. Conf. on Robotics and Automation.* IEEE, 2010, pp. 3384–3391.

[13] V. Richter, B. Carlmeyer, F. Lier, et al. "Are you talking to me? Improving the robustness of dialogue systems in a multi party HRI scenario by incorporating gaze direction and lip movement of attendees". In: *Int. Conf. on Human-Agent Interaction.* 2016.

[14] J. Wienke and S. Wrede. "A Middleware for Collaborative Research in Experimental Robotics". In: *Int. Symp. on System Integration.* IEEE, 2011, pp. 1183–1190.