



Article

# Behaviour of True Artificial Peers

Norman Weißkirchen <sup>1,\*</sup> and Ronald Böck <sup>1,2,\*</sup>

<sup>1</sup> Cognitive Systems Group, IKT, FEIT, Otto von Guericke University, 39106 Magdeburg, Germany

<sup>2</sup> Center for Behavioral Brain Sciences, 39106 Magdeburg, Germany

\* Correspondence: [norman.weisskirchen@ovgu.de](mailto:norman.weisskirchen@ovgu.de) (N.W.); [ronald.boeck@ovgu.de](mailto:ronald.boeck@ovgu.de) (R.B.)

† These authors contributed equally to this work.

**Abstract:** Typical current assistance systems often take the form of optimised user interfaces between the user interest and the capabilities of the system. In contrast, a peer-like system should be capable of independent decision-making capabilities, which in turn require an understanding and knowledge of the current situation for performing a sensible decision-making process. We present a method for a system capable of interacting with their user to optimise their information-gathering task, while at the same time ensuring the necessary satisfaction with the system, so that the user may not be discouraged from further interaction. Based on this collected information, the system may then create and employ a specifically adapted rule-set base which is much closer to an intelligent companion than a typical technical user interface. A further aspect is the perception of the system as a trustworthy and understandable partner, allowing an empathetic understanding between the user and the system, leading to a closer integrated smart environment.

**Keywords:** assistant system; companion system; human–machine interaction; exploration–exploitation; cognitive architecture



**Citation:** Weißkirchen, N.; Böck, R. Behaviour of True Artificial Peers. *Multimodal Technol. Interact.* **2022**, *6*, 64. <https://doi.org/10.3390/mti6080064>

Academic Editors: Christos Troussas, Cleo Sgouropoulou, Akrivi Krouska, Ioannis Voyiatzis and Athanasios Voulodimos

Received: 13 June 2022

Accepted: 26 July 2022

Published: 2 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The general idea of an assistant system, independent from its field of employment and intended function, is the support of their users in certain activities and objectives. As an additional requirement, such a system should also provide an easy and natural interaction to allow for an easy application of the assisting capabilities. Such a system, either as a device, tool or service is a staple of modern human–machine interaction (HMI) research, as seen in Biundo et al. [1]. The current state-of-the-art, especially in the area of virtual assistance systems which are the main exemplary focus of this work, are capable of providing (simple) automations and data-retrieval tasks. Regarding these developments, we see an evolution from simple text-based settings, being mainly related to desktop systems, toward systems which can be applied in real environments (as seen in Biundo et al., Hasegawa et al. and Deng et al. [1–3]). Such embodied devices allow for a more direct interaction since they have a physical representation (e.g., in Ötting et al. and Marge et al. [4,5]). Unfortunately, there is still a gap between this representation and the underlying concept of behaviour (as an overall control paradigm)—the main aspect of the current manuscript—and communication skills. According to Ötting et al. [4], two indicators for behaviour can be distinguished, namely *task performance* and *cooperation*. Current systems are mainly focussed on task performance, which indicates how well a technical system and its user handle a task or perform together as a team, as seen in Blackler et al. [6]. Therefore, such devices are often optimised for voice control or similar human-like interactions for an easy *understanding* and *integration* into the lifestyle of the user. Additionally, there is the idea of a real (semi-)autonomous system capable of providing continuous support and oversight of their user activities, as well as a more seamless integration into their user's lifestyle without the need for constant manual activation, as seen in Wendemuth et al. [7]. This idea is directly linked to the second indicator, which is mentioned in Ötting et al. [4] and also is the focus of

the manuscript: cooperation. For this, the approach requires much more autonomy on the part of the system, more than most current state-of-the-art systems/devices can provide, as it represents a step from the current reactive approach to a more proactive paradigm of capabilities and structures. Such a system would fill the niche of a *true* companion system, or a true peer as seen presented by Biundo et al. and Weißkirchen et al. [1,8], and would continuously and pre-emptively care for their user, specifically without the overt input necessary from the user. This not only subsumes the typical personal assistance systems, but also general assistance systems in industrial environments (e.g., smart manufacturing and service stations). Current systems instead focus more on integrated Smart Home Solutions as examined in Thakur et al. [9], or more efficient voice-controlled applications in a personal environment as presented in Valli et al. [10]. Additionally, even with this approach, with the current developing area of new and improved interface options, such as eye-tracking-based systems or even brain–computer interfaces, this can lead to extensive further improvements for interface technologies as examined in Vasseur et al. and Simanto et al. [11,12]. The requirements for a responsible and sufficient control instance increases with this further integration of human and technical system. As a result of the inclusion of new human–machine interface technologies, studies imply a greater impact of interactive media and information on the mental state of the user, which can be shown through higher attentiveness and brain activity as shown in Katona et al. [13–15]. Without allowing a technical system to analyse and rate its own impact, this can lead to potential harmful influences on the user. These systems by themselves, as a result, do not lead to a direct improvement of the capabilities of the assistance itself, which are then constrained by their reactive capabilities.

We have already proposed an approach capable of providing this kind of human-like assistant capability (see Section 3.1 for further details, as well as [8]), specifically a system equipped with a human-like decision-making process and their own set of priorities. Such a system works alongside a user, in the sense of collaboration and cooperation. This system provides “peer”-like capabilities (cf. [8] and Section 3.1) as an assistant, since it aims to continuously search for possible ways to assist (in the sense of partner at eye level) the designated user. Further, this kind of system tries to solve potential problems before they arise, or at least before they pose an imminent impairment. This is tackled through a combination of a comprehensive situational awareness, an adaptive and trainable representation of the most likely aims and priorities of the user, and—most importantly—the independent objectives of the system itself which actively control the way the system may solve potential impasses between the other aspects of the system. Based on these aspects, we establish the (cooperative) behaviour of a peer-like system or True Artificial Peer as shown in Weißkirchen et al. [8]) as a meta-level overseeing the main goals, achievements, and actions, providing an overarching strategy, being a consequent adaptation of the ideas presented in Schulz et al. [16] who argue for a (biologically inspired) ability of strategy changes as “a core competence of [technical, adaptive] systems”. Therefore, the behaviour triggers underlying concepts like sensor and activation control, dialogue management, etc. that handle the specific tasks in a very particular manner, ideally in an adaptive fashion. Regarding the “proactivity levels” presented in Meurisch et al. [17], which are separated into reactive, proactive or autonomous levels, we are dealing rather with the autonomous part of the range. In this sense, we aim for an extension of the system capabilities towards more system-individual capabilities, which is more than the current view on a rather task-oriented adaptability which was already discussed in Chaves et al. [18]. This meta-information can also be included at the level of report necessary for each decision, depending on the level of trust the user affords the system.

The advantage of this kind of method is not only related to a better integration into the daily lives of the user, but also its employment of a perceived empathy towards the interlocutor, often lacking in contemporary applications.

In contrast, in combination with the internal objectives and characteristics of the system, it allows for the system to experience empathetic reactions, and furthermore the

same from the user towards the system. This is achieved by effected and affected actions during an HMI, resulting in anticipatory decisions from the system's side as discussed in the works of Thun et al., Valli et al. and Vinciarelli et al. [10,19,20].

Before going into details regarding True Artificial Peers and our realisation of behaviour in the mentioned context, we provide an overview of our understanding of the relevant terms:

*Efficiency* is usually linked to the time necessary to solve or complete a task. In terms of communication, this refers to the number of turns present in an interaction to obtain the expected information as explained in Mitev et al. [21]. From our perspective, in HMI, the combination of both mentioned aspects leads to a holistic understanding of collaboration in settings where humans and technical systems interact with each other. Given interlocutors collaborating as partners and peers, a team can be compiled being more likely focussed on the task, which of course depends on the current setting and task as also discussed in Hancock et al. and Mitev et al. [21,22].

The *satisfaction* of the user is defined, according to Ötting et al. [4], "as the extent to which user responses (physical, cognitive, and emotional) meet user needs and expectations". For True Artificial Peers, this is directly linked to the aspect of how the system's internal goals and objectives are related to the interlocutors expectations.

From our perspective, user satisfaction is also connected to *acceptance*. Venkatesh et al. [23] define acceptance as "the attitudinal antecedent of usage", thus arguing that the mere usage is an intrinsic acceptance of the (system's) capabilities and limitations. Taking this into account, an accepted interaction is considered as any communication which is not broken down after a few turns.

Satisfaction as well as acceptance are coupled to indicator *trust*. In relation to Lee et al. and Ötting et al. [4,24], trust is the user's belief that the technical system will help to achieve common and shared goals "in a situation of uncertainty" as cited by Ötting et al. [4]. From this, we argue that trust can be achieved in either a highly capable and highly secure way—in the sense of being application-oriented—or in an interaction to establish common goals. In this manuscript, we foster the latter approach, allowing the system to also benefit from own and shared goals.

In our work we also mention the term "*empathy*", while there is a variety of definitions in human–human and human–mechanical interactions, for example, in Cuff et al. [25]. We use it specifically concerning the traceability of actions. The empathy factor in this case is the ability of a human user, or interaction partner, to assume what the system may do next. The same empathy also includes the ability of a technical system to assume a human decision-making process. This is of course a small aspect of true human empathy but is still an improvement on typical technical user profiling which often just approximates repeating actions and the assumption of a user that the system may latch onto these repeating actions, which reflects the current trend for improvement of assistance systems and smart environments as also examined in Murad et al. [26].

The general development is towards a system capable of understanding human emotions, intentions and decisions as described in Yoon et al. and Pelu et al. [27,28]. For "empathy" of user states, it is more important to recognise emotions or mental states, while "empathy" for an interacting peer is more focussed on retracing the human decision-making process. While the detection of emotional states with good results is already a staple of machine-learning-based classifiers as, for example, carried out in Schuller et al. [29], the interpretation of these emotions into a human-like "empathy" is more complicated as it requires technical alternatives for the understanding of the recognised emotion classes. Emotional classes range in that case from discrete classes, such as fear or happiness, to more indirect representation, such as a valence-arousal axis representation as presented in Russel et al. [30] and can give helpful indicators for the general state and satisfaction of the user.

This research primarily concerns itself with which different options a system may engage the user with. With this, a system can, for example, control interactions in a more

efficient and expedient manner. One of the main aims of the behaviour control is the choice of optimal interaction strategy for each situation. Especially during the initialisation and self-adjustment to their user, it is instrumental to generate an overview over the priorities and specificities of the user and the situation. In contrast, the system has to provide a certain amount of satisfaction, allowing a continued use of the system, specifically to impede a potential breakdown of interaction due to inconsequential actions. These two aims can be opposite in their implementation, as the data-generation process can be repetitive and error-prone, while a higher user satisfaction is often coupled with fast and correct decisions from a potential assistant system.

The current state-of-the-art, as well as the base for this work is given in Section 2. This includes explanations for our “peer”-like concept, and the conceptual similar BDI architecture. Given this introduction and motivation of our work, the main contributions are briefly summarised here, and will be discussed in detail in Section 3:

- C1: Providing an extension of the True Artificial Peer concept.
- C2: Providing a perspective to situation-adaptive characteristics of technical systems in interactions.
- C3: Providing a modelling approach for behaviour of technical systems, combining and extending concepts of BDI (cf. Section 2.1) and ACT-R (cf. Section 2.2).
- C4: Providing a framework for the realisation of autonomous behaviour of technical systems in general and True Artificial Peers in particular.

Finally, an outlook and conclusion are provided in Section 4.

## 2. Materials and Methods

In this section, we briefly reflect the foundations of the developed approach in relation to the current handling of “behaviour modelling”. Further, we also highlight the aspect of True Artificial Peers, the core concept and nucleus of the behaviour, which was introduced in Weißkirchen et al. [8] by the manuscript’s authors.

### 2.1. Belief, Desire, Intention Systems

To differentiate our system from an important contemporary method, we include an explanation of the BDI system, which uses similar terminology but on a different level and applicability than our system. An approach to solve the need for an independently intelligent agent, capable of generating and following their own set of actions based on external and internal states, is the so called belief–desire–intention (BDI) Model as presented in Bratman et al. [31]. This mirrors in certain aspects our idea of “peer”, which should also be independent in its decision-making processes. The intelligence in this case is the ability to compare and choose a preferable option from a pool of alternatives based on external circumstances. The name is compounded on the three main parts (see Figure 1) necessary in the decision-making process. Further, it shows relations to the True Artificial Peers architecture as discussed in Section 3.1.

The main objective of BDI systems is to provide three separated aspects, each consisting of interdependent databases and memories, which allow for a somewhat human-like approach for decision making under external information and internal assumptions. In detail, this consist of:

*Belief* part: This contains all factual knowledge concerning the technical agent itself, but also including external information as well. This includes static background information, like the declaration of certain facts, names, or unchanging dates. Furthermore, dynamic information, like the current time or temperature, is stored.

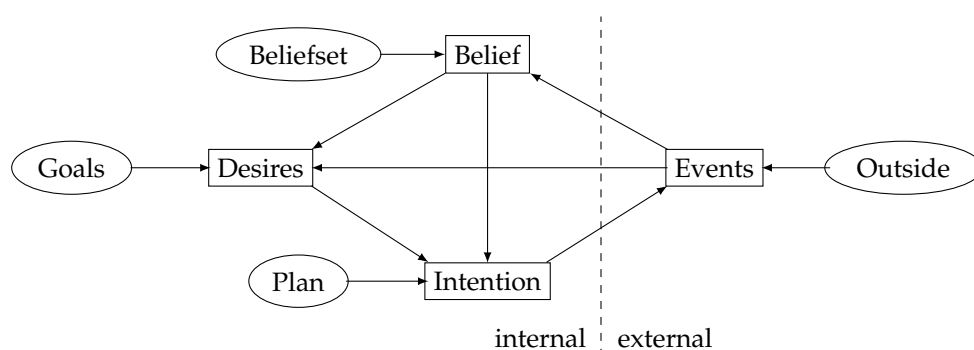
*Desire* part: Independently from the belief part, this contains the desire structure, where the system saves and retrieves the current system’s main objectives. This influences the model’s decision-making process, but is not the decision itself. The action of the system may be chosen to support the desire, or the taken action may conform to a desire, but the decision to take the action is not dependent on a congruent desire. This distinction is

necessary, so that the system may employ different actions over a longer period of time and retain the accompanying desire for several action steps.

*Intention* part: The structure directly controls the specifically chosen actions, as it contains the information on “what” has to be done. Here the decision itself takes form, which is based on the direct intention recalled in this part, modified by the current *belief* and *desire*.

This approach leads to a system which is capable of continuously recognising the surroundings (belief), choosing an overarching objective (desire) and selecting the appropriate plan or action (intention). A visual representation of these connections and their components in a typical BDI is given in Figure 1. Besides the core principle, the system is equipped with the necessary sensors and actors, connected to the recording and decision-making processes, to facilitate the flow of information into the system and the ability to influence the external world as examined in Rao et al. [32].

The BDI approach, while similar on a rudimentary level, differs greatly from our approach, on methodological and functional capabilities. The areas of belief, desire and intention are transferable in our system by different functional elements. These are, for example, the underlying cognitive architecture and the control unit, as explained in detail in Sections 3 and 3.1. The belief aspect specifically is primarily taken over by the external sensors and learning architectures, while the desire and intention aspects are controlled by the cognitive architecture controlling the decision-making process of the system.



**Figure 1.** Our simplified interpretation of the BDI architecture, inspired by works of Jacobson et al. [33]. The beliefs (knowledge) influence the desires and intentions. The goals lead to the next desire with the highest priority which in turn leads to the next intention, based on the available plans. The ability to react to external and system-initiated events presents a method of an intelligently reacting system.

## 2.2. Cognitive Architectures

Different cognitive architectures or systems are applied not only in academic research but also in several applications as shown in Kotseruba et al. and Sweller et al. [34,35], modelling various aspects of mental processing and cognition-based control. Regarding those modelling approaches, the cognitive processing can be reproduced, but the higher level of behaviour is yet rather fragmentarily mapped. Usually, behavioural reactions are often based on learnt and experienced circumstances as shown in Araiba et al. [36]. From a practical perspective, these issues are also linked to cognitive biases (learnt/adopted prototypical behaviour patterns or heuristic-based “shortcuts” in human cognition) allowing a “fast” reaction on current situations similar to the study in Doell et al. and Kotseruba et al. [34,37]. Interestingly, this aspect reflects a spotlight on the ability to handle human cognition and the derived behaviour, since each person is afflicted with such cognitive biases as explained in Tiedemann et al. [38]. From a technical perspective, such susceptibility for biases in human cognition is hard to comprehend, as in relation to technical systems, a well-defined logic-based behaviour is preferable. However, Tiedemann et al. and Kotseruba et al. [34,38] show that modelling architectures are also afflicted with biases, which might be seen as a drawback, but could be used to achieve a more



human-centric realisation of technical systems as in the sense of True Artificial Peers as presented in Weißkirchen et al. [8]. This might allow a creation of an “empathy” which enables (1) a better understanding and interpretation of the human counterpart and (2) a relationship to the artificial peers showing some “splens and quirks”.

As presented in Kotseruba et al. [34], a multitude of possible cognitive architectures is given. In our investigations, we focus on ACT-R as presented in Bothell et al. and Ritter et al. [39,40], utilising the fundamental descriptive power for the generation of behaviour in technical systems. Therefore, we briefly introduce this cognitive architecture: “ACT-R is [a cognitive architecture] widely used by a large worldwide community and applied to areas as diverse as airplane flying [...], intelligent tutoring [...], skill acquisition [...], and list memory” as explained in Dimov et al. [41]. The system allows a modelling and handling of human cognition on various levels. In particular, the use of particular memory buffers simulates the human short- and mid-term memory, as known from human information processing. These memories regulate the information and action exchange, indeed being considered also as a bottleneck. The processing of information and working rules is inspired by a probabilistic approach, mainly based on BAYESIAN theory as shown in Bothell et al. and Ritter et al. [39,40]. The processing in ACT-R is grounded on chunks (i.e., representations of knowledge), whose activation is related to the temporal and weighted influence of connections. The activations as described in Bothell et al. [39] are shown in Equations (1) (activation of a chunk) and (2) (basic activation per chunk):

$$A_i = B_i + \sum_j W_j S_{ij}, \quad (1)$$

where  $A_i$  is the activation of chunk  $i$ ,  $B_i$  is the basic activation of chunk  $i$ ,  $W_j$  is the weight of the goal's elements  $j$  (indicating the context) and  $S_{ij}$  is the connection's strength or “strengths of association from the elements  $j$  to chunk  $i$ ” ([39]);

$$B_i = \ln\left(\sum_{j=1}^n t_j^{-d}\right), \quad (2)$$

where  $t_j$  “is the time since the  $j^{\text{th}}$  practice” as shown in Bothell et al. [39] of the current chunk and  $d$  is the decay parameter. Besides the probabilistic handling of cognitive processes—to the best of our knowledge—no generation of behaviour or human-indicated “empathy” can be found in ACT-R, so far. Based on the fundamental ideas of BDI (cf. Section 2.1) and the processing of chunks (cf. Equation (1)), we discuss the way towards behaviour of True Artificial Peers in Section 3.

### 2.3. Assistant Systems

The main areas of applications for peer-like system are assistant system, or more generally, systems designed to work closely related to human interlocutors as examined in Biundo et al. and Wendemuth et al. [7,42]. As stated, our approach is designed to allow for longer sustainable interactions between technical and human agents, as well as a better understanding and satisfaction between the communication partners (cf. Section 3). To rank our investigations, we briefly discuss related work in the sense of technical systems being established as assistant systems. A conceptual framework of assistant systems or (sometimes called) companion systems is laid in the work of Wilks et al. [43,44], which was further extended to companion technologies as presented in Biundo et al. and Wendemuth et al. [7,42].

In contrast, the currently typical, and “traditionally” mentioned, assistant systems are mainly developed as commercial smart assistants, such as Amazon Alexa, Microsoft Cortana or Apple Siri which have a wide market appeal (as seen in various publications e.g., [17,45–48]). These systems employ usually voice-based interfaces, which in the beginning were primarily employed in a personal organiser system. This technology, primarily based on the continuous improvements in the speech-to-text methodologies shown in Zaidi et al. [49], allows for an easy and fast interaction and input from the user. However,

as discussed in Marge et al. [5], there is still a need for improvements and “dialogue-capable [...] systems”. For this, based on the (commercial) framework, being around, the idea of assistant system advanced towards more varied features and capabilities, often implemented in different “smart environments” settings as described in Cook et al. [50]). The typical smart homes as shown by Thakur et al. [9,51] allow for an integration of a technical assistant into the house itself and most equipped systems. This allows for a level of assistance closer to a butler or caregiver than a simple organising tool, resulting in applications in home-care situations as explained by Thakur et al. [9,52]; see their work for an overview). Such principles can also be transferred to smart factories as described in Lu et al. [53], where technical systems work in close proximity to the human employees.

In a broader sense—as, for instance, discussed in Biundo et al. [42]—technical systems that are intended to be assistant technologies need abilities to assess the user and his/her behaviour as well as to handle knowledge. We see a multitude of work considering specific aspects of the aforementioned broader issues, therefore only a spotlight can be given. Regarding the assessment of the user, a wide spectrum of modalities is investigated by Biundo et al. and Wendemuth et al. [7,42], mainly focussing on audio-visual signals. This can be elaborated, using, for instance, the special issue of Dibeklioglu et al. [54], highlighting detection and classification approaches as well as discussing insights to the current methods used. The survey of Yang et al. [55] complements the aspects, providing an interpretation of user behaviour. A more general overview of classification approaches and applied features is given in Böck et al. [56].

A final aspect should be considered as well, namely the modelling of knowledge in assistant systems. A first insight is presented by Gorodetsky et al. [57], who explain an ontology-based setting for data handling in personal technology. In the work of Le et al. [58], the modelling is extended to content-driven and relationship-based approaches, regarding information transfer and changes of behaviour. Recent works in this context use evidence-based models as done by Olson et al. [59] and also consider contextual and situational parameters as shown by Böck et al. [60]). Based on the review of the state-of-the-art and related work, we introduce our perspective on assistant systems and develop a theoretical concept as well as a possible framework to generate and adapt behaviour in True Artificial Peers.

### 3. Results

To establish a True Artificial Peer with its very own characteristics, objectives and goals, we propose the behaviour of these peers in the current section. For this, we define the term “behaviour” as follows:

Behaviour describes the actions, and more importantly the reactions, the system takes concerning their human interaction partner or partners in the case of multi-user applications. These actions include primarily the validation by the user of the steps taken till now, as to assure that the user is sufficiently informed about the system’s intention, as well as the internal shift of one set of objectives to another.

In the following, we emphasise and elaborate the theoretical concepts and methods, aiming for behaviour in True Artificial Peers.

#### 3.1. True Artificial Peers

The idea of a True Artificial Peer is based on our former work presented in Weißkirchen et al. [8]. There we extended the concept of a companion technology in Biundo et al. [1] which not only follows the passive role of a command interpreter, but is also equipped with their own set of objectives and priorities. These objectives are held concurrent to those from the human interaction partners.

Depending on the particular design, this can be used to provide a certain amount of stability for the interaction between humans and the system. For instance, orders or applications in conflict with a former specified set of safe priorities will either be ignored or actively communicated to the user in the case of a user mistake. This reduces the real

problem of wrongful activations and misunderstood command recognition. As seen in the article of the Verge [61], this is a really occurring current problem which leads, for example, to false marketplace activation in an assistant system with negative consequences for a user. This may only become more serious if the functionality of an assistant system is even more deeper integrated in the daily lives and such a system may control further personal assets.

Another important aspect is the decision under uncertainty or decisions without direct user input, as current systems rely primarily on the constant outside influx of orders and commands to continue working. Even under assumed infallibility of the interpretation of these inputs, the user still would need to constantly interact with the system to ensure a satisfactory result of the delegated task as the system lacks the capability to perform several independent steps for an overarching task. While this current model may reduce the workload of certain tasks to a degree, by providing an efficient and fast user interface, for example, to provide crucial information in a short time or to remind the user of important events and dates, it most often simply shifts an interaction from one control method to another. By providing an assistant system with their own objective set and the additional ability to generate further adapted sets based on an interaction, the system receives the ability to act during minimal information situations, either to engage the user for clarification or even to solve problems based on former experiences. This not only solves problems during communication but also allows the system to provide assistance when it is not specifically asked for. In this sense, the system would be reasonably assured of the intention of the user using a kind of “empathy” (cf. Section 3.2).

The last aspect was already described in Weißkirchen et al. [8]; it builds the foundation of the current research, and is the issue of traceability of the decision-making process. Every new task or objective for an assistant system can be interpreted as a handing over of responsibility from the human user to the technical system. While this can be a relatively small responsibility for current systems, it may inadvertently increase for more integrated applications, which control further aspects of the user lifestyle. This, of course, can and should lead to certain wariness, especially if the system is faulty in its examination of the user objectives. To reduce this uncertainty, it is imperative that the user is still the informed partner of the interaction, since the final responsibility of all actions remains with him/her as described in Beldad et al. and Mohammadi et al. [62,63]). However, the full informational breakdown of every decision the system makes, even for repeating actions, would simply shift the current dynamic of a user from constantly engaging the system (to give information) to constantly receiving information, which is still a high workload in this case to supervise the system continuously.

To reduce this constant alienation, we propose a specific set of “behaviour” on the system’s side which may change based on engagement of the user, level of information and relative level of habituation, helping to facilitate an improved level of trust from the user towards the system during any interaction.

### 3.2. Perspectives on Empathics

To explain the underlying principle of our behaviour approach, we introduce the concept of “empathy” and distinguish two main “emphatic” aspects in HMI: First the “empathy” of the system itself, concerning their human user. This includes aspects of a typical user profiling, but advances upon the idea of typical and untypical situations and reactions. Specifically, it assumes that a system should be able to recognise and react to situations which are without direct precedent or opposed to the typical profile. It not only allows that the system recognise the active influx of commands; it rather actively designs a user profile including typical actions and reactions either through recording given information or preferably by engaging the user in an interaction designed to extract or generate new information. Through this, the system may be able to map the decisions and preferences of the user on a continuous representation instead of a list of singular occurrence which may be recalled in case they reappear. This results in the continuous learning aspect not only recognising the action itself, but also the surrounding situation necessitating the



action. For example, a typical user interaction on a current assistant system may include the search for a specific restaurant based on distance, price and available cuisine. Based on that and also former search results, a current system may be able to generate a preference for a particular restaurant; for example, a short distance from home was the primary search criterion from the user. A changing interest from the user or a specific occurrence, which changes these priorities, cannot easily be included in this kind of representation of the user. While a system may rewrite or over-write a set of information, it would have problems differentiating a singular occurrence or a shift of priorities. This leads either to further specification by the user or non-optimal results by the system.

In our approach, the system would include questions and separate options to measure the dependence between distance, price and available cuisine to enrich the representation as well as include the information of specific re-occurrences which may influence the decision process of the user. For example, such changing (single) events could be an anniversary or bank holiday. Importantly, it would also recognise a change of typical behaviour, such as a missed lunch, a traffic jam or similar obstructions to the usual profile. It would then engage in solving the apparent problem of a potential hungry user as well as, if possible, the source of the obstruction. This would be based on architectures designed to approach human decision processes, as, for example, ACT-R or similar systems. The particular procedure is visualised in Figures 2 and 3.

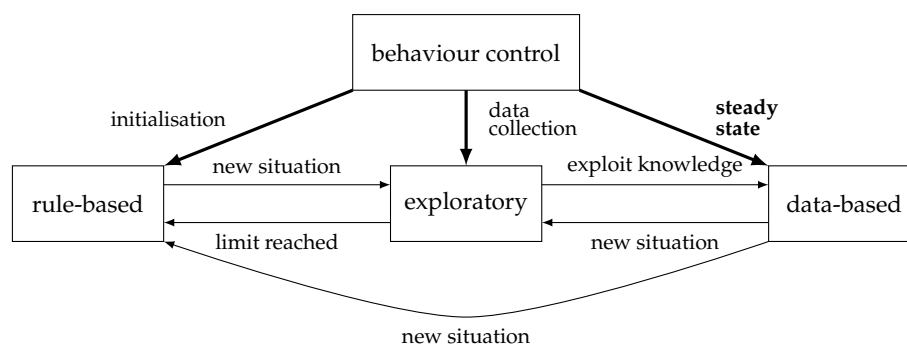
The second aspect is the “empathy” from the user concerning the system. This has to be separated from the natural adaptation most users undergo during the use of technical systems. Such adaptation can be easily recognised when using voice commands; users approach the system most often in a comparably unnatural and stiff way. To reduce faulty detections and wrong interpretations from the system’s side, they reduce contractions and clearly list their command grammatically, so that the technical interpreter can parse the information more easily.

In contrast, we envision “empathy” concerning the decisions and actions of the system. The decisions taken by the system are based on reproducible information and this information can be presented to the user at any time. Moreover, the users themselves may recognise arising problems or situations of lacking information before the situation occurs. In a practical application, this means that the user would “know” which commands have to be used and which could be abridged. A recurring order for a specific dish as takeout food could be phrased as “give me the usual” instead of a clear listing of commands. The empathy in this case would be the assurance of the user that the system would recognise “the usual” the same way the user would intend it to. More in depth, it also includes the assurance to the user that the system reports and explains the decision-making process. In a situation where the system is lacking information, it will additionally not stop working, but will actively engage and solve this problem. In the process, it will also display the current state the system is in, which further includes the user in the interaction. This allows not only for a more natural interaction, but also presents a better integration of the technical system into the user’s lifestyle. This we termed “peer level” in Weißkirchen et al. [8] in the True Artificial Peer as it approaches a human–human level of interaction.

To facilitate both these empathic dynamics, the system has to assure at the same time: (1) the ability to change the inner architecture based on new situations and (2) remaining mostly static in their decision-making process so as not to lose the understanding from its user. To allow these contrasting requirements, we propose “behaviour” as a controlling aspect of the internal dynamic adaptation and the external static appearance, which make the system understandable.

### 3.3. Types of Behaviour

We distinguish three primary types of behaviour the system can employ at any moment, as can also be seen in Figure 2. Each of these contain their own set of priorities during any (inter-)action between the system and the user, these will be characterised in the following.



**Figure 2.** A general representation of the concept. The overarching behaviour control unit oversees the incoming reactions from the user. In conjunction with the information level of the system, the unit chooses the most appropriate behaviour. The main situations occurring are the first initialisation, the need for data collection and the final learned reactions specific for the user standby state. As can be seen, new situations tend towards the exploratory state, except when the interaction between system and user breaks down.

The first “behaviour” is grounded on a *rule-based approach*, alternatively called a pre-designed approach. This comprises most state-of-the-art systems, which are (mainly) pre-programmed to react to each stimulus every time in the same manner. Specific users or specific situations usually do not influence the system’s action. Rather, users have to define preference by themselves to account for particular characteristics or actions, often during the user profile generation. In this case, the behaviour is objectively traceable but may pose problems when the current user is not aware of the underlying rules. Therefore, it leads to the aforementioned adaptation of the user to the system since the user adapts his/her interaction towards the desired reaction, which may lead to unnatural expression on the user’s side. To improve adaptation on the human side of the HMI, which may frustrate the user in turn, the system should be capable of communicating its internal rules implicitly or explicitly. This is especially important if unnatural or overly acted behaviour from the user side is otherwise necessary. While this approach may allow for better acceptance, it still does not allow for “true peer” level adaptation of a system. As an example for a typical rule-based interaction, the user would need to specify the desired action, the specific parameters, and acknowledge the process after call back from the system. A command could be something like: “System, order a pizza with salami toppings at 3.00 p.m. to my address at the nearest pizzeria.” Which translates to: “(Addressee: System), (action:order) (parameter object: a pizza with salami toppings) (parameter time: at 3.00 p.m.) (parameter location: to my address) (action location: at the nearest pizzeria).” This would be parsed towards a specific set of rules concerning the task under these parameters.

Second, we distinguish the rule-based behaviour from the “opposite behaviour” based on an *exploratory approach*. The exploratory behaviour allows for system actions during situations of high uncertainty and low information. The awareness of the situation and generated conclusions may rapidly change to react to the novel generated information during this state. Unless the system achieves enough information to generate a new rule, which requires an adequate understanding of the situation. While this method allows the latter generation of new rules, it may incur a level of heightened load on the user, as the active engagement rises significantly. To soften this impact, the system either informs the user from its change into the exploratory stage or otherwise minimises the taken steps to such an extent that the user may still be able to intervene. Without direct information of the behaviour change, some users may (subjectively) rationalise this behaviour, potentially assuming objective rules where there are none yet generated. Generally this behaviour may result in a decrease of trust and empathy from the user, the longer the system employs this approach, as it lacks the clearly defined rules of the other approaches. While the taken steps reduce the impact on the trustworthiness, this behaviour may not be followed indefinitely, before the user may stop the interaction. This reflects a form of exploitation/exploration

dilemma, as the system has to find a way to optimise the employments of exploratory and rule-based behaviour in a way to lengthen the possible interactions. The main method to alleviate some of the problems is the explicit way in its communication to reduce possible decisions which are opposed to the desire of the user. As an example, the system would, based on its knowledge of the user behaviour and former activities, ask: “Are you hungry? Would you like to order something to eat?” The intention of the system to engage with this question is not only based on a timer from the last known meal of the user, but also includes typical daily routines, known interjections, which might have led to a skipped meal, and measurable behaviour changes of the user which might imply hunger on their part. The system needs to anticipate and engage the user, ideally before a problem or situation becomes imminent, to convey its empathy towards the users’ decision-making process. This might lead either to a rejection or an acknowledgment by the user. In particular, the acknowledgment would result in further interactions, clarifying the type, time and further parameters of the ordering. This in turn would follow a known pattern, which the user recognises as desire of the system to learn more about the user and his/her situation. Importantly, the system would also engage in questions like: “Do you generally prefer this kind of food?” or “When is the usual time you like to eat?”. These questions are used to generate a deeper understanding of the user decisions for the system. Therefore, a continuous feedback evaluation is necessary to stop these questions as soon as the user appears to be dissatisfied as described in Nielsen et al. [64]. In case of a denial by the user, the system would attempt to re-evaluate its knowledge of the user and his/her situation. Practically changing its belief towards a more correct version, which better represents the needs of the user. Importantly, this approach not only employs typical user profiling, but also recognises changes from the norm, taking positive, negative and new situations into account. As a result of the continuing learning process, the time when the system engages this behaviour will give way to the last type of behaviour, as new rules are generated specifically for the user.

The third “behaviour” is grounded in a *data-based* approach. It is an intermediate level combining perspectives of the two previous behaviours since the system generates new rules (rule-based behaviour) based on the exploratory phase (exploratory behaviour). While it mirrors the rule-based approach, the rules here are directly based on typical user profiles, situational awareness and personal preferences of the user. This allows for a far more integrated and adapted system than a typical off-the-shelf (as given by a generally trained system) method would achieve. It also allows the system to generate own priorities and objectives for external situations, which are not directly tied to the user, but which are necessary for the continued functioning of the system. The important part is the implicit generation of rules without the direct supervision of either the original designer or the specific user. This fits our understanding of “peer” level, as the user may subjectively understand the rules perfectly, since they are based on the generated data during common interactions, while for a different human observer, not directly being involved in the exploratory interaction, the decision making process may still appear random or non-deterministic. We assume that a system, operating in this way, would elicit more trust and acceptance from the user, based on the completeness and correctness of the situational awareness generated during the common exploratory phase. This completeness can be measured as the amount of additional data which supports the achieved decision by the system, while the correctness would primarily depend on the feedback from the user. In fact, we did not perform user studies, yet, but can however, argue based on literature dealing with similar aspects. In Ötting et al. [4], the authors validate several hypotheses considering issues of HMI and how this is influenced by the autonomy and adaptivity of the technical system (cf. especially hypotheses 2 to 4). Given the respective findings, we see that adaptability has a positive effect on satisfaction and acceptance of the user. For autonomy, the study reveals no negative influences on acceptance. Unfortunately, for trust a positive effect could not be confirmed; however, this “need[s] to be interpreted with caution because [it is] based on a small amount of effect sizes”, see Ötting et al. [4]. In contrast, the meta-analysis of

Hancock et al. [22] shows that for trust additional characteristics (e.g., appearance) might influence the human expectations and assessments. In this sense, finally the entire setting—and not only the system’s behaviour—influences the overall trust in the technical device, which is though beyond the scope of the manuscript. Additionally, the user’s expectations influence the way the system is used, trusted and perceived. In their study—related to health care and well-being with AI systems’ support—Meurisch et al. [17] report that they “revealed significant differences between users’ expectations in various areas in which a user can be supported”, where expectations can be in a range from technically feasible to unrealistic based on the user’s technical understanding and knowledge. For mental health support, for instance, “most users tend to prefer reactive support or no support at all” see Meurisch et al. [17]; rejecting adaptive systems is usually correlated to expectations on privacy violation. Since the study comprises data from European and North American countries, the outcomes reflect (somehow) cultural differences. This should be kept in mind also for the approach; we suggested, as, for instance, participants from Canada preferred less proactive systems as shown in Meurisch et al. [17]. From Figure 5 in [17], we see rather negative feelings of the users the more the system tends to autonomy, which slightly contradicts the findings of Ötting et al. [4]. However, the negative view “can be partly explained by their attitudes and beliefs towards the particular AI system” of the study’s participants, where generally an openness toward a novel system can be seen. Given these studies performed by Hancock et al., Meurisch et al. and Ötting et al. [4,17,22], we conclude that our approach can contribute to the discussion on trust in autonomous systems and might lead to a better understanding as well as levelling of expectations and (system) outcomes. This is also related to the particular (difficult) selection of parameter settings, especially in the timing of behaviour changes. Therefore, the transition time between different behaviour stages should be chosen to minimise the additional load on the user, as each change requires additional mental and cognitive loads. The resulting interaction would be as described on the peer level, for example, by either the user requesting “the usual (stuff)” or by the system asking “if the usual (stuff)” shall be provided. On top of this direct interaction, there is also the possibility of the system providing the assistance without the given command. Both the system and the user are sufficiently sure what “the usual (stuff)” means, comprises and entails, providing the highest user satisfaction with the least necessary input. As the system generally learns the needs and wishes of the user, it will start to prepare “the usual” ahead of time, as well as recognise when it is not needed during a sudden change of situation. One used term in this research is “proactivity”, this term is used in different cases with different meaning behind it, for example, in Chaves et al. [18]. Generally, in conjunction with AI tools such as assistant system, it describes different levels of independent decision making. This potential independence is taken from the view of a user, and goes from fully reactive support, proactive decision making after checking with the user and at last autonomous decision without direct user input as described by Meurisch et al. [17]. The first step is fully part of the rule-based and break-down situation where every action is communicated by the user themselves. The check-up with the user is mainly part of the exploratory, while it is also part of the data-based stage in conjunction with the ability to follow its own decision-making process. Additionally, the decision of the system may even go a bit beyond this paradigm of proactivity and even autonomously, in cases when the decisions are not directly part of a user-support process. The architecture allows the system to engage and solve problems even in the case where no user is present or when there are transient users, such as in supervision function in factory settings or in other open environments.

The term of proactivity is also used in conjunction with the less general capabilities of a dialogue manager, specifically a chatbot as described by Chaves et al. [18]. Here proactivity also describes the effect of a system engaging a user without direct former input or signal. Importantly, while this also describes an interaction process, this aspect is subsumed into our general behaviour control. The interaction may as well be on a dialogue basis, but can also be based on actions and reactions, mimics or other indicators of user state, intention

and non-verbal interaction. Here again this engagement can also start without the user being present, autonomously improving the situation for the system itself. This interaction is not reduced to only dialogue from the system, but includes, among other things, sensor platforms, technical appliances or even semi-autonomous drones/agents.

The currently used behaviour is decided by the “behaviour control” or “control unit” (as seen in Figure 2). The specific methods to decide the change are given in Section 3.4, but it generally either mainly follows an algorithm based on information and user satisfaction, while an override by the system also remains possible.

As visualised in Figure 2, the particular behaviour is chosen and controlled by a *behaviour control unit*, which is explained in detail in Section 3.4. The control unit initially selects a rule-based paradigm, being a starting point for the exploratory paradigm when a new situation arises, given that the user is still cooperative. If the user shows a lower level of satisfaction, the system changes back to the less exploratory dependent but traditionally used rule-based approach. In contrast, given a suitable amount of analysable data, the system will switch to the data-based paradigm, which also contains rules, but is generated from specific user information.

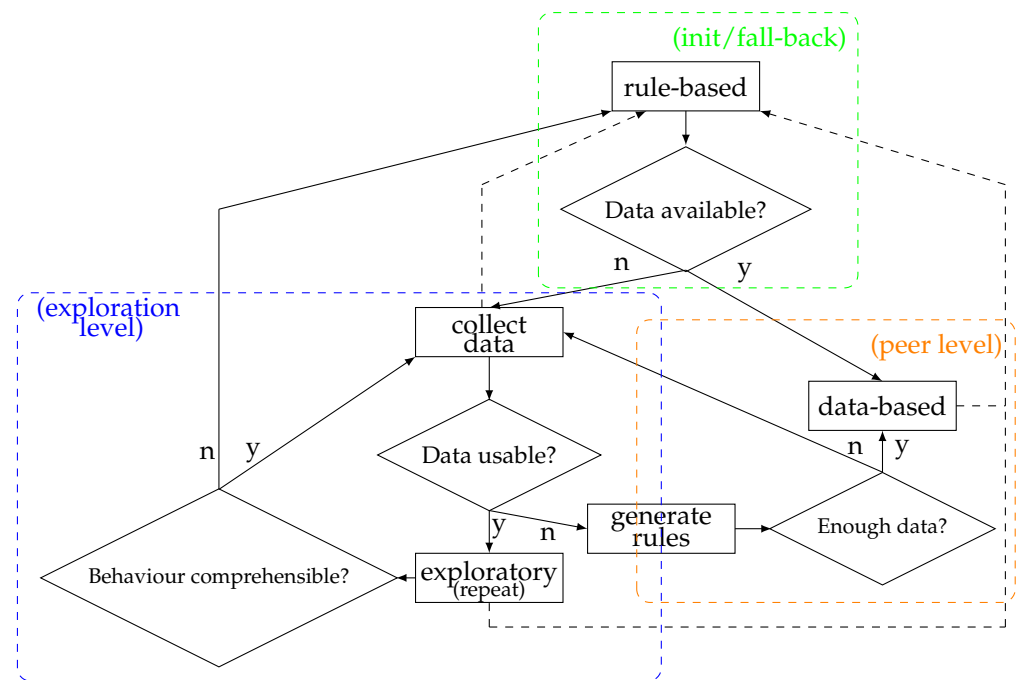
Given the considerations in Section 3.2 as well as in the current section, an advanced conceptual relation of True Artificial Peers can be achieved (cf. contribution C1). In particular, Figure 2 conceptualises our perspective on an adaptive and situation-related technical system. This allows an argumentation concerning how those systems in general and True Artificial Peers in particular could be set up (cf. contribution C2). Moreover, it lays foundations for the adaptation of behaviour which will be discussed in Section 3.4, realising the framework for the control unit.

### 3.4. Adaptation of Behaviour through the Control Unit

During the lifecycle of such a system, the behaviour may change dynamically as soon as the topic or situation changes during the interaction, based on the level of information and the reaction from the user. This is controlled by a delay between each change, so as not to impair the general satisfaction. This is observed and directed by the control unit, which declares the relevant behaviour for each time step at which the system is active. By default, the system may employ the rule-based approach as a final fall-back strategy and as the first approach during initialisation. During any interaction, the system recognises a new situation, either because of a command, which was not used before, or the user uses the command in an event which would imply a change of the general priority order. To resolve this situation the system applies the explorative behaviour to fill this lack of information, either by reaffirming the most likely solution, based on former interactions or by directly asking the user for clarification for new approaches. Depending on the level of complexity, the interaction may evolve to further topics. This can be seen in Figure 3, where the possible developments are shown.

Depending on the level of uncertainty, for example, the inability of the system to apply any former information to the particular new problem, the system decides the behaviour based on the satisfaction of the user. Given a general agreeableness from the user, the system may change freely towards the most possible exploratory behaviour. If no intervention by the user is detected, the system automatically applies exploratory behaviour to collect appropriate new data to compensate for the lack of information. In case of imminent negative reaction of the user or detected user dissatisfaction, the system immediately reverts to a standard rule-based approach. As the system continuously generates novel information, these rules are replaced by adapted “rules” generated based on user-specific data. Given this adaptation process, the proposed system is much more grounded in the current situation and exceeds approaches which have only a general rule-based method.





**Figure 3.** A visualisation of the internal control units and the respective state changes in terms of a flow chart. The main influences for the change are the available data or information level, and the satisfaction of the user with the current behaviour. Any lack of information leads toward the exploratory behaviour, which is ideally repeated as long as necessary for a full understanding. Given negative user reactions, the system reverts to a typical rule-based approach (dashed arrows represents the fall-back).

This kind of interaction and behaviour can be described in a mathematical way. The used formula is chosen to mirror the typical activation description as employed in the ACT-R architecture as explained by Bothell et al. [39] and Equation (1). This is not only used as a visualisation but also as a potential option to combine both systems. By using the already given cognitive architecture, this could be added relatively easily in a modular fashion to a typical workstep of ACT-R. Information which contains similar patterns or is retrieved together often in short timeframes automatically receives an increasing connection weight. This not only allows for the system to retrieve one data segment but further in-depth information, which may contain relevant background or topical data.

Introducing the term *knowledge value*  $K(T)$  for each specific topic  $T$  allows to combine information in the current situation, interaction or similar occurrences:

$$K(T) = \sum_{i=1}^n (I_i(T) + \sum_{j=1}^m (w_{i,j}(T) \cdot S_{i,j}(T))) \quad (3)$$

where

$I_i(T)$  are all beliefs or information directly concerning  $T$ , where  $n$  is the amount of available information,

$S_{i,j}(T)$  is all the contextual information concerning  $T$ , where  $m$  is the amount of contextual information for each  $i$ ,

$w_{i,j}(T)$  is the weighted importance which connects the context to the original topic  $T$ .

Each of these are parts in the expanding memory; during exploration, new data are generated and connections are created as possible. By including not only the directly connected information, but also the contextual information, the system aims for a deeper understanding. This context is part of the exploratory process. Specifically, this includes the extended knowledge, for example, different preferences dependent on time, location or preliminary actions.

This leads to high  $K$  values for topics  $T$  for which many direct and contextual information items as described by Böck et al. [60] are given. In contrast, low  $K$  values were reached in cases where the topic is unknown.

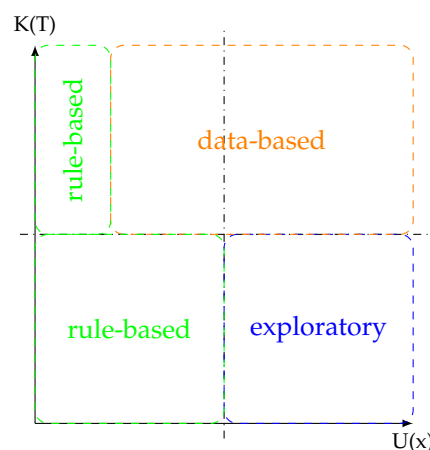
In addition, the system generates a user (-satisfaction) score  $U(x_i)$  which is based on the feature vector  $x_i$  for each user  $i$ .

$$U(x_i) = f(x_i^1, x_i^2, \dots, x_i^n) \quad (4)$$

The function  $f$  is dependent on the available external sensors which are used for user observation. These can be unimodal, like voice or visual, but also multimodal. The composition of this score depends on the specific system, but contains measures for the user characteristics such as shown in Böck et al. and Vinciarelli et al. [20,60]. In particular,  $x_i$  covers feature values for emotion, mental load and/or similar indicators of user satisfaction. Each of these values can either be taken directly from a connected sensor array, or indirectly by mapping extractable sensor information to the user states; this can be done with machine learning solutions as shown by Schuller et al. and Weißkirchen et al. [29,65]). The combination of these values for a general user satisfaction is then related to the personal perception of the user concerning different states, or alternatively the situation. For example, during a dialogue, user emotion is more important; during an assisted task, the mental load of the user is a more important indicator.

It also covers direct observations concerning the reactivity and interactivity with which the user replies to each inquiry. While the different aspects of the satisfaction score may change depending on typical user behaviour and type of inquiry, it assures that the user may not “give up” on the interaction. This includes the aspect whether the user agrees to continue with the current course of action. A high value indicates satisfaction and consent, while a low score implies dissatisfaction concerning the last actions of the system.

The knowledge value in combination with the user score spans a behaviour space, where each change of situation may lead to a transition of behaviour. As visualised in Figure 4, the previously described behaviours can be matched to areas in the space, covering the reasonable areas. As both satisfaction and knowledge ranges are normalised in the range of  $(0,1]$ , negative values are not possible.

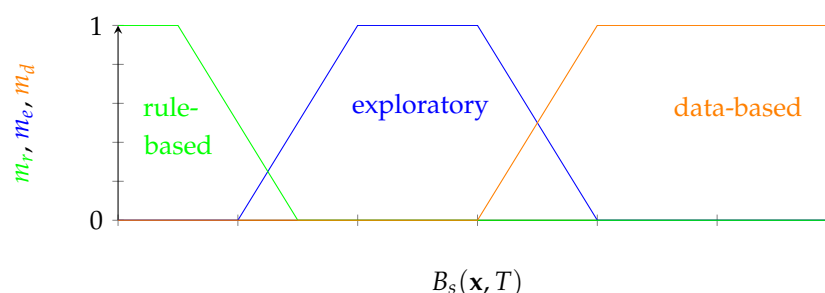


**Figure 4.** The figure highlights those areas in which the system operates with respect to the knowledge  $K(T)$  and satisfaction score  $U(x)$ . The lower parts present the general low information state, any change here depends on the satisfaction and reaction of the user. The upper part represents a high information state. As long as the reaction of the user is not categorically opposed to the interaction, the system prefers to remain with the adapted data-based behaviour; only in a breakdown does the system revert to the rule-based approach. In the case of a new situation and user interaction, the system transits to a satisfied exploratory stage to improve its information level.

Finally, we achieve behaviour values  $B_s(x_i, T)$  which combine the characteristics of both the user and the situation concerning the topic  $T$  for each interaction step  $s$ :

$$B_s(x_i, T) = U_s(x_i) + (U_s(x_i) \cdot K_s(T)) \quad (5)$$

$U_s$  and  $K_s$  are the same as explained but mapped to a specific time step. The estimation of  $B_s(x_i, T)$  can be expressed as, for instance, membership functions known from Fuzzy Logic, as shown in Figure 5. Each membership  $m_i$  selects the particular behaviour, executed by the system (cf. also Figures 3 and 4). Importantly, the membership function overlap—being usually applied in Fuzzy Logic to allow flexibility in final outputs—enables the system to not necessarily switch directly to its new behaviour, but allows a smooth transition. Otherwise, this could be perceived as “fleeting” or jumping in behavioural expressions. The gradual change, in contrast, is smoother and further is expected to be more natural or less abrupt for the user. The smoothness is primarily based on the continuous interaction flow, since the stepwise change is of course discrete. When changing his/her behaviour, the user will recognise the increase in interaction initiated by the system. This curiosity primes the user for a more in-depth explanation. The system also remains in this state for a recognisable amount of time, allowing the user to adapt to this change. A typical system would simply choose the most likely option, declare the inability to parse a command (in the hope that the user may re-phrase the request more clearly) or simply not act at all. This aspect is essential since—as also known from psychology—behaviour is a longer lasting characteristic of both humans and future technical systems. In contrast, a device is able to select and react on sensor inputs, which might be acquired in short intervals (some milliseconds to seconds) that might result in abrupt changes. In Böck et al. [66], this issue is discussed and respective approaches to handling this aspect are presented. These can be combined or also integrated in the suggested Fuzzy-like method, see Figure 5. This will result in a more naturalistic interaction and communication with the user.



**Figure 5.** The membership classes based on the value of  $B_s(x_i, T)$ . Instead of a space, the behaviour preference can also be presented as one value  $B_s(x_i, T)$ , which combines the aspects of information level and user satisfaction.

Regarding the main contributions stated in Section 1, we summarise: The current section provides a basic concept concerning the adaptation of behaviour of technical systems. This is based on the current knowledge (cf. Equation (3)) the system derived from its beliefs and the situation/context. For this, a scoring-like handle is achieved, combining main ideas of the BDI and ACT-R model, showing a theoretical concept (cf. contribution C3). In our research, the system’s behaviour was mapped to an adaptive behaviour value (cf. Equation (5)), which can be interpreted and operationalised from a Fuzzy Logic perspective, see Figure 5. Considering the theoretical argumentation in combination with Figure 3, a concept for a framework is given, realising the behaviour of technical systems (cf. contribution C4). The assistance systems mentioned in the state-of-the-art, compare Section 2.3, allow for a greater integration of technical systems into the daily lives of their human users, but at the same time (still) lack the ability to truly interact on a human-like, naturalistic level (in the sense of Valli et al. [10]). This is a weakness, especially when such a system is faced with its own set of objectives and priorities which it

has to accomplish concurrently or in spite of the task requested by the user. This potentially perceived distance between the assumed responsibilities of such assistance systems and their real capabilities to recognise and solve problems ought to be bridged. Therefore, systems need to engage their users more intensely without repulsing engaged (interaction) partners. Our solution allows (1) a system that engages as much as possible, trying to avoid the aforementioned critical state, and (2) the system to be better integrated in the decision-making process of the user.

#### 4. Discussion

In the current manuscript, we sketched a procedure and method to establish behaviour in True Artificial Peers as described by Weißkirchen et al. [8]. Those peers are technical systems or devices that extend current state-of-the-art systems, intended to be assistive devices in a general sense, see Section 2.3; this is described in Cowan et al., Marge et al. and Weißkirchen et al. [5,8,45]. From our perspective, True Artificial Peers do not have a passive role in an interaction; they rather take action by themselves, following their own objectives and goals, see Section 3.1; this is described by Weißkirchen et al. [8]. Given this additional quality of technical systems, these systems need their very own behaviour that results also in a better understanding of the system's characteristics by the user. In particular, a valuable interaction relates to a (grounded) understanding of the interlocutors' characteristics as described by Thun et al. and by Marge et al. [5,19], which is encouraged by interpretable and consistent behaviour (cf. contributions C1 and C2 in Section 1). Therefore, we aimed for such behaviour, allowing for (1) the active pursuit of the system's own objectives and goals and (2) an interpretable understanding of the system's reactions in an "empathic" way (cf. Section 3.2). To reach this goal, three types of behaviour were considered (see Section 3.3), namely rule-based, data-based and exploratory behaviour, which are controlled by a central unit called behaviour control (see Figure 2 and Section 3.4). The relation and transitions between the particular behaviour settings are visualised in Figure 3. Any transition is based on the behaviour value  $B_s(x_i, T)$  (cf. Equation (5)) that can be realised and interpreted, for instance, in the sense of membership functions, being well-known from Fuzzy Logic (see Figure 5). This enables the system, on the one hand, to derive suitable selections of behaviour and its characteristics, and also the human interlocutor, and, on the other hand, to interpret the system's reactions by implicitly predicting the memberships (see Figure 5). Further, this type of modelling results (usually) in a smooth transition between the respective behaviour types. The generation of behaviour in technical systems was, regarding Section 3, elaborated in a theoretical way (cf. contribution C3) as well as sketched in the sense of a framework and algorithm, visualised in Figure 3 (cf. contribution C4).

Since we presented mainly the theoretical considerations of our approach for the behaviour of True Artificial Peers and only discussed the beneficial implications in Section 3.3 building on work of, for instance, in Ötting et al. [4], we consequentially plan to integrate the approach for application-based research into (some) assistive devices, especially in (1) voice-based assistant systems and in (2) a setting related to ambient assisted living. We consider these two particular settings for the following reasons: A voice-based device allows direct communication between the interlocutors (cf. Section 3.3 and also by Marge et al. [5]); instructions, needs, desires, etc. can be negotiated; and finally, the "shared goal" can be achieved. In contrast, an ambient assisted living setting links the "interaction" partners in a different way. The "system", compiled from various sensors and multiple actuators, needs more "empathy" towards the inhabitant(s) see also Section 3.2, which further has to be combined with the objectives and goals to support the user(s), where rather an implicit exchange of information is in favour. Therefore, we have the ability to study the interplay of different behaviour types, as visualised in Figure 3 and discussed in Section 3.3. Furthermore, this allows investigations into the influence of the contextual information and history, as described in Böck et al. [60], being used in the transition between the three behaviour types.

Generally, the interchange with practical achievements, implementing our concept, allows more fine-tuning of the underlying theoretical foundations as well as a validation of the method, also in the sense of Ötting et al. [4], and Marge et al. [5]. Therefore, respective user studies in both a lab environment (mainly in short-term interactions) and “in the wild” (long-term interactions) are planned.

**Author Contributions:** Conceptualization, N.W. and R.B.; methodology, N.W. and R.B.; validation, N.W. and R.B.; formal analysis, N.W. and R.B.; investigation, N.W. and R.B.; resources, N.W. and R.B.; data curation, N.W. and R.B.; writing—original draft preparation, N.W. and R.B.; writing—review and editing, N.W. and R.B.; visualization, N.W. and R.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** We acknowledge support by the projects “Intention-based Anticipatory Interactive Systems” (IAIS) funded by the European Funds for Regional Development (EFRE) and by the Federal State of Sachsen-Anhalt, Germany, under the grant number ZS/2017/10/88785. Further, we highlight the support by the project “Adaptive Strategies in Assistive Technologies in Multi-Person Interaction” (ASAMI) funded by the Federal State of Sachsen-Anhalt, Germany, under the grant number: I 138. We acknowledge support for the Book Processing Charge by the Open Access Publication Fund of Magdeburg University.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

ACT-R	Adaptive Control of Thought-Rational
BDI	Belief, Desire, Intention (Architecture)
HMI	Human Machine Interaction

## References

1. Biundo, S.; Wendemuth, A. An Introduction to Companion-Technology. In *Companion Technology: A Paradigm Shift in Human-Technology Interaction*; Biundo, S., Wendemuth, A., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 1–15. [\[CrossRef\]](#)
2. Hasegawa, D.; Cassell, J.; Araki, K. The Role of Embodiment and Perspective in Direction-Giving Systems. In Proceedings of the AAAI Fall Symposium: Dialog with Robots, Arlington, VA, USA, 11–13 November, 2010.
3. Deng, E.; Mutlu, B.; Mataric, M.J. Embodiment in Socially Interactive Robots. *Found. Trends Robot.* **2019**, *7*, 251–356. [\[CrossRef\]](#)
4. Ötting, S.K.; Masjutin, L.; Steil, J.J.; Maier, G.W. Let’s Work Together: A Meta-Analysis on Robot Design Features That Enable Successful Human–Robot Interaction at Work. *Hum. Factors* **2020**, online first. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Marge, M.; Espy-Wilson, C.; Ward, N.G.; Alwan, A.; Artzi, Y.; Bansal, M.; Blankenship, G.; Chai, J.; Daumé, H.; Dey, D.; et al. Spoken language interaction with robots: Recommendations for future research. *Comput. Speech Lang.* **2022**, *71*, online first. [\[CrossRef\]](#)
6. Blackler, A.; Popovic, V.; Desai, S. Research methods for intuitive interaction. In *Intuitive Interaction: Research and Application*; Blackler, A., Ed.; CRC Press: Boca Raton, FL, USA, 2018; pp. 65–88.
7. Wendemuth, A.; Kopp, S. Towards cognitive systems for assisted cooperative processes of goal finding and strategy change. In Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy, 6–9 October 2019; pp. 4302–4307.
8. Weißkirchen, N.; Böck, R.; Wendemuth, A. Towards True Artificial Peers. In Proceedings of the 2020 IEEE International Conference on Human-Machine Systems (ICHMS), Rome, Italy, 7–9 September 2020; pp. 1–5.
9. Thakur, N.; Han, C. An Intelligent Ubiquitous Activity Aware Framework for Smart Home. In *Human Interaction, Emerging Technologies and Future Applications III*; Ahram, T., Taiar, R., Langlois, K., Choplin, A., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 296–302. [\[CrossRef\]](#)
10. Valli, A. The design of natural interaction. *Multimed. Tools Appl.* **2008**, *38*, 295–305. [\[CrossRef\]](#)



11. Vasseur, A.; Léger, P.M.; Senecal, S. *Eye-Tracking for IS Research: A Literature Review*; In Proceedings of the Pre-International Conference on Information Systems (ICIS) Workshop, Special Interest Group on Human Computer Interaction 2019, Munich, Germany, 15 December 2019.
12. Saha, S.; Mamun, K.A.; Ahmed, K.; Mostafa, R.; Naik, G.R.; Darvishi, S.; Khandoker, A.H.; Baumert, M. Progress in Brain Computer Interface: Challenges and Opportunities. *Front. Syst. Neurosci.* **2021**, *15*. [[CrossRef](#)] [[PubMed](#)]
13. Katona, J. Examination and comparison of the EEG based attention test with CPT and T.O.V.A. In Proceedings of the 2014 IEEE 15th International Symposium on Computational Intelligence and Informatics (CINTI), Budapest, Hungary, 19–21 November 2014; pp. 117–120. [[CrossRef](#)]
14. Katona, J.; Ujbanyi, T.; Sziladi, G.; Kovari, A. Examine the effect of different web-based media on human brain waves. In Proceedings of the 2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Debrecen, Hungary, 11–14 September 2017; pp. 407–412. [[CrossRef](#)]
15. Katona, J.; Kovari, A.; Heldal, I.; Costescu, C.; Rosan, A.; Demeter, R.; Thill, S.; Stefanut, T. Using Eye-Tracking to Examine Query Syntax and Method Syntax Comprehension in LINQ. In Proceedings of the 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Mariehamn, Finland, 23–25 September 2020; pp. 437–444. [[CrossRef](#)]
16. Schulz, A.L.; Woldeit, M.L.; Ohl, F.W. Neurobiological Fundamentals of Strategy Change: A Core Competence of Companion-Systems. In *Companion Technology: A Paradigm Shift in Human-Technology Interaction*; Biundo, S., Wendemuth, A., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 145–166. [[CrossRef](#)]
17. Meurisch, C.; Mihale-Wilson, C.A.; Hawlitschek, A.; Giger, F.; Müller, F.; Hinz, O.; Mühlhäuser, M. Exploring User Expectations of Proactive AI Systems. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2020**, *4*, 1–22. [[CrossRef](#)]
18. Chaves, A.P.; Gerosa, M.A. How Should My Chatbot Interact? A Survey on Social Characteristics in Human–Chatbot Interaction Design. *Int. J. Hum. Comput. Interact.* **2021**, *37*, 729–758. [[CrossRef](#)]
19. Schulz von Thun, F. *Miteinander reden 1-Störungen und Klärungen*; Rowohlt: Reinbek, Germany, 1981.
20. Vinciarelli, A.; Pantic, M.; Boulard, H. Social Signal Processing: Survey of an Emerging Domain. *Image Vis. Comput.* **2009**, *12*, 1743–1759. [[CrossRef](#)]
21. Mitev, N.; Renner, P.; Pfeiffer, T.; Staudte, M. Towards efficient human–machine collaboration: Effects of gaze-driven feedback and engagement on performance. *Cognit. Res. Princ. Implic.* **2018**, *3*, 51. [[CrossRef](#)]
22. Hancock, P.A.; Billings, D.R.; Schaefer, K.E.; Chen, J.Y.C.; de Visser, E.J.; Parasuraman, R. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. *Hum. Factors* **2011**, *53*, 517–527. [[CrossRef](#)]
23. Venkatesh, V.; Bala, H. Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decis. Sci.* **2008**, *39*, 273–315. [[CrossRef](#)]
24. Lee, J.D.; See, K.A. Trust in Automation: Designing for Appropriate Reliance. *Hum. Factors* **2004**, *46*, 50–80. [[CrossRef](#)]
25. Cuff, B.; Brown, S.; Taylor, L.; Howat, D. Empathy: A Review of the Concept. *Emot. Rev.* **2016**, *8*, 144–153. [[CrossRef](#)]
26. Khan, M.; Khan, M.T.R.; Saad, M.M.; Kim, D. A User Profile-based Smart Home Energy Management System. In Proceedings of the 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, 21–23 October 2020; pp. 646–651. [[CrossRef](#)]
27. Yoon, N.; Lee, H.K. AI Recommendation Service Acceptance: Assessing the Effects of Perceived Empathy and Need for Cognition. *J. Theor. Appl. Electron. Commer. Res.* **2021**, *16*, 1912–1928. [[CrossRef](#)]
28. Pelău, C.; Dabija, D.C.; Ene, I. What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. *Comput. Hum. Behav.* **2021**, *122*, 106855. [[CrossRef](#)]
29. Schuller, B.; Vlasenko, B.; Eyben, F.; Rigoll, G.; Wendemuth, A. Acoustic Emotion Recognition: A Benchmark Comparison of Performances. In Proceedings of the 2009 IEEE Workshop on Automatic Speech Recognition & Understanding, Moreno, Italy, 13 November–17 December 2009; pp. 552–557. [[CrossRef](#)]
30. Russell, J. A Circumplex Model of Affect. *J. Pers. Soc. Psychol.* **1980**, *39*, 1161–1178. [[CrossRef](#)]
31. Bratman, M. *Intention, Plans, and Practical Reason*; Harvard University Press: Cambridge, MA, USA, 1987.
32. Rao, A.; Georgeff, M. BDI Agents: From Theory to Practice. In Proceedings of the First International Conference on Multiagent System, San Francisco, CA, USA, 12–14 June 1995; pp. 312–319.
33. Jakobson, G.; Corp, A.; Parameswaran, N.; Buford, J.; Lewis, L.; Ray, P. Situation-Aware Multi-Agent System for Disaster Relief Operations Management. In Proceedings of the 3rd International ISCRAM Conference on Information Systems for Crisis Response and Management, Newark, NJ, USA, 14 May 2006; pp. 313–324.
34. Kotseruba, I.; Tsotsos, J.K. 40 years of cognitive architectures: Core cognitive abilities and practical applications. *Artif. Intell. Rev.* **2020**, *53*, 17–94. [[CrossRef](#)]
35. Sweller, J.; van Merriënboer, J.J.G.; Paas, F. Cognitive Architecture and Instructional Design: 20 Years Later. *Educ. Psychol. Rev.* **2019**, *31*, 261–292. [[CrossRef](#)]
36. Araiba, S. Current Diversification of Behaviorism. *Perspect. Behav. Sci.* **2019**, *43*, 157–175. [[CrossRef](#)]
37. Doell, C.; Siebert, S. Evaluation of Cognitive Architectures Inspired by Cognitive Biases. *Procedia Comput. Sci.* **2016**, *88*, 155–162. [[CrossRef](#)]
38. Tiedemann, M.R. Cognitive Biases in Cognitive Architectures as Indicators for Human-Like Behavior. Master’s Thesis, Otto von Guericke University Magdeburg, Magdeburg, Germany, 2018.

39. Bothell, D.; Byrne, M.D.; Douglass, S.; Lebiere, C.; Qin, Y. An Integrated Theory of the Mind. *Psychol. Rev.* **2004**, *111*, 1036–1060.
40. Ritter, F.E.; Tehranchi, F.; Oury, J.D. ACT-R: A cognitive architecture for modeling cognition. *WIREs Cognit. Sci.* **2019**, *10*, e1488. [CrossRef] [PubMed]
41. Dimov, C.; Khader, P.H.; Marewski, J.N.; Pachur, T. How to model the neurocognitive dynamics of decision making: A methodological primer with ACT-R. *Behav. Res. Methods* **2020**, *52*, 857–880. [CrossRef] [PubMed]
42. Biundo, S.; Wendemuth, A.; (Eds.) *Companion Technology—A Paradigm Shift in Human-Technology Interaction*; Springer: Cham, Switzerland, 2017.
43. Wilks, Y. Artificial companions. *Interdiscip. Sci. Rev.* **2005**, *30*, 145–152. [CrossRef]
44. Wilks, Y. *Artificial Companions as a New Kind of Interface to the Future Internet*; Technical Report; Oxford Internet Institute, University of Oxford: Oxford, UK, 2006.
45. Cowan, B.R.; Pantidi, N.; Coyle, D.; Morrissey, K.; Clarke, P.; Al-Shehri, S.; Earley, D.; Bandeira, N. “What Can i Help You with?”: Infrequent Users’ Experiences of Intelligent Personal Assistants. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*; ACM: New York, NY, USA, 2017. [CrossRef]
46. Kinsella, B. Amazon Echo Device Sales Break New Records, Alexa Tops Free App Downloads for iOS and Android, and Alexa Down in Europe on Christmas Morning. voicebot.ai. 2018. Available online: <https://voicebot.ai/2018/12/26/amazon-echo-device-sales-break-new-records-alexa-tops-free-app-downloads-for-ios-and-android-and-alexa-down-in-europe-on-christmas-morning/> (accessed on 26 December 2018).
47. Osborne, J. Why 100 Million Monthly Cortana Users on Windows 10 Is a Big Deal. TechRadar. 2016. Available online: <https://www.techradar.com/news/software/operating-systems/why-100-million-monthly-cortana-users-could-be-a-bigger-deal-than-350-million-windows-10-installs-1325146/> (accessed on 20 July 2016).
48. Kleinberg, S. 5 Ways Voice Assistance Is Shaping Consumer Behavior. Think with Google. 2018. Available online: [https://www.thinkwithgoogle.com/\\_qs/documents/5604/1178-CES-Voice-Research-PDF.pdf](https://www.thinkwithgoogle.com/_qs/documents/5604/1178-CES-Voice-Research-PDF.pdf) (accessed on 28 January 2018).
49. Zaidi, S.F.N.; Shukla, V.K.; Mishra, V.P.; Singh, B. Redefining Home Automation Through Voice Recognition System. In *Proceedings of the Emerging Technologies in Data Mining and Information Security*; Hassanien, A.E., Bhattacharyya, S., Chakrabati, S., Bhattacharya, A., Dutta, S., Eds.; Springer: Singapore, 2021; pp. 155–165.
50. Cook, D.; Das, S. *Smart Environments: Technology, Protocols and Applications*; Wiley: Hoboken, NJ, USA, 2005. [CrossRef]
51. Thakur, N.; Han, C.Y. A Context Driven Indoor Localization Framework for Assisted Living in Smart Homes. In *Proceedings of the HCI International 2020—Late Breaking Papers: Universal Access and Inclusive Design*; Stephanidis, C., Antona, M., Gao, Q., Zhou, J., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 387–400.
52. Thakur, N.; Han, C.Y., A Review of Assistive Technologies for Activities of Daily Living of Elderly. In *Elderly Care: Current Issues and Challenges*; Eckstein, S.M., Ed.; Nova Science Publishing: Hauppauge, NY, USA, 2020; pp. 61–84.
53. Lu, Y.; Xu, X.; Wang, L. Smart manufacturing process and system automation—A critical review of the standards and envisioned scenarios. *J. Manuf. Syst.* **2020**, *56*, 312–325. [CrossRef]
54. Dibeklioglu, H.; Surur, E.; Salah, A.A.; Dutoit, T. Behavior and usability analysis for multimodal user interfaces. *J. Multimodal User Interfaces* **2021**, *15*, 335–336. [CrossRef]
55. Yang, H.; Lee, H. Understanding user behavior of virtual personal assistant devices. *Inf. Syst. e-Bus. Manag.* **2019**, *17*, 65–87. [CrossRef]
56. Böck, R. Affects in Groups: A review on automated affect processing and estimation in groups. *IEEE Signal Process. Mag.* **2021**, *38*, 74–83. [CrossRef]
57. Gorodetsky, V.; Samoylov, V.; Serebryakov, S. Ontology-Based Context-Dependent Personalization Technology. In *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, Washington, DC, USA, 31 August–3 September 2010; Volume 3, pp. 278–283. [CrossRef]
58. Le, L.T.; Anthony, B.J.; Bronheim, S.M.; Holland, C.M.; Perry, D.F. A Technical Assistance Model for Guiding Service and Systems Change. *J. Behav. Health Serv. Res.* **2016**, *43*, 380–395. [CrossRef]
59. Olson, J.R.; Schurer-Coldiron, J.; Parigoris, R.M.; Zabel, M.D.; Matarese, M.; Bruns, E.J. Developing an Evidence-Based Technical Assistance Model: A Process Evaluation of the National Training and Technical Assistance Center for Child, Youth, and Family Mental Health. *J. Behav. Health Serv. Res.* **2020**, *47*, 312–330. [CrossRef]
60. Böck, R.; Wrede, B. Modelling Contexts for Interactions in Dynamic Open-World Scenarios. In *Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Bari, Italy, 6–9 October 2019; pp. 1475–1480.
61. Liptak, A. Amazon’s Alexa Started Ordering People Dollhouses after Hearing Its Name on TV. TheVerge. 2017. Available online: <https://www.theverge.com/2017/1/7/14200210/amazon-alexa-tech-news-anchor-order-dollhouse> (accessed on 20 March 2017).
62. Beldad, A.; De Jong, M.; Steehouder, M. How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust. *Comput. Hum. Behav.* **2010**, *26*, 857–869. [CrossRef]
63. Mohammadi, N.; Paulus, S.; Bishr, M.; Metzger, A.; Könnecke, H.; Hartenstein, S.; Weyer, T.; Pohl, K. *Trustworthiness Attributes and Metrics for Engineering Trusted Internet-Based Software Systems*; Springer: Cham, Switzerland, 2014; Volume 453, pp. 19–35.
64. Nielsen, F.; Minker, W. Assistive and Adaptive Dialog Management. In *Companion Technology: A Paradigm Shift in Human-Technology Interaction*; Biundo, S., Wendemuth, A., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 167–186.

- 
65. Weißkirchen, N.; Böck, R.; Wendemuth, A.; Nürnberger, A. Significance of Feature Differences in the Distinction of Mental-Load. In Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 7–10 October 2018; pp. 2608–2613. [[CrossRef](#)]
  66. Böck, R. *Anticipate the User: Multimodal Analyses in Human-Machine Interaction towards Group Interactions*; TUDpress: Dresden, Germany, 2020.