New ITG Guideline for the Usability Evaluation of Smart Home Environments

Sebastian Möller, Klaus-Peter Engelbrecht, Stefan Hillmann, Patrick Ehrenbrink

Quality and Usability Lab, Telekom Innovation Laboratories, TU Berlin, 10587 Berlin, Germany
Email: {sebastian.moeller,klaus-peter.engelbrecht,stefan.hillmann,
patrick.ehrenbrink}@telekom.de
Web: www.qu.tu-berlin.de

Abstract

As a complement to the 2011 ITG Guideline 2.1-01 on the evaluation of terminal devices, a new ITG guideline on the usability of smart home environments has been drafted. This guideline addresses services and applications which allow for an implicit interaction between user and system taking place in an environment equipped with networked sensors and actuators, partially also without any dedicated user interface. The present paper reviews the current draft of the guideline by describing its scope, the underlying system set-up, the usability aspects and the evaluation criteria which might be considered relevant.

1 Motivation and Introduction

In 2011, the ITG has agreed upon a guideline for the evaluation of terminal devices for communication applications [1]. This guideline provides an introduction to quality and usability as well as their underlying dimensions, classifies applications according to their function, usage scenario, and input- and output modalities, and further provides an introduction to evaluation methods which are useful in the product design cycle, following the usability engineering lifecycle [2]. The guideline than provides practical recommendations on how to select an appropriate evaluation method for a given purpose, how to design evaluation tasks, and how to carry out an evaluation is a proper way. Practical application examples of evaluations conclude the guideline.

Whereas the 2011 guideline might provide helpful information for typical applications running on terminal devices (such as smartphone apps), it is limited when it comes to applications or services which are integrated into an environment, potentially equipped with sensors and actuators which are interconnected via an IT network. Such applications frequently allow for *implicit* interactions between user and systems, where the effect of a user action is perceived only via a change in the smart environment, and where system actions are provoked implicitly by sensors reacting to changes of the environment itself, or to user actions which are not explicitly directed to the system. Sometimes, services do not even provide a dedicated user interface which would allow for a proper evaluation, but are fully integrated into the environment. Proactive behavior of the service together with the absence of a user interface can then lead to the impression that the environment may be *smart* or *intelligent*. Recent developments lead to multimodal interaction concepts, often one modality enhanced with capabilities of speech.

In order to cope with the limitations of the 2011 guideline, researchers and practitioners gathered around the department “services and applications” of the ITG (ITG FB 2) to start work on a new guideline related specifically to the evaluation of smart home environments. This guideline is conceptualized as a complement to the existing guideline, but should be formulated in a way which makes it understandable as a stand-alone document to practitioners, like designers and suppliers, who are the main intended addressees. The current draft of the guideline [3] is the outcome of a series of expert workshops which have taken place in 2013, as well as open calls for contributions launched through the ITG website. It has been submitted to the ITG FB 2 for further discussion, to be approved by the ITG later in 2014.

It is the aim of the present paper to provide a synopsis of the rationale and the content of this new ITG guideline. Although the guideline does not specifically address speech as an interaction modality, many of its considerations also refer to speech. We will start with a definition of the scope of the guideline in Section 2. Section 3 provides a schematic view of the interaction between a user and a smart home system used throughout the guideline. Section 4 addresses different aspects of usability which are specifically connoted for smart home systems. Section 5 addresses criteria which can be used for an adequate evaluation. Section 6 concludes with a list of evaluation methods addressed by the guideline, as well as open questions conceivable for future work.

2 Guideline Scope

The guideline addresses aspects of the *usability* of smart home environments. In doing so, it takes a user-centered viewpoint and does not consider purely functional testing of software or hardware components. The aspects of usability which are covered by the guidelines are further described in Section 4.

The applications or services which are covered by the guideline are meant to be installed in a *smart home* environment. Such an environment is thought to be equipped with sensors and actuators, including complete devices or appliances (such as entertainment devices, household appliances), which are interconnected by an IT network. In this environment, *services* are offered which provide different functions to the user or to other system components. Services may also address meta-functions such as the administration of the smart home environment itself. The smart environment might not necessarily be limited to a (private) home, but may also be used in a work context, although aspects purely relevant for work environments (such as productivity or employees) are not covered by the guideline.

In such an environment, the interaction between user and system/environment may happen in an *implicit* way, meaning that the user does not necessarily intentionally interact with the system. Instead, system actions might be

triggered by sensors reacting to changes in the environment, with and without intervention from the user. For example, a smart home environment may decide to switch on the light because a daylight sensor registers reduced daylight, or because the user opens the front door which is equipped with a sensor. In both cases, the system action is provoked without the user operating a dedicated *interface*. In fact, smart home environments may even work without any user interface, rendering classical methods of interface testing difficult to apply.

The word "decide" used in the paragraph above may imply that the system is sometimes considered to be *intelligent*, in that it has cognitive abilities and takes decisions on behalf of the user. Whereas the properties of intelligence are still debated, it is important to note here that a system may actually be considered as "intelligent" by its users, thus intelligence becomes a criterion for the evaluation, see Section 5. A system may also behave in a *persuasive* way, in that it attempts to induce a specific behavior in the user, e.g. to do more sports or to save energy; such persuasive services are also covered by the guideline.

Smart home systems have frequently been used in an Ambient-Assisted Living (AAL) scenario, where certain medical, care-taking and/or comfort functions are provided to users with reduced abilities. Whereas such scenarios are not excluded in the guideline, it has to be noted that particular usability aspects and criteria which are relevant for AAL (such as user safety, reliability of medical treatments, ethical aspects of care-taking by a machine) are not covered.

3 Interaction

The guideline considers three actors in smart home environment, see Fig. 1: The user, the smart home and the service.

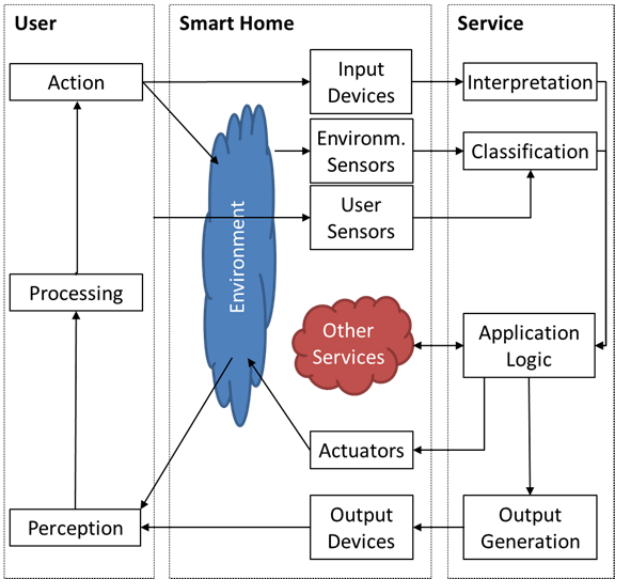

Figure 1: Schematic representation of an interaction in the smart home environment.

A user interaction with the system may be triggered in a classical way through input and output devices. As an example, a microphone integrated in the smart home may be considered as an input device for speech, and a loudspeaker as an output device. Input and output may however also take place through general-purpose devices such a smartphones (with microphone and loudspeaker integrated), with general-purpose user interfaces. Input to the system may also stem from environmental sensors, or from user-mounted sensors which cannot be considered as user interfaces, such as position sensors, heartbeat sensors, and alike. In a similar vein, the output from the system may refer to actuators which influence the environment, and may later be perceivable by the user or not (such as switching on the light, or changing the heating setting). Services may also communicate with other services (e.g. placing an order, asking for human intervention).

Depending on the way input and output take place, it is advisable to distinguish between *implicit* and *explicit* interaction. An explicit interaction from the user is triggered through an input device, whereas an implicit interaction may be triggered through sensors. The user may also use sensors for explicit input (e.g. shielding a brightness sensor to provoke switching the light on).

In order to cope with a multitude of input and output capabilities, the service needs to interpret and fuse the information coming from the user, and to classify the information coming from the input sensors. The semantic information is then used by the application logic to fulfill the service, and to further interact with the user. Both targets can be reached through the output generation logic, which should determine also the best modality for information output (fission).

4 Usability Aspects

The aim of the guideline is to provide practical guidance to usability evaluation in the smart home. ISO 9241-11 [4] defines usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". This definition mainly considers the functional part of it, using effectiveness (i.e. the accuracy and completeness with which the functions are fulfilled) and efficiency (i.e. the resources spent in relationship to the function fulfillment) as criteria. User satisfaction is considered as a result of function fulfillment, but not further specified in that standard.

In order to complement this traditional view on usability with non-functional, sometimes called *hedonic* aspects [5], the concept of *User Experience (UX)* has been brought up. ISO 9241-210 [6] takes a comprehensive view and defines user experience as "a person's perceptions and responses that result from the use or anticipated use of a product, system or service". Although this view might be very broad [7], it shows that classical perspectives might be too limited in covering all aspects which are considered to be decisive for a user to successfully use a service offered in a smart home context.

The guideline further differentiates between *ease of use* and *joy of use*. Whereas ease of use describes the subjective perception of the user that the system can be used without any effort [8], and includes the aspects effectiveness, efficiency, intuitiveness and learnability, joy of use [9] describes the conscious positive experience during the interaction. Such an experience is influenced by the aes-

thetics, the discoverability and the personality of the system, the latter being particularly relevant for systems enabling speech interaction.

In the smart home environment, it is a characteristic that services are used in an intermitted way, sometimes over a longer period of time, either by one user (who might get used to the services) or by many users (sometimes convening in the same environment, making a separation of users difficult). This renders the following usability aspects of major importance:

- *Consistency*: Different services offered in the same environment should be offered in a consistent way.
- *Transparency*: The user should be able to understand system behavior in order to identify and correct errors. The control over the system should in the ultimate instance remain with the user.
- *Obtrusiveness*: The changes in the environment provoked by the system should not bother the user.
- *Personalization*: The environment should adapt or be adaptable to preferences and requirements from different users, without being inconsistent.
- *Absence of barriers*: Especially with respect to users with particular abilities or competences.
- *Adequacy to multiple users*: With respect to multiple users engaged in the environment simultaneously, but also with respect to users with different levels of expertise.
- *Trust* and *Security*: The user should have trust in the system and feel secure and safe when using it.
- *Robustness* of the input and interaction logic.

With respect to the usage context in the home, the following additional requirements on the environment have to be taken into account:

- It should address real-world needs of its users.
- It should operate on devices and the environment securely.
- It should respect the users' privacy requirements.
- It should respect the users' feeling of being at home.
- It should be designed in a way which excludes as far as possible negative consequences with respect to changes in the users' behavior.

The usability engineering lifecycle illustrates how user requirements can be taken into account during the entire development process of a new smart home service, consisting of analysis, design, prototyping, expert evaluation, empirical testing, iterative design, and feedback from field. This approach motivated by Nielsen [2] was already proposed in [1].

5 Evaluation Criteria

Services offered in a smart home may span over a variety of domains. Depending on the domain, some criteria may be more important than others, so that they dominate the perception of the user. In addition, there are criteria which are of importance for special types of services offered in the smart home.

The following domains are explicitly mentioned in the guideline, together with specific evaluation criteria:

- *Communication services*, including telephone or messaging. As these services are frequently used by many users in the home, absence of barriers, multi-user adequacy and personalization are important criteria. Security and privacy also need to be respected.
- *Safety services*, such as protection against theft and fire. These services should be reliable, testable, and transparent (with respect to their operation) also to naïve users.
- *Energy management services*, e.g. for controlling electricity consumption or heating. These services should be consistent across devices of different providers, and should be transparent with respect to their effect. As they can be expected to be used less frequently, they should be easily learnable, and perhaps be personalized to different users.
- *Control services* for different household appliances, e.g. for operating lamps, blinds, etc. As these services can be expected to be used frequently, effectiveness and efficiency play a major role. In addition, they should be adequate for multi-user operation, for personalization, and they should be transparent to their users. Joy of use may be an additional criterion for such comfort services.
- *Entertainment services*, e.g. for operating the TV, music collection, etc. As above, these services should be effective, efficient, personalizable, and adequate for multi-user operation.
- *Health and fitness services*, e.g. for organizing training units and persuading the user of health-adequate behavior. As these services may have fatal consequences, safety, security and privacy are key criteria. The services should also be informative for their users, in providing relevant information about user behavior and potential consequences.

In addition to domain-specific criteria, there are criteria which are adequate for particular types of services. The following services are mentioned in the guideline:

- *Intelligent Services*: The criterion of "intelligence" is mainly interesting in a phenomenological sense, i.e. whether a user considers the service to be intelligent or not. Adaptivity, use of sensor data or understanding of spoken language (see below) may lead to such a perception, as well as the ability to understand complex phenomena or to interpret complex information.
- *Adaptive Services*: This includes adaptation to the context of use and the usage environment, as well as the adaptation to the user (including preferences, requirements and abilities). Adaptive behavior may result in reduced predictability and transparency, reduced controllability, intrusiveness, and reduced privacy [10]. Adaptive systems are difficult to evaluate, as traditional paradigms which suppose that the behavior of a system is only influenced by explicit interaction from the user fail. Instead, it may be more appropriate to decompose an adaptive system into its underlying components and evaluate each of them separately [11][12].
- *Persuasive Services*: In addition to the criteria mentioned above, persuasive interfaces should also be evaluated with respect to the effect they have on the behavior of the user. The effect may depend on how information regarding the consequences of the user's own behavior is displayed to him/her, and whether appropriate intrinsic motivators are found.

- *Sensor-based services*: These should react to changes of the environment or of the user, without the user's explicit interaction. Important criteria for such services are not only whether the reactions of the system are correct, in the sense that they act according to the interaction logic of the system, but also whether the reactions are considered as appropriate by the user, whether they are unobtrusive, whether the systems remains controllable even in the absence of correct sensor information, etc.
- *Actuator-based services*: Actuators may serve the change of the environment or the information for the user, irrespective of whether this information is actually picked up. Particularly interesting is the case when the user is not able to perceive the system action; in those cases, it might be necessary to provide additional display possibilities to satisfy the controllability criterion.
- *Generic input and output devices*: Such devices allow different services to be controlled via a common interface. Common criteria are intuitiveness and consistency, apart from standard usability criteria.
- *Speech-based services*: Speech-based services for smart homes have been investigated e.g. in [13]. For these systems, aspects like intelligence, adaptivity and transparency have been in constant discussion. Relatively new is the fact that speech-based interfaces process increasingly implicit input, such as emotions, or personality automatically extracted from the speech signal. Thus, it is expected that the concepts addressed in the guideline will be of value for the spoken language processing community.

6 Conclusions and Future Work

In its main body, the guideline describes standardized methods which can be used for evaluating smart home environments according to the criteria stated above. These include cognitive walkthrough, heuristic evaluation, model-based evaluation, simulation, layered evaluation, thinking aloud, focus groups, Wizard-of-Oz tests, as well as field studies. These methods are briefly reviewed, relevant literature is given, and they are classified according to several criteria which facilitate the selection of an appropriate methodology by the practitioner.

Although the guideline is intended for designers and appliers of smart home environments, it shows that a number of fundamental research questions need to be answered before disposing of a full-blown set of evaluation methods, covering all relevant aspects. In particular, methods for evaluating adaptive systems are still in their infancy. Further work is also necessary in order to address perceived security, privacy and safety of smart home environments. Many of these aspects can only analyzed in sufficient depth via long-term studies, spanning over months or years, in order to cover changes and adaptations in user behavior which have to be considered in the evaluation as well. Finally, smart home systems can be expected to have consequences for the long-term behavior of users, and the consequences of such changes are still guessable at best.

As the writing of the guideline was a collaborative work by different experts, input and suggestions for improvement are welcome, through the authors or through FB 2 of the ITG. The authors gratefully acknowledge the input received from the editing group up to now. The work has been carried out in the frame of the ongoing projects Universal Home Control Interface@ConnectedUsability (funded by BMWi) and Forschungscampus Connected Technologies (funded by BMBF).

References

[1] VDE-ITG-Richtlinie ITG 2.1-01 (2011). *Messung und Bewertung der Usability von Kommunikationsendeinrichtungen*, Informationstechnische Gesellschaft im VDE, Frankfurt am Main.

[2] J. Nielsen, *Usability Engineering*, Academic Press, San Diego CA, 1993.

[3] Draft VDE-ITG-Richtlinie (2014). *Messung und Bewertung der Usability in Smart Home-Umgebungen*, Informationstechnische Gesellschaft im VDE, Frankfurt am Main.

[4] ISO 9241-11, *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) — Part 11: Guidance on Usability*, International Organization for Standardization, Geneva, 1998.

[5] M. Hassenzahl, N. Tractinsky, "User Experience - a research agenda", *Behavior & Information Technology* 25(2): 91–97, 2006.

[6] ISO 9241-210, *Ergonomics of Human-system Interaction — Part 210: Human-centred Design for Interactive Systems*, International Organization for Standardization, Geneva, 2010.

[7] N. Bevan, "What is the difference between the purpose of usability and user experience evaluation methods", *Workshop User Experience Evaluation Methods in Product Development (UXEM'09)*.

[8] F. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology", *MIS Quarterly* 13(3): 319–340, 1989.

[9] R. Schleicher, S. Trösterer, "The 'Joy-of-Use'-Button: Recording Pleasant Moments While Using a PC", in: *Proc. Human-Computer Interaction – INTERACT 2009*, Vol. 5727/2009, Springer, Heidelberg, 2009.

[10] A. Jameson, "Adaptive Interfaces and Agents", in: *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (A. Sears and J. A. Jacko, eds.), pp. 433–458, CRC Press, Boca Raton FL, 2006.

[11] A. Paramythis, S. Weibelzahl, J. Masthoff, "Layered evaluation of interactive adaptive systems: framework and formative methods", *User Modeling and User-Adapted Interaction* 20(5), 383–453, 2010.

[12] Van Velsen, *The Knowledge Engineering Review*, Cambridge University Press, Vol. 23:3, 261–281, 2008.

[13] S. Möller, J. Krebber, P. Smeele, "Evaluating the Speech Output Component of a Smart-Home System", *Speech Communication* 48:1-27, 2005.