Learning Smart in Home

Acceptance of an intelligent personal assistant as virtual tutor

Alexander Kiy
alexander.kiy@uni-potsdam.de

Dustin Wegner

Ulrike Lucke
ulrike.lucke@uni-potsdam.de

Universität Potsdam
Institut für Informatik & Computational Science
Komplexe Multimediale Anwendungsarchitekturen

Abstract

Developments in the Smart Home sector are increasingly supporting users in their everyday lives and at home. The devices are intelligent, networked and integrated into the environment. With Intelligent Personal Assistants (IPAs), systems have recently made their way into the living room at home. However, the part of formal and informal learning with Intelligent Personal Assistants is not yet exhausted. Yet, IPAs offer an innovative approach to rethink for example classic learning methods with flashcards or partner learning. Are they feasible to support learning at home? IPAs can be used to provide knowledge and can be used for a variety of ways for assessment. This article presents an IPA using the wide-spread Amazon Echo as a realization of a virtual learning partner. Therefore, the prototype combines aspects from the domain of Intelligent Tutoring Systems and the domain of Learning Analytics with the help of a Learning Record Store (xAPI). The presented prototype was evaluated according to the Technology Acceptance Model. The articles closes with conceivable scenarios how IPAs will be able to adapt the physical environment according to the learning goal.

Keywords: e-learning; Intelligent Personal Assistant; Amazon Alexa; Learning Record Store; Technology Acceptance Model

1. Smart Home as a realization of pervasive learning environments

Intelligent and interconnected devices, so-called smart devices, have been appearing in almost all areas of life for several years. Under the catchword Internet of Things, these devices, the underlying infrastructures and processes offer more convenience in everyday areas. In the meantime, they are increasingly finding their way into the home living room, where they are referred to as Smart Home. According to Satpathy, the Smart Home is "a home which is smart enough to assist the inhabitants to live independently and comfortably with the help of technology [...]. In a smart home, all the mechanical and digital devices are interconnected to form a network, which can communicate with each other and with the user to create an interactive space" [1]. The devices are therefore increasingly

interconnected and hardly present for the users, so that an all-pervasive networking of everyday life occurs through the use of intelligent objects in the sense of a pervasive environment [2].

The so-called Intelligent Personal Assistants (IPAs) represent a Natural User Interface (NUI) to the Smart Home. IPAs have also been integrated into operating systems and devices, such as Apple Siri , Samsung Bixby or Google Assistant . IPAs offer a variety of functionalities that are constantly growing. These so-called skills can be used to execute a variety of functions to retrieve information or to control smart home devices. For this purpose, the voice input is transferred to a corresponding cloud infrastructure of the respective provider. The voice recording is semantically analyzed and then corresponding functionalities are executed in the cloud, on the local hardware or on connected smart devices. Through the collection of user data and a subsequent analysis using artificial intelligence, an attempt is made to simplify the daily life of the users by offering contextual support, functions and recommendations. In the meantime, specialized hardware is also offered, such as Amazon Echo with its associated IPA Alexa , or Google Home with the IPA Google Assistant.

However, little attention has so far been paid to the support of informal and formal teaching and learning processes in the context of smart homes. The primary focus is still on energy management, entertainment, communication and home automation [3]. IPAs offer a possibility to rethink traditional learning with flashcards or with a classical learning partner in a familiar environment, with technology and interaction taking a back seat. But are IPAs even suitable to support independent learning and assessment at home?

This paper presents an approach to use IPAs based on the Amazon Echo and the associated IPA Alexa to design formal and informal learning in the form of a virtual tutor. For this purpose, the technical and didactic requirements as well as the special features of agent-based and pervasive learning with the help of NUIs are discussed. With the help of the Technology Acceptance Model (TAM), the implemented proof of concept is evaluated with regard to user acceptance. In doing so, the fundamental question of whether users would be prepared in principle to use such a technology for home learning will be investigated.

2. Using IPA's as virtual tutors for teaching and learning

A didactic conception of the virtual tutor is crucially for the later acceptance by the users. For this purpose the didactic and technical components must be separated from each other and designed in a modular way so that teachers and learners can adapt the IPA to the teaching and learning scenario to their wishes as far as possible. As part of the requirements analysis, recent developments in the area of Intelligent Tutoring Systems (ITS) and the *design recommendations* applicable to ITS [4] were adapted for the design and implementation of the prototype. Furthermore current findings on teaching and learning research [5, 6] and virtual agents [4, 7] were used.

2.1. Teaching and learning settings for the use of an IPA

Learning with IPAs is a special realization of learning with virtual agents/tutors. IPAs allow the user to interact in a natural way without necessarily having to interact with input or display devices. Even though IPAs primarily address auditory learning types due to their voice-based operation, it is possible to enrich and extend the teaching and learning setting with visual and haptic interaction components by using suitable smart devices (e. g. smartphone, tablet or Smart TV).

As part of this work an IPA is being designed and scientifically evaluated especially to support independent learning. Characteristic for independent learning is that the learner "makes the essential decisions about whether, what, when, how and why he learns and can have a grave and consequential influence on the latter" [6]. In this context, there are various prerequisites for the learner, the learning media, the learning material, the learning tools and the learning environment [6], which are taken into account when designing independent teaching and learning processes.

When studying independently the learner can decide, amongst other things, about goals, content, general conditions, the learning process and the assessment. Therefore, the IPA has to keep the didactic parameters freely configurable by the user as far as possible. Furthermore, the IPA has to provide feedback, assistance and guidance, enable the adaptivity of the learning environment, ensure the usability of the learning tools used, and an option to structure learning materials [6]. Overall, the selection of a suitable interoperable format for the learning material should always be considered (ibid.). Learning tasks can be worked on by the learner in a changing sequence. These variable learning paths must also be considered during conception and implementation. Last but not least, the virtual tutor has to enable diagnostic, formative and summative feedback and self-assessment to support independent learning.

Essentially, a teaching and learning setting with a virtual tutor can be limited to the learner, the teacher (or virtual tutor), and a set of learning objectives (tasks). For the technical implementation already recognized standards and data formats exist for the individual aspects, which can be used to differentiate the didactical setting. The didactic representation of the learning tasks was based on the wellknown learning target taxonomy of Anderson & Krathwohl [8]. For the conception of the IPA, first the extent to which an IPA could be used to audit and verify more complex taxonomically levels was clarified (beyond knowledge and understanding and a differentiation with regard to factual knowledge, conceptual knowledge and procedural knowledge). The Computer Supported Evaluation of Learning Goals (CELG) Taxonomy Board by Mayer [5] is ideal for transferring or mapping learning tasks to a suitable form with an associated task type. The following table provides an overview of an exemplary assignment of task types that can be mapped using an IPA according to the factual, conceptual and procedural knowledge categories (see Table 1). However, neither true or false questions, nor single choice questions are suitable for the reproduction of the procedural knowledge category (P), just as multiple choice questions for the reflection and evaluation level. Nevertheless, other task types can be used to cover the relevant procedural knowledge category, as it can be seen in Table 1. Two example questions are listed below to illustrate a sequence and a mapping with the associated answer options:

Question I: "How is most of the surface in Germany covered? Sort descending!"
Options I: "Forest", "Water", "Agriculture", "Settlements"

Question II: "Match two of these trees to their species family!"
Options II: "coniferous tree", "deciduous tree", "copper beech", "scots pine"

Task types	Repro-duce	Understand & Apply	Reflect & Eval-uate	Cre-ate
True/False	X(P)			
Single Choice	X(P)			
Multiple Choice	X	X	X(P)	
Markings	X	X	X	
Sequences	X	X	X	
Mapping	X	X	X	
Crossword Puzzle	X	X		
Cloze	X	X		
Free Text		X	X	X

Table 1. Mapping of task types to cognitive process categories using the factual, conceptual and procedural knowledge category; (P) means that the task type for the corresponding taxonomic level is not suitable for mapping the procedural category.

For the implementation of the IPA further requirements have to be considered which result from the auditory implementation as NUI and will be discussed in more detail in the next section.

2.2 Requirements and Learning Design of a Virtual Tutor

For the auditory implementation of the virtual tutor particular aspects concerning the dialogue design have to be considered. The dialogue design is fundamentally determined by the *high-level design*, in which the dialogue strategy is defined, and the *detailed design*, in which the differentiation of elements takes place [7]. According to the high-level design, strategies for error tracking are to be taken into account through the use of so-called universals, functions such as Help, Repeat, Go-Back, which can be called at any point in the dialogue. Finally, with regard to the detailed design attention must be paid to minimize the cognitive load. For this purpose, as few elements as possible are to be kept in the working memory at all times. This can also lead to a division of learning tasks into individual sub-tasks. Further requirements resulting from the NUI design include a natural linguistic

output of the system and a dialogue that is as error-free as possible and limited to the essentials. Based on the already mentioned Design Recommendations from the field of ITS [4], the requirements are summarized in the use case diagram in fig. 1.

A total of 28 functional and four non-functional requirements (usability, portability, maintainability and interoperability) were considered. In order to support independent learning, attention was paid to the transferability of the solution and the free adaptability of the didactic design. For this reason a modular and loosely coupled structure is proposed using platform-independent technologies and enforcing the use of recognized standards.
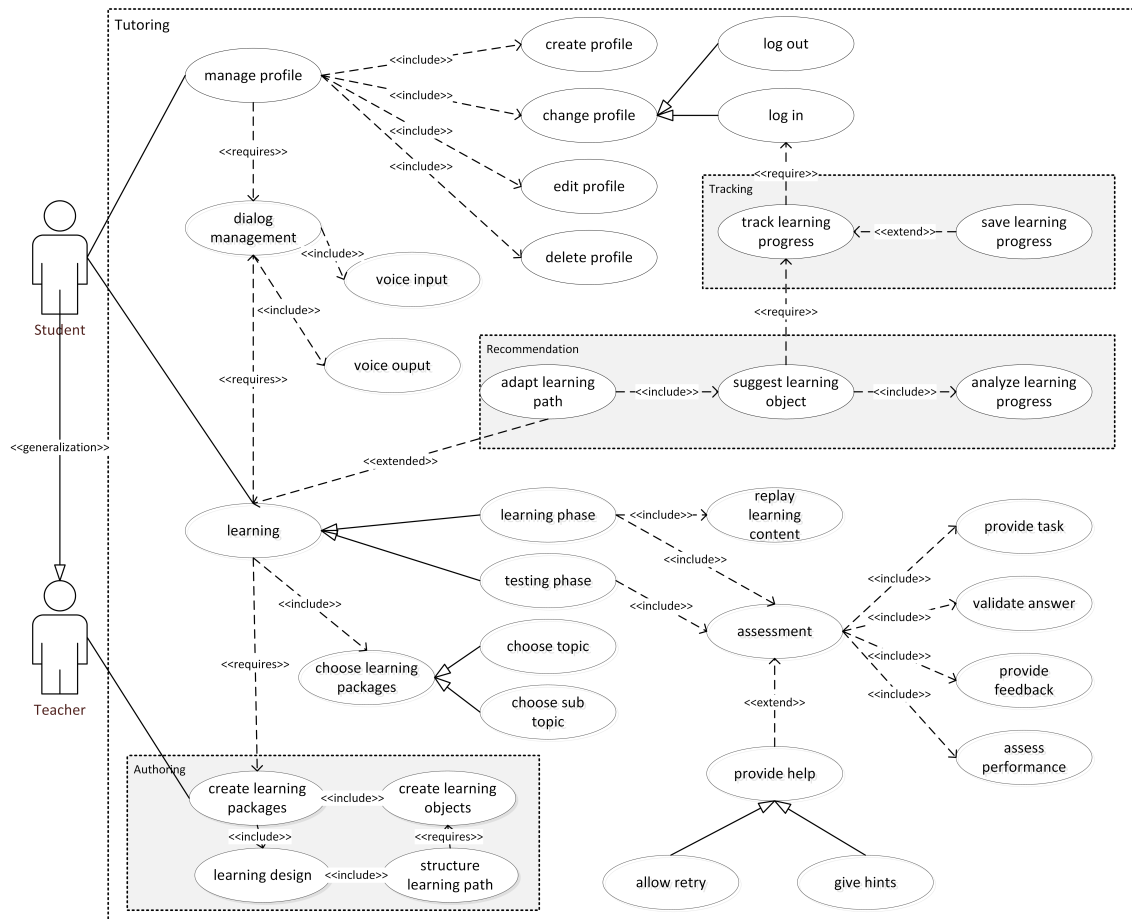

Fig. 1. Use Case diagram grouped by functions

2.3 Target Platform – Amazon Echo

The Amazon Echo is a smart speaker based on the IPA Alexa. The device has a microphone array, which is permanently switched on and filters incoming speech signals according to a defined activation word ("Alexa"), unless it was set to the mute mode. The input is transferred to the Alexa Voice Service (AVS) (usually abbreviated as Alexa) via the REST API (see Fig. 2). AVS is service/agent behind Amazon Echo's hardware, embedded as a part of Amazon Web Services (AWS). The audio recording is sent to the speech analysis component (SLU), where automatic speech recognition including semantic

analysis takes place. After a successful analysis, a suitable function (so called skill) is searched for and assigned to the Alexa Skill Service. A distinction is made between so-called *Custom Skills, Smart Home Skills, Video Skills* and *Flash Briefing Skills* [9]. In particular, *Custom* and *Smart Home Skills* are of interest for our scenario, because with their help developers can extend the possibilities of Alexa. For example, data can be retrieved from the Internet via web services. Possible outputs are synthesized using the Text to Speech (TTS) module and returned to the user auditory.
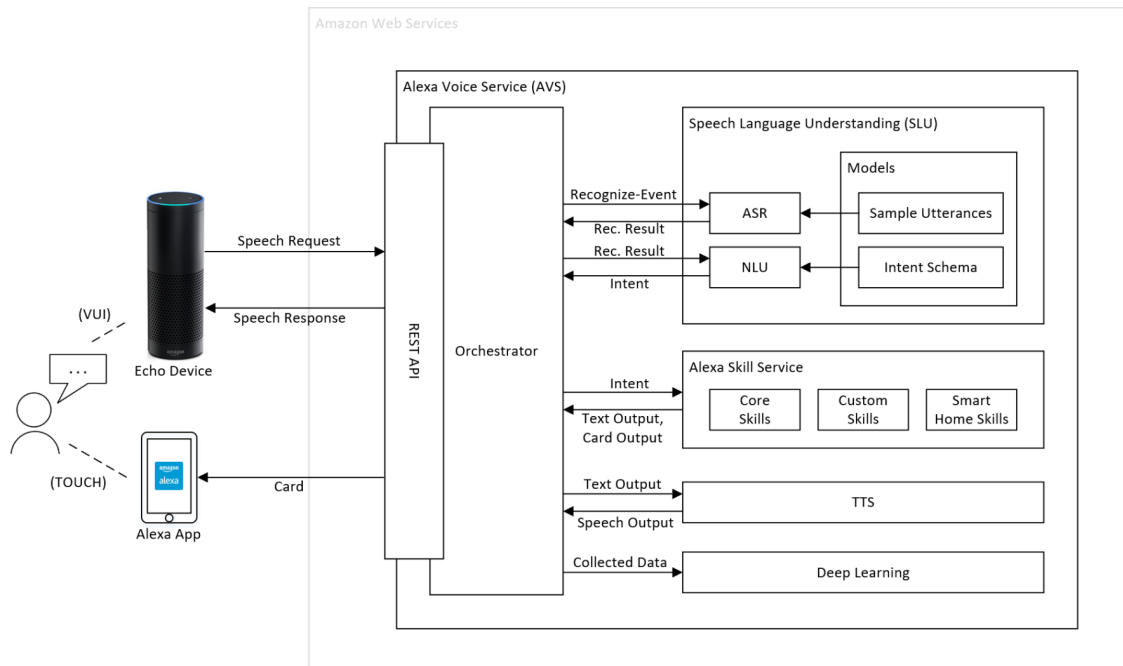

Fig. 2. Sketch of the Alexa Voice Services (AVS) with the related components Speech Language Understanding (SLU), Alexa Skill Service and TTS und Deep Learning

Outputs can also be made using the Alexa App or other Smart Devices. Finally, the processed data is passed on to the AWS Deep Learning Infrastructure to expand Alexa's knowledge base and sustainably improve the speech recognition rates and the user experience [9].

3 Implementation of the IPA-based tutor

The implemented prototype consists of six loosely coupled components (see Fig. 3). The authoring system (1) is used to submit learning packages, learning objects and didactic design. These are then persisted in a database (2) specialized in XML files via the central RESTful web service (6). The IPA Client (3) can use the web service to update user profiles and to retrieve stored learning data from the database. A Learning Record Store (LRS) (4) is used to store the learning progress using xAPI statements. The overall architecture is completed by a recommendation component (5) for proposing suitable learning objects.
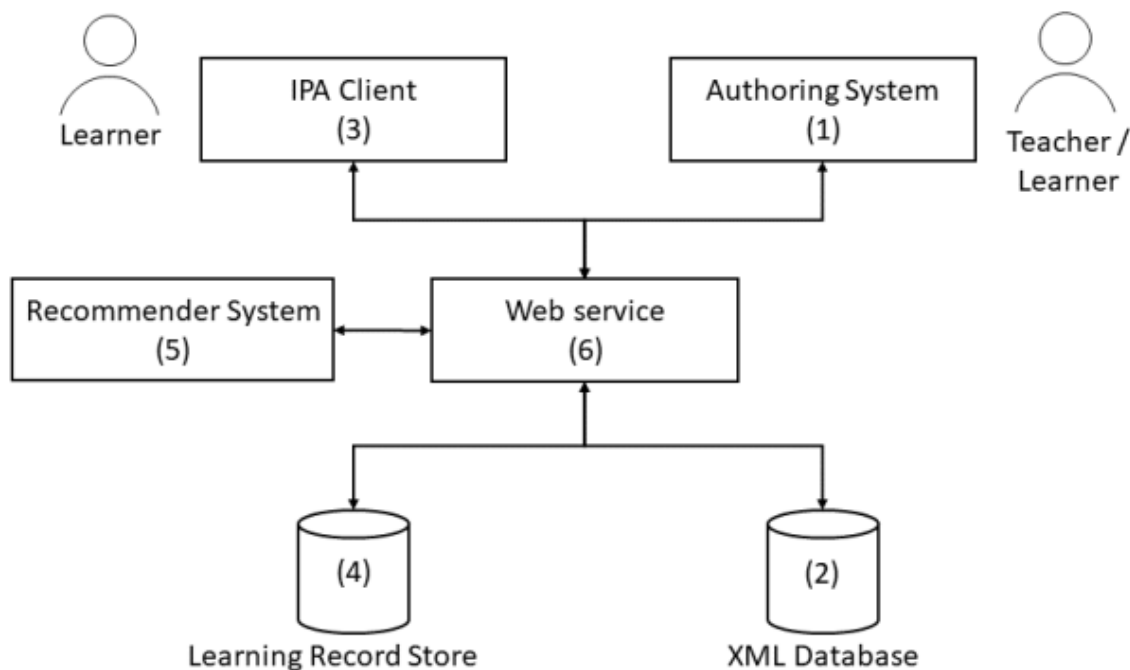
Fig. 3. Architecture consisting of authoring system, xml database, IPA client, learning record store, recommender system and a RESTful web service

Only recognized standards were used to ensure reusability and adaptability of the system. The standards *Sharable Content Object Reference Model* (SCORM) and *Common Cartridge* (CC) by the *IMS Global Learning Consortium* (IMS) were shortlisted. Unlike SCORM, CC integrates the *IMS Question and Test Interoperability* (QTI) specification. With *IMS Learning Design* (LD), CC also offers the option of mapping the didactic setting and with the *IMS Learner Information Package* (LIP) the "learning outcomes" of learners. These functions are essential for the implementation of independent learning and assessment functions. The learning packages are mapped with *IMS Content Packaging* (CP) and the individual learning objects with QTI. IMS MD fully integrates *the IEEE Learning Object Metadata* (LOM) specification and adds further information such as language, keywords, subject area, format and difficulty to the learning objects specified in QTI. Therefore the IMS MD specification is used to manage the metadata. In the following sections, the single components of the system are now described in more detail.

3.1 Authoring system and XML database

The authoring system is used to create and maintain learning packages, learning objects and the learning design in the corresponding formats CP, QTI and LD. It is kept minimal in the prototype and has no frontend. Content objects are imported as a ZIP archive, then checked for syntactic correctness using the IMS XML schema definitions and transferred to the RESTful web service for subsequent storage in the database. In addition, there is a simplified way of creating learning objects without the need to use external authoring tools (e. g. TAO6). For this purpose learning objects can be stored in the form of JSON files,

where the associated data is converted to the specifications using a template engine. An example of a used question with the task type "sequence" is listed below in the compressed JSON format and the associated QTI format.

```json
{
    "id": "LA_0011",
    "title": "LA_0011",
    "objectives": [{"title": "none"}, {"title": "none"}],
    "prerequisites": [{"ref": "none"}],
    "description": "Aus der Vogelperspektive ist Deutschland gruener als man denken mag.",
    "feedback_pos": "Gut aufgepasst! Mach weiter so.",
    "feedback_neg": "Schade, das ist stimmt leider nicht ganz. Auf Landwirtschaft folgen Wald, Siedlungen
                     und zuletzt das Wasser. Das war aber auch nicht leicht.",
    "type": "test-object",
    "file": "qti_LA_0011.xml",
    "size": "0",
    "difficulty": "easy",
    "language": "de",
    "author": "D.W.",
    "creationdate": "17.10.2017",
    "keywords": ["interaction:order", "topic:deutschland", "subtopic:deutschland", "difficulty:easy"],
    "next": [
        {"succeed": "true", "interaction": "fact", "difficulty": "easy", "ref": "LA_0012"},
        {"succeed": "false", "interaction": "fact", "difficulty": "easy", "ref": "LA_0013"}
    ],
    "content": {
        "interaction": "order",
        "question": "Wodurch ist in Deutschland die meiste Flaeche bedeckt? Sortiere absteigend!",
        "responses": [
            {
                "id": "resp0001",
                "response": "wald",
                "position": "2",
                "score": "1"
            },
            {
                "id": "resp0002",
                "response": "wasser",
                "position": "4",
                "score": "1"
            },
            {
                "id": "resp0003",
                "response": "siedlungen",
                "position": "3",
                "score": "1"
            },
            {
                "id": "resp0004",
                "response": "landwirtschaft",
                "position": "1",
                "score": "1"
            }
        ]
    }
}
```

Listing 1: json-lp-11.txt

```xml
1  <?xml version="1.0" encoding="UTF-8"?>
2  <assessmentItem xmlns="http://www.imsglobal.org/xsd/imsqti_v2p2" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3   xsi:schemaLocation="http://www.imsglobal.org/xsd/imsqti_v2p2  http://www.imsglobal.org/xsd/qti/qtiv2p2/imsqti_v2p2.xsd"
4   identifier="ITM_LA_0011" title="LA_0011" adaptive="false" timeDependent="false">
5      <responseDeclaration identifier="RESPONSE" cardinality="ordered" baseType="identifier">
6          <correctResponse>
7              <value>resp0004</value>
8              <value>resp0001</value>
9              <value>resp0003</value>
10             <value>resp0002</value>
11         </correctResponse>
12     </responseDeclaration>
13     <outcomeDeclaration identifier="SCORE" cardinality="single" baseType="float" />
14     <itemBody>
15         <orderInteraction responseIdentifier="RESPONSE" shuffle="true">
16             <prompt>Wodurch ist in Deutschland die meiste Flaeche bedeckt? Sortiere absteigend!</prompt>
17             <simpleChoice identifier="resp0001">wald</simpleChoice>
18             <simpleChoice identifier="resp0002">wasser</simpleChoice>
19             <simpleChoice identifier="resp0003">siedlungen</simpleChoice>
20             <simpleChoice identifier="resp0004">landwirtschaft</simpleChoice>
21         </orderInteraction>
22     </itemBody>
23     <responseProcessing template="http://www.imsglobal.org/question/qti_v2p2/rptemplates/match_correct" />
24 </assessmentItem>
```

Listing 2: qti_LA_0011.xml

All authoring tools that support IMS CP, IMS LD Level B and IMS MD are suitable for creating the learning objects. IMS LD Level B is required to map the adaptive learning paths.

Since all data is available in XML format, the document-oriented XML database BaseX was chosen. At the time of implementation in 2017 BaseX was the only representative, which fully supported the *W3C's XPath* and *XQuery Recommendations*. This is necessary for storing and querying individual elements, which is particularly relevant for accessing individual QTI objects within learning packages.

3.2 Client of the Intelligent Personal Assistant - the Alexa skill

The IPA client extends Alexa's language assistant by functions that can serve as a virtual tutor for the user. For the used product Amazon Echo an Alexa skill was implemented, which represents the essential functions of the user interface. Part of the skill is the modeling of the *interaction model*, which contains all dialog directives. An interaction model consists of *intents, utterances, slots* and *slot types*. In simplified form, intents stand for the user's intentions, utterances are statements or formulations to which Alexa should react and call the associated intents. Slots are again parameters in utterances of a data type. Slot type can be, for instance, DE_City or DE_First_Name. In addition to fixed slot types, custom slot types can be implemented. In the following, Tab. 2 shows the fully implemented Interaction Model.

Intent	Example-Utterances	Slots
WhoisIntent	"I am [name]", "My name is [name]", "[name]"	name
LoginIntent	"Sign in [name]", "Please log in [name]", „Log in [name]"	name
LogoutIntent	„Log out", "Please, log me out", "Sign off"	no slot
YesIntent	"yes", "yep", "okay"	no slot
NoIntent	"no", "nope"	no slot
RecommendationIntent	"select topic [topic] ", "select [topic]", [topic]	topic
AnswerIntent	" answer [answer]"	answer
ContinueIntent	"next", "skip", "following"	no slot
HintIntent	"hint", "tip"	no slot
RepeatIntent	"Repeat", "Again", "I did not understand you"	no slot
HelpIntent	"Help", "Help me", "What should I do"	no slot
CancelIntent	"Cancel"	no slot
StopIntent	"End", "Stop", "Quit", "Finish"	no slot

Table 2. Interaction Model of the IPA client (Alexa skills) with intents, utterances and slots

To implement an Alexa skill either so called One-Shot-Models or Invocation-Models can be used. Using a One-Shot-Model the interaction with the IPA is finished after an input by the user followed by an output of the system. Contrary, the Invocation-Model allows it to aggregate several One-Shot-Models one after another to realize dialogs with the IPA. More complex questions can lead to an increased thinking time and therefore to a longer response delay of the user. If the user does not respond within 10 seconds the Invocation-Model of the Alexa skill is automatically closed, which leads to a loss of information according to the user and his current preferences, selected topic and current learning task. This limitation of Alexa is avoided by the implementation of a session manager, which makes it possible to resume a learning session by just a log in (cp. Resume&Pause.mp3). All possible interaction states are listed in fig. 4, which summarizes the functionality of the Alexa skill in combination with the interaction model presented above.
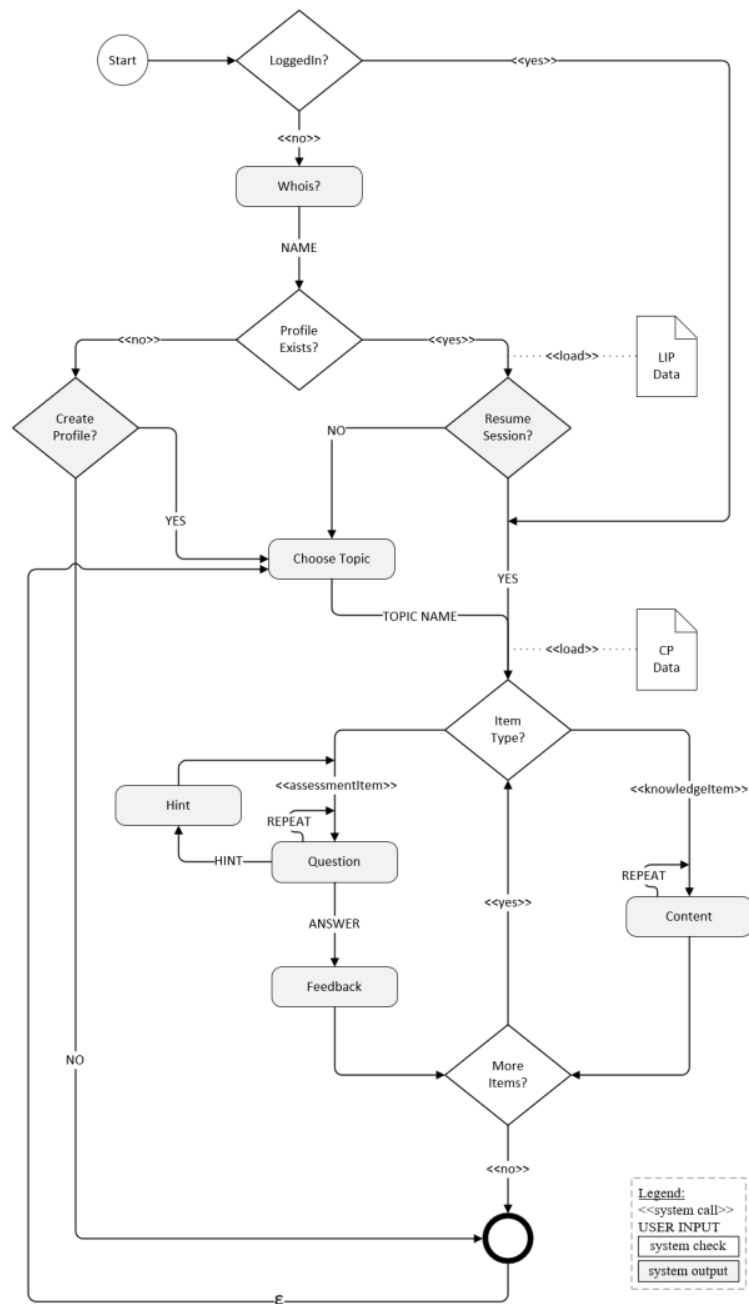
Fig. 4. Call-Flow diagram indicating the different states and used input for learning and assessment

3.3 Learning Record Store & Recommender System

An LRS is used in the overall architecture to store the learning progress achieved. The progress is stored as an xAPI statement. For the existing prototype the LRS of Learning Locker was used, since it fulfills all necessary specifications according to ADL and offers an

extended API to aggregate results. In the prototype it is sufficient for the learner to be the sole actor, i.e. to consume the learning objects. Additionally, the context and outcome are saved. Therefore, an xAPI statement of the following form is used:

[Name] experienced [LearningObject] in context [Topic, Difficulty, Type] with [positive/ negative] result.

This data is used by a recommender system to suggest suitable learning packages to the learner. The title of the desired topic and the name of the learner is transferred to the RESTful web service, which provides the learner information in the LIP format and the learning package in CP format. The learning design is extracted from the organization section of the learning package. Then, using conditions of the individual learning objects, a graph is generated that generates a *Recommendation Value* based on the LRS information. It is calculated from the users learning performance (ratio of correct and incorrect answers according to the task type and the attempts) and the defined difficulty of the learning object.

3.4 Web service as a central piece for connecting the individual components

The web service represents the central communication component for the data exchange in the architecture. For the implementation we decided to choose the lightweight Python Flask Framework . The Alexa client was also implemented with the Flask framework because it allows the use of the Flask-ASK framework, which fully implements the Alexa Skills Kit. A total of ten functions (see Fig. 5) were implemented that allow the retrieval or creation of elements from the XML database, the recommendation system and the Learning Record Store.

POST	/api/topic/titles get all available topics
POST	/api/learner/name persist LIP package
GET	/api/learner/name get LIP package of a user
POST	/api/learner/name/activity persist last used activity of a user
GET	/api/learner/name/activity get current activity of a user
GET	/api/recommendation/name/topic get recommendation for user
POST	/api/learner/contentpackage/topic persistent IMS contentpackage
POST	/api/item/file persist single learning object
POST	/api/lrs/statement update learner progress
GET	/api/lrs/statement return learner progress

Fig. 5. Available endpoints of the RESTful web service

4 Evaluation of the user acceptance

The implemented solution is a novel scenario using a relatively new technology making it difficult to compare or classify effectiveness. Thus, the acceptance of the users regarding the use of the IPA for independent learning was first evaluated. It should be determined whether the users accept the conceived teaching and learning scenario with a virtual tutor at all and whether the prototype can be used for further developments.

In acceptance research the Technology Acceptance Model (TAM) has been proven to be a fundamentally reliable model for recording user acceptance with its various extensions (TAM 2/TAM 3) [10]. The TAM 2 extends the original TAM by adding determinants from the categories social inputs and cognitive instrumental processes. The social inputs include the input variables Subjective Norm (SN), i.e. the perception of the social environment, the social status (Image - IMG), the perceived influence of the social status through the use of technology, the experience (EXP) and the voluntariness (VOL). The category of cognitive instrumental processes includes job relevance (REL), i.e. the perception of how much the technology supports the task, the output quality (OUT) and the result demonstrability (DEMO).

The actual use or acceptance depends significantly on the two variables perceived usefulness (PU) and perceived ease of use (PEOU). Usefulness is the degree to which a person thinks that the use of the technology improves labour productivity. Usability describes the perceived effort required to learn. The variables of social status (IMG) and

subjective norm (SN) were not considered for the study design. For the survey nine hypotheses, one for each variable, were stated. An adapted questionnaire of Venkatesh and Davis [10] was used and adapted to the study area of the virtual tutor (see Tab. 3).

Det.	Item	Question	x̄	σ
ITU	ITU	Assuming that the virtual tutor is available in my smart home, I would use it.	2,00	0,71
PU			2,67	0,58
	PU1	The use of the virtual tutor creates an added value for me in domestic learning.	2,50	0,87
	PU2	The use of the virtual tutor in domestic learning improves my productivity	3,08	0,64
	PU3	The use of the virtual tutor increases the effectiveness of my domestic learning process.	2,75	0,60
	PU4	I think the use of the virtual tutor in domestic learning makes sense.	2,33	0,62
PEOU			2,27	0,58
	PEOU1	The interaction with the virtual tutor is simple and understandable.	2,00	0,58
	PEOU2	The use of the virtual tutor does not require any specific mental effort from me.	2,33	0,62
	PEOU3	I think the virtual tutor is easy to use.	2,42	0,76
	PEOU4	The virtual tutor allows me to get the information I expect.	2,33	0,62
VOL	VOL	I used the virtual tutor voluntarily.	1,25	0,43
EXP	EXP	I have experience in using voice-controlled systems.	2,25	0,83
REL			3,33	0,66
	REL1	The use of the virtual tutor is important for domestic learning.	3,67	0,85
	REL2	The use of the virtual tutor is relevant to domestic learning.	3,00	0,58
OUT			2,42	0,57
	OUT1	The output quality I get from the virtual tutor is high.	2,50	0,76
	OUT2	I have no difficulties with the output quality of the virtual tutor.	2,33	0,47

DEMO			2,42	0,53
	DE-MO1	I have no difficulties explaining the benefits of a virtual tutor to others.	2,58	0,49
	DE-MO2	The benefits of using the virtual tutor seem obvious to me.	2,25	0,72

Table 3. Mapping of defined items to determinants and the corresponding question for the items; Mean values and standard deviation of the sampling

The acceptance tests were conducted on 12 tech-savvy subjects (seven male, five female). At the beginning, the subjects were introduced to the basic handling of Amazon Echo and Alexa. A test series (cp. Demo.mp3) consisted of 18 learning objects and lasted about 25 minutes. The used learning design and the corresponding learning objects are listed in the content package [11] (cp. folder content). The missing elements like the positive and negative feedback can be observed in the full content package.
The questions were then completed in the form of an online survey with a Likert scale of 1 (strongly in favor) to 5 (disagree).

The perceived usefulness (PU) was evaluated with a mean of 2.67 ($\sigma = 0.58$). Thus, the prototype could not convince clearly but is perceived as rather useful. The perceived ease of use ($\bar{x} = 2, 27$; $\sigma = 0.58$) suggests that the system is considered to be relatively user-friendly. The intention to use (ITU; $\bar{x} = 2$; $\sigma = 0.71$) indicates an intention to use with given availability. The relevance (REL) with a mean value of 3.33 ($\sigma = 0.66$) is rather negative, so that the virtual tutor does not seem to have any decisive relevance for learning (see Tab. 3 for all results).

For a better explanation of these input variables a correlation analysis according to Bravais-Pearson was then performed. The correlation results relevant to the hypothesis test are shown in Fig. 6. An additional regression analysis was carried out to identify the variables that make the highest contribution to the Intention to Use (ITU). All variables are suitable as predicator as expressed by the adjusted coefficient of determination ($⟦r-⟧\ ^2$) (see Fig. 6).
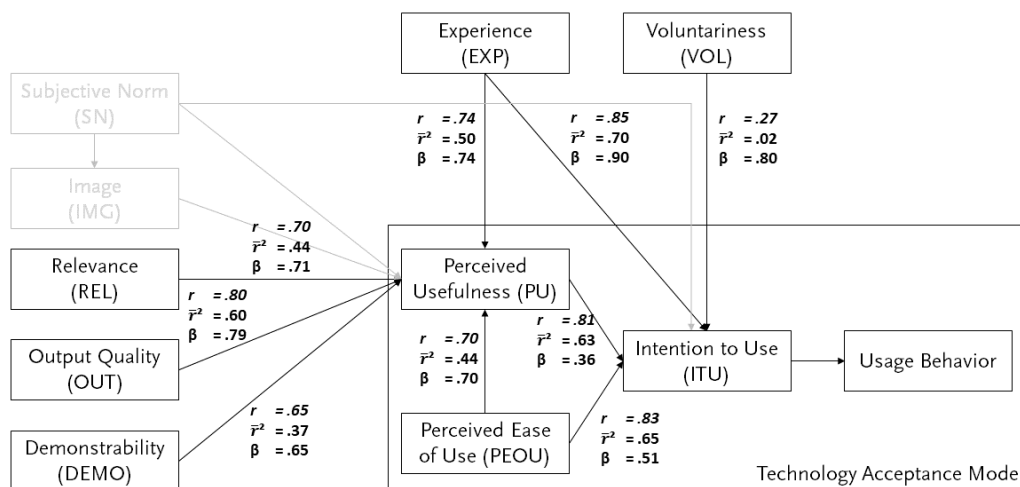
Fig. 6. Adapted TAM2 with regression coefficients and corresponding correlation coefficients

The intended use (ITU) depends significantly on the previous experience (EXP; β = 0.90), the perceived added value (PU1; β = 0.88) and the perceived ease of use (PEOU3; β = 0.87).

However, a strong limitation of this study is that only 12 subjects were included. Therefore, no valid statistical conclusion can be drawn. The results are thus only to be understood as reference values and must be supplemented by further test subjects.

5 Summary and Outlook

An IPA based on Amazon Alexa was designed and as a proof of concept implemented and evaluated. The technical and didactic principles, the requirements and the main components of the resulting architecture were briefly outlined. For developing the prototype, the latest technologies and recognized standards were used. The prototype represents an innovative approach in using the existing pervasive possibilities to design native teaching and learning scenarios in the form of a virtual tutor. The modular architecture ensures that a substitution of Amazon Alexa by an open-source component like Mycroft Mark II or Jasper is possible to address data privacy concerns.

The evaluation results using the technology acceptance model (TAM) should be viewed with caution due to the small number of test subjects. If the system is available, however, there seems to be an increased intention to use it. Amazon Echo with a predefined interaction model has limited possibilities to implement a free dialogue design. Thus, only short conversation scenarios and the realization of smaller learning units seem meaningful. Other limitations refer to the predefined interaction model and slot types, which limits the virtual tutor by compile time. Since the natural language unit uses pattern matching by

probability calculation, wrong answers can be detected as correct (e.g. questions using the task type sequence with five answer options where the first three are reproduced by the learner in the correct order and the fourth and fifth option are in the wrong order).

The proposed prototype provides a narrowly defined range of functions. Enhancements are conceivable in which the IPA controls additional devices in the Smart Home in order to create a learning-friendly climate in the context of the physical and virtual environment. Different environmental settings for learning or for testing are possible. Further developments could deal with the adaptation and the integration of the environment. It is conceivable, for example, that the brightness level could be automatically adjusted by controlling darkening, that disturbing or distracting devices are switched off or put into mute mode, or that smart devices (tablet, smart TV) relevant for the respective learning scenario could be started. By including and analyzing the learning progress, it would also be conceivable to suggest breaks from learning or to offer further learning aids. In order to expand the didactic and recommendation functions, the Deep Learning infrastructure can be used for future work. When dealing with commercial cloud infrastructures, data protection aspects must still be taken into account.

6 References

1. Satpathy, L.: Smart Housing: Technology to Aid Aging in Place. New Opportunities and Challenges. Master's Thesis, Mississippi State University, Starkville, MS, USA, 2006.

2. Abicht, L.; Spöttl, G.: Qualifikationsentwicklungen durch das Internet der Dinge: Trends in Logistik, Industrie und Smart House". Bertelsmann, 2012.

3. Wolff, A.: Smart Home Monitor 2017, 27. März 2018. https://www.splendid-research.com/smarthome.html (last visit 2019-08-22).

4. Sottilare, R.; Graesser, A.; Hu, X.; Goodwin, G.: Design Recommendations for Intelligent Tutoring System - Volume 5: Assessment Methods, 2017. https://gifttutoring.org/attachments/download/2410/Design%20Recommendations%20for%20ITS_Volume%205%20-%20Assessment_final_errata%20corrected.pdf (last check 2019-08-22)

5. Mayer, H. O.; Hertnagel, J.; Weber, H.: Lernzielüberprüfung im eLearning. deGruyter Oldenbourg, München, 2009.

6. Strzebkowski, R.: Selbständiges Lernen mit Multimedia in der Berufsausbildung - Medien-didaktische Gestaltungsaspekte interaktiver Lernsysteme. Diss., Berlin, 2006.

7. Cohen, M. H.; Giangola, J. P.; Balogh, J.: Voice User Interface Design. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 2004.

8. Anderson, L. W.; Krathwohl, D. R. (Eds.): A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives. Longman, New York, 2001.

9. Amazon: Build Skills with the Alexa Skills Kit, 27. März 2018, url: https://developer.amazon.com/docs/ask-overviews/build-skills-with-the-alexa-skills-kit.html (last check 2019-08-22).

10. Venkatesh, V.; Davis, F. D.: A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. In: Manage. Sci., Vol. 46/2, Feb. 2000, pp. 186-204, url: http://dx.doi.org/10.1287/mnsc.46.2.186.11926 . (last check 2019-08-22)

11. Kiy, A.; Wegner, D.: University-of-Potsdam-MM/Alexa: Alexa Skill (Version v1.0). Zenodo, 2018. http://doi.org/10.5281/zenodo.1441158 . (last check 2019-08-27)