



KRISTINA

A Knowledge-Based Information Agent with Social Competence
and Human Interaction Capabilities

H2020-645012

D2.1

Requirements for Dialogue Management in Adaptive Human-Machine- Communication

Dissemination level:	Restricted
Contractual date of delivery:	Month 6, August 31, 2015
Actual date of delivery:	Month 6, August 30, 2015
Workpackage:	WP2 - Adaptive Dialogue Management
Task:	T2.1
Type:	Report
Approval Status:	Draft
Version:	1.0
Number of pages:	23
Filename:	D2.1-Requirements_for_DM_2015_08_31_v1.0docx

Abstract

This document provides a description of the interaction characteristic in the KRISTINA use cases relevant for the dialogue manager and formulates a list of requirements which have been derived from these characteristics.

The information in this document reflects only the author's views and the European Community is not liable for any use

that may be made of the information contained therein. The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.



co-funded by the European Union



History

Version	Date	Reason	Revised by
0.1	05.08.2015	Initial structure	Stefan Ultes, UULM
0.2	19.08.2015	Draft extension	Louisa Pragst, UULM
0.3	20.08.2015	Draft completion	Stefan Ultes, UULM
0.4	21.08.2015	First complete version	Wolfgang Minker, UULM
0.4	24.08.2015	Internal Review	Elisabeth André, UAU
0.5	28.08.2015	Revised version	Louisa Pragst, UULM
1.0	30.08.2015	Preparation for the submission	Leo Wanner, UPF

Author list

Organization	Name	Contact Information
UULM	Stefan Ultes	Stefan.ultes@uni-ulm.de
UULM	Louisa Pragst	Louisa.pragst@uni-ulm.de
UULM	Wolfgang Minker	Wolfgang.minker@uni-ulm.de
UAU	Elisabeth André	Andre@informatik.uni-augsburg.de
UPF	Leo Wanner	leo.wanner@upf.edu



Executive Summary

At the core of the KRISTINA agent resides the dialogue manager which is responsible for managing the overall interaction. Here, the most important task is to select the next system action. Due to the use cases in KRISTINA, the conversation with the agent encompasses cultural and emotional aspects which should be reflected by the dialogue manager.

Hence, a characterisation of the general KRISTINA interaction in all use cases is presented. Here, only the aspects relevant for dialogue management are considered. The characterisation is based on the literature, the first round of recordings, the general task description, and the identified user requirements. Furthermore, the languages of the use cases are considered showing that especially regarding the cultural and emotional aspects, there may be substantial differences in interaction style.

Based on the identified characteristics, a set of potential requirements will be presented which the dialogue manager may need to match. Again, this includes requirements derived from the emotional and cultural aspects of the interaction. Which of these requirements will need to be implemented eventually depends on the final use case scenarios.



Abbreviations and Acronyms

DM Dialogue management



Table of Contents

1	INTRODUCTION	7
2	CHARACTERISTIC OF THE KRISTINA INTERACTION	8
2.1	Emotion.....	8
2.2	Culture	8
2.2.1	Arabic	11
2.2.2	Turkish.....	12
2.2.3	Polish and German.....	12
2.3	User	13
2.4	Domain	13
3	REQUIREMENTS FOR DIALOGUE MANAGEMENT.....	15
3.1	Emotion.....	15
3.2	Culture	17
3.3	User	18
3.4	Domain	18
3.5	List of Requirements.....	19
4	CONCLUSIONS	20
5	REFERENCES	21
	APPENDIX 1	23



1 INTRODUCTION

The interaction between the users and the system as described by the use cases of the KRISTINA project entails certain challenges which go beyond conventional dialogue system-based communication. There, most systems are built to provide easy access to tasks and services, e.g., train booking, not regarding the user's peculiar interaction style arising from the emotional context or the cultural background. The KRISTINA agent, though, will be designed to handle these phenomena. These cultural and emotional aspects are very important to be taken into account as they play a big role when conversing about sensitive medical topics.

To achieve the goal of rendering the KRISTINA agent culture-aware and emotion-sensitive, we will first have a closer look at the actual interaction styles with respect to the cultural background and the emotional contexts. Furthermore, there are also aspects of the interaction which arise from the general application domain of the KRISTINA agent. Finally, the user requirements identified in D8.2 also need to be taken into account.

Only if we have identified the peculiarities which differentiate the interaction with the KRISTINA agent from other types of interaction, we are able to derive the requirements they put on the dialogue manager.

All of these aspects of the interaction will be described in more detail in the next section. The requirements for the dialogue manager, which may be derived from those aspects, will be described in Section 3.



2 CHARACTERISTIC OF THE KRISTINA INTERACTION

Identifying the characteristics of the interaction style which are peculiar for the KRISTINA agent is necessary for deriving the requirements associated with these characteristic aspects. Based on the literature, impressions from the first recordings, and the general task description, we will characterise the interaction with respect to dialogue management in this section. Here, we will start with the emotional context. Subsequently, we will provide general aspects of cultural influence on the interaction style before relating these aspects to the cultures and languages of the KRISTINA project. Finally, we will continue with more task related characteristics.

2.1 Emotion

The medical domain poses a difficult area of conversation as often sensitive and personal topics need to be discussed. Several studies (e.g. [1] [2] [3]) show that taking into account the emotional state can improve the user acceptance of the system. As characteristics of the interaction regarding emotions, the authors identify system reactions presented in response to the detection of specific emotions in these studies. They were expressing thankfulness, praising, calming, motivating, apologising, and positive/negative wording of the response. For having an emotional agent which is able to not only react to emotions but to show emotions, Butler et al. [4] demonstrated that not showing emotions may prevent social bonding.

To describe emotion, Plutchik [5] defines a set of eight emotions arranged in a circle (emotion wheel). Others rely on the "big six" emotions happiness, anger, disgust, sadness, surprise, and fear [6]. While both previous definitions categorise and label the emotions, the pleasure arousal dominance (PAD) [7] scale aims at creating a three-dimensional emotion space.

However, in the KRISTINA project, only positive vs. negative sentiment or high vs. low arousal may be considered as a starting point.

The display of emotions as well as the appropriate reaction to them is—as many other aspects of communication—highly culturally dependent [8]. Therefore, the next section examines the impact of cultural aspects on dialogue management.

2.2 Culture

It has been observed by many researchers, (e.g. [9]) that different cultures prefer different communication styles. As a person is usually accustomed exclusively to the communication style of their own culture, it is reasonable to assume that talking to members of different cultures may lead to misunderstandings. The high amount of literature regarding business etiquette in foreign countries with the goal of reducing misunderstandings between business partners supports this claim. Similar to the business domain, it is desirable to reduce cultural misunderstandings in the medical domain as a high amount of sensitive information is discussed.



Therefore, unlike most dialogue systems, the KRISTINA agent will be specifically designed to adapt to the culturally dependent communication styles of its users. By adjusting the system's behaviour to the culture of the user, misunderstandings will be reduced and the agent will appear more familiar and therefore more trustworthy to the user. This will render dialogues between the KRISTINA and the user about sensitive medical topics more user-focussed and helpful.

In order to be able to create an adaptive dialogue manager (DM) taking into account specific cultural phenomena, we will first identify certain aspects of communication that are both relevant to dialogue management and influenced by the culture of the speaker(s). Elliott et al. [9] compared communication styles of different cultural groups in the United States regarding, among other aspects, animation/emotion, thought pattern/rhetorical style, directness/indirectness, identity orientation, turn taking/pause time and time. We based our comparison of the cultures considered in the KRISTINA project on these aspects, as they cover a variety of cultural differences relevant for dialogue management. For example, Hofstede's model of culture [10], which is often used as base for the culture-aware agents, mainly covers aspects regarding values and expectations that would influence the semantic content of a chosen system action. Such aspects are included in this list as animation/emotion, identity orientation and time (Hofstede's model of culture [10] is more exhausting in this regard). Additionally, rhetorical differences, which are also important for dialogue management, such as thought pattern/rhetorical style, directness/indirectness and turn taking/pause time are considered. In the following, a short description of each style aspect is presented ordered by their relevance to dialogue management:

Animation/Emotion:

The display of emotions and the apparent involvement in a topic may be perceived very differently across cultures. While in some cultures, showing strong emotions indicates that the user is adamant about their opinion, the same behaviour may be seen as uncontrolled or exaggerated in other cultures. Here, being aware of the culture will help the DM to provide appropriate emotional behaviour, e.g., by interpreting the user's emotions correctly or having the KRISTINA agent itself showing emotions.

Thought Pattern/Rhetorical Style

This aspect refers to the way arguments are presented in a discussion. Kaplan [11] describes five cultural thought patterns, which Elliott et al. [9] reference in their work. These styles are characterised by linearity, parallelism, circularity and digression respectively. While, when using a linear rhetorical style, arguments are presented sequentially and hierarchically sorted, following the parallel style arguments do not have a hierarchy and are presented in a parallel manner, using coordinators rather than subordinators. An example of the linear style would be 'You should drink more. It is hot, therefore your body needs more water', in contrast to the parallel argument 'You should drink more. It is hot and your body needs more water'. The circular style often does not elaborate the main topic, but discusses a multitude of topics, that might have an impact, e.g. 'About 60% of the human body is Water. Water is important for the proper functioning of your body. Finally, in the digressive style not only the main topic, but also surrounding topics are discussed, as in 'You should drink



more. Eating enough is also very important, but when it is so hot, your body needs more water'. Taking this into account will help the DM to provide the necessary information to the user more appropriately so that the user is more likely to accept it.

Directness/Indirectness:

While in some cultures it is favoured and expected to directly express your opinion or your intent, others prefer a more indirect communication style where the listener deduces the intent from the context. As an example, the statement 'Take an aspirin.' could be more indirectly expressed as 'Aspirin can help with headaches'. Taking this into account for the DM results in a suitable presentation of the information provided for the user.

Identity Orientation

People have a set of values along with a certain way they perceive themselves which both influence their decisions. In some cultures, these values and the self-perception tend to be more group oriented: the status of oneself depends on the status of the family and decisions are often made considering the well-being of the group. In contrast to that, there exist more individualistic cultures, in which the status depends on the achievements of the individual person. Decision in such cultures tend to be made considering the own well-being before that of the group. This is important for the DM as by knowing the user's identity orientation appropriate conversation topics, arguments and proposals can be chosen in compliance with the user's values. For example a group oriented user could be motivated by 'You're a big help for your family.', whereas an individualistic user might be more motivated by 'It is impressive how you are able to handle all of this.'

Turn Taking/Pause Time

There are many different ways to signal the conversation partner that it is their turn to speak. Often, pauses are used. There, after some period of silence that may greatly vary in length depending on the culture, the conversation partner may speak. Starting to speak before enough time has passed is often considered as an indication that the speaker does not think before talking. In other cultures, it is perfectly normal to interrupt each other in a conversation. Besides turn taking cues derived from the speech signal, other communication channels may be used as well. Here, some cultures use eye contact as turn-taking indicator, for example by looking at someone when they are supposed to speak. For DM it is important to be able to differentiate between normal turn taking behaviour and impolite interruption by the user, to be able to react appropriately.

Time

Punctuality does not have the same value in every culture. In some cultures, it is considered highly impolite to be late to an appointment. In others, though, the time to meet up is considered to be more like an approximate. There, waiting for the 'right' time to meet is deemed proper and insistence on punctuality would be rude behaviour. Reflecting this in the DM is important as reminding the user of appointments may be part of the medical domain and should be done in a polite way. Also, users should be made aware of conventions different from their own, for example by telling them 'In Germany, it is considered rude to be late for an appointment.'



Considering these aspects, we have analysed which role these cultural communication aspects play for cultures relevant to the KRISTINA project. In order to better grasp the differences between the cultures, we have designed a template (A.1) and asked people, who have witnessed the recordings of the use case partners, to complete it based on their impressions. Based on literature and this template, we have identified the following aspects to be important for the type of communication necessary for the KRISTINA use cases.

2.2.1 Arabic

Animation/Emotion According to Zaharna [12], in Arab communication, evoking an emotional response in the listener is important. This could be achieved by an animated and emotional presentation of what is said.

In the recordings we could observe a stronger use of gestures in male participants than in female participants. Taking into account the model of culture presented by Hofstede [10], this gender specific difference in behaviour could be a result of different expectations regarding men and women, indicating a masculine culture.

Thought pattern/Rhetorical style Kaplan [11] characterises the Arab rhetorical style as parallel. Furthermore, Feghali [13] notes that expressive and verbose language as well as rhetorical patterns such as exaggeration, assertion and stylised repetition are being used.

The recordings did not reveal any noticeable insights regarding the rhetorical style.

Directness/Indirectness The communication style is described as being indirect [13]. Often the intent of the speaker is concealed and needs to be deciphered by the listener.

The recordings did not reveal any noticeable insights regarding the directness of statements.

Identity Orientation According to Feghali [13], the Arab culture is group orientated. Additional values are hospitality and honour. We assume these values to influence decision making. Thus, those should be considered when designing the dialogue and choosing a dialogue strategy.

The recordings did not reveal any noticeable insights regarding the identity orientation.

Turn Taking/Pause Time The recordings show more interruptions from participants with Arab background – especially males – than from participants with other backgrounds.

Time Feghali [13] characterises the predominant approach to time as polychronic, with more emphasis on building relationships than adhering to a schedule.

The recordings did not reveal any noticeable insights regarding the approach to time.

While Feghali [13] gives a very detailed summary of Arab communication patterns, issues with the findings are pointed out as well. One of the main issues of the existing research is the definition of Arab culture and that there is no such thing as ‘the’ Arab culture. The KRISTINA project focuses on North African Arab culture, the findings of the literature need to be confirmed for this specific part of the Arab culture.



2.2.2 Turkish

For this section, information found in the business guides “Cross Cultural Communication, Know about Turkey” and “Cultural Information – Turkey” by the Centre for Intercultural Learning, Canada, was included.

Animation/Emotion According to the aforementioned business guides, Open display of emotions is not usual in the Turkish culture. Especially the facial expression remains neutral, while the voice shows a broader range of emotionality. Gestures are used frequently. (e.g. [14] [15])

The use of gestures observed in the recordings varies on an individual level.

Thought pattern/Rhetorical style Enginarlar [16] reports an elaborative language use with frequent metaphors, idioms, clichés, set phrases, or proverbs in the Turkish culture. It seems to share common aspects with the Arab culture in this respect.

In the recordings we observed a frequent use of filler words, such as *şey* and *yani*.

Directness/Indirectness The Turkish communication style is characterised as indirect in business guides (e.g. [15]).

The recordings did not reveal any noticeable insights regarding the directness of statements.

Identity Orientation The Turkish culture is often described as group oriented, placing a high value on the family. (e.g. [14] [15])

The recordings did not reveal any noticeable insights regarding the identity orientation.

Turn Taking/Pause Time In the recordings we observe that holding eye contact is common for the speaker as well as the listener. As turn taking indicators gaze as well as pauses are used. Interruptions are uncommon.

Time Business guides describe Turkish culture as polychronic. Nevertheless, punctuality is valued in business environments. (e.g. [14] [15])

The recordings did not reveal any noticeable insights regarding the approach to time.

2.2.3 Polish and German

During the recordings, German and Polish communication patterns seemed to be similar. Also, they appeared to conform to the European American communication patterns described by Elliott et al. [9], which seems reasonable given the roots of this culture. Therefore, both cultures are discussed in this section, with differentiations given where necessary. Furthermore, the findings of Elliott et al. [9] regarding the European American culture are included.

Animation/Emotion In the European American culture the display of Emotions tends to be restricted, especially in public [9].

As Polish participants mentioned that Germans seem to be cold, it is possible that in Polish culture the display of emotions is expected to be more pronounced than in German culture. This assumption could be taken into consideration and evaluated in future recordings.



Thought pattern/Rhetorical style Elliott et al. [9] characterises the European American communication style as linear, concise and logical.

The recordings did not reveal any noticeable insights regarding the rhetorical style.

Directness/Indirectness The communication style is described as being direct and concise [9].

The recordings did not reveal any noticeable insights regarding the directness of statements.

Identity Orientation Elliott et al. [9] state, that the European American culture is individualistic.

The remark by Polish participants that German people seem to be cold might indicate that the German culture is more individualistically oriented than the Polish culture. As individualistic cultures do not place as high a value on their families as group oriented cultures, individualistic behaviour could be interpreted as lack of compassion and care by a person with group oriented values. In future recordings this assumption could be reviewed.

Turn Taking/Pause Time In European American communication, turn taking is signalled by the speaker looking in the eye of the listener and ceasing to speak, with only brief pauses [9].

The recordings show only little interruptions from participants with a Polish background.

Time Elliott et al. [9] note, that the European American approach to time is monochronic. Punctuality is important and failure to arrive in time can be considered rude.

The recordings did not reveal any noticeable insights regarding the approach to time.

2.3 User

Deliverable D8.2 specifies the initial user requirements identified by the use case partners. Regarding the dialogue management the following user requirements were identified:

The KRISTINA agent is supposed to communicate with the user via speech, visually supported by a human avatar on a screen. Depending on user preferences or limitations, the communicated information may be presented in different forms.

Furthermore, the system should be able to accompany the user in different moods. For this it is necessary to correctly identify different emotions and react in an appropriate way. The system's conversation style as well as its mood should be adaptive to the specific context and type of user.

In order to identify misunderstandings, incoherent dialogue moves need to be detected and the system should react appropriately.

Finally, a warning to the monitoring personnel should be issued in case the system encounters difficulties it cannot handle.

2.4 Domain

Based on the use cases of KRISTINA, several aspects of the domain may be observed which are relevant for dialogue management. In the KRISTINA use cases, different tasks and their



interaction domains are described: getting general medical information from the system, getting information about the health care system, getting personalised medical information, and getting personal information about the patient.

Although no medical diagnosis or treatment recommendations will be part of the KRISTINA interaction, the possible conversations still cover the wide range of the medical domain. Moreover, the patient's own medical history and medical conditions add to general information creating an even wider domain resulting in a mixture of personalised and general knowledge. The personalised knowledge—medical as well as other personal information—may be acquired over time through interaction. Furthermore, both types of knowledge may be addressed for information retrieval tasks. Hence, it is not known during design-time which information is necessary to fulfil the queries. Even more so, the system should be able to present individualised responses.

Finally, to increase the user's trust into the KRISTINA agent, the agent may chat with the patient or relatives about the patient's respective personal background and history.



3 REQUIREMENTS FOR DIALOGUE MANAGEMENT

Given the characteristic of the KRISTINA interaction, several requirements for the dialogue management may be deduced. Those may be grouped into “emotion”, “culture”, “user”, and “domain” requirements and will be presented in this section.

3.1 Emotion

Regarding emotions, the KRISTINA agent faces two major challenges: reacting to the user’s emotions in an appropriate way and expressing suitable emotions itself. This section examines the requirements the DM needs to fulfil in order to handle those challenges.

A computational model of emotion for both the user and the system should be utilised. For the user model, approaches as presented by Pittermann et al. [17] [18], André et al. [19] and Gnjatović et al. [20], that modify the system behaviour according to the user’s emotional state, could be suitable. The first two use a dimensional emotion model, the last one discrete emotion labels. In contrast, systems that display emotions, and possibly even adapt their behaviour to the system emotion, often are based on the appraisal model of emotion. Examples for such systems are presented by Gratch et al. [21] and Dias et al. [22]. In order to create computational models suited for the KRISTINA project, we will evaluate the usability of the given examples. It will be especially important to make sure that both models can work together and with the cultural computational model. An example where emotion and culture are modelled in one system can be found in [23].

The dialogue strategy needs to be adaptable to the emotions of the user in order to create appropriate reactions of the system. Therefore, the dialogue manager needs to provide mechanisms which allow for emotion-aware dialogue strategy rendering. Examples of emotion-adaptive reactions of the dialogue manager are listed in the following to give a better impression of the kind of adaptation needed. Of course, the goal within KRISTINA is not to provide strict rules for adaptation. Hence, these examples may only be seen as a starting point. Consequently, for better results, the system may be equipped with the ability to learn a policy automatically to reflect actual human reactions to emotions.

Joy

If a user displays joy, the system may also show signs of happiness thus mirroring and intensifying the user’s emotions. Furthermore, the user being joyful also indicates that the topic of the discussion is pleasant for the user which may be taken into account for selecting the next system action. Also, the system may ask further questions concerning the source of the feeling to get a better understanding of the underlying mechanism:

User: I could get my father to drink more water.

Kristina: I am happy to hear that. How did you do that?

Sadness

In the case of sadness, the system may again mirror the user’s feeling by displaying mild sadness. This sign of compassion may improve the user-system relationship. Furthermore, the system may try to console the user and to offer a different perspective. Rhetorically,



indirect statements seem less aggressive and may therefore be preferred. Repetition may have a calming effect.

User: I don't think the medicine is helping. It isn't getting any better.

Kristina: It is not that bad. This medicine usually takes a few days to work, it is too early to worry. And if it does not work we can still try ...

Insecurity

If the user feels insecure, the system may remind him of his achievements. Explanations may be broken down to easily understandable pieces and different verbalizations could be tried. Direct statements leave little room for interpretation and therefore offer a certainty that can support the user.

User: I'm not sure I will be able to get a wheelchair for Mr. Smith. I don't know how to proceed.

Kristina: Don't worry. You already got him a walker, it is almost the same. First you have to ...

Shame

If the user is ashamed, the system may ensure the user that everything is normal and may remind him that other people experienced the same. This could be achieved by providing appropriate statistics.

User: I don't want to use condoms. What will the people think when I buy them?

Kristina: It is not unusual to buy condoms. 87% of the people use condoms when in a new relationship. And it offers many benefits.

Anger

An angry user should be calmed down. If the user is angry because of the system, the reason for that feeling should be discovered. The system may apologise if it is the reason, provide explanations for its behaviour and proposes to act differently in the future. If the user is angry because of something else, the system may listen to the complaints, offer a different perspective or advice if possible.

User: I am not talking to that guy again.

System: What happened? Why are you so upset?

User: He never lets me finish my sentences. How can he be so rude?

System: I see why this is upsetting. But you should know that he is not doing that to be rude. In his culture it is perfectly acceptable to interrupt each other while talking.

In addition to reacting appropriately to the user's emotions, the KRISTINA agent needs to express emotions itself in order to appear as a relatable and natural dialogue partner. This requirement is further emphasised by the fact that some cultures expect a pronounced display of emotions. In order to properly express suitable emotions, the dialogue manager



will be extended by an emotion generation module and a corresponding policy for emotion generation will be trained. This emotions generation module will be responsible for enhancing the system actions with an emotion that is suitable considering the cultural background and the emotional state, of the user as well as the overall situation and the chosen system action. The appropriate representation of the emotion, e.g. by facial expression, gestures or verbalisation, is not part of the dialogue management and will be handled by the subsequent modules.

3.2 Culture

Regarding culture, some computational models for agent behaviour have been presented in literature, e.g. [24], [25] and [26]. Often Hofstede's model of culture [10] is utilised for such computational model. While some aspects can be adopted for the computational model of culture used in KRISTINA, especially regarding differences in values, others have to be introduced in order to address all cultural aspects of dialogue management, as none of these models considers rhetorical differences. In order to enhance the existing computational models of culture to meet the needs of dialogue management, we can utilise the insights we gained in the previous section. There, we have identified aspects of communication that are dependent on culture. Those aspects can be influenced by the dialogue management in the following ways:

Animation/Emotion:

For adaptive behaviour regarding the display of emotions, the DM may choose different system moves. While, in cultures with a restrained communication style, the dialogue manager will mostly stick to purely 'functional' dialogue moves, more expressive communication styles require dialogue moves expressing emotion, if applicable even without any functional or semantic content.

Thought Pattern/Rhetorical style

The rhetorical style that is to be implemented has a great impact on the strategy pursued by the dialogue manager. An argument can be built during several dialogue moves and needs to be consistent and to follow one strategy in order to be convincing.

Directness/Indirectness:

Depending on the specific architecture adaption of directness might be achieved either by different dialogue moves chosen by the dialogue manager or by speech generation expressing the same intent in different ways.

Identity Orientation:

Knowledge about the identity orientation can be used by the dialogue management in order to determine what to propose and which arguments to use to convince the dialogue partner.

Turn Taking/Pause Time

Turn taking is relevant to the dialogue manager insofar as interruptions are concerned. Dialogue management should be able to correctly interpret interruptions and handle them in a suitable way. Directing gaze or maintaining pauses is not part of dialogue management.



Keeping pauses for the right amount of time might be a specifically hard challenge for the real-time ability of any spoken dialogue system.

Time

The importance of punctuality might be considered when reminding users of appointments. Suitable system moves for every culture may be implemented and chosen accordingly.

3.3 User

Some of the user requirements established in D8.2 have already been addressed in the previous sections. Adaptivity to the emotional state and cultural background enables the KRISTINA agent to accompany the user in different moods as well as different types of users and contexts.

The user requirements state, that visual output in form of an avatar is to be used for KRISTINA. Dialogue management itself is modality independent. However, in a turn-based dialogue, idle or listening behaviour may easily be modelled if only speech is used as communication channel: the system is simply silent. For visual output, though, the avatar needs more complex behaviour capabilities. Simply showing a frozen avatar is not appropriate. Hence, continuous idle and listening behaviour of the visual representation of the system needs to be generated.

Another user requirement is the identification of incoherent user moves with the goal of identifying misunderstandings. In order to achieve this, a list of possible next user moves should be available in the knowledge base. When encountering a user move that is unexpected, maintaining multiple dialogue states synchronously may enable easy recovery by offering alternatives of what has been said.

Finally, a warning to the monitoring personnel should be issued in case the system encounters difficulties it cannot handle.

3.4 Domain

As the interaction of the KRISTINA use cases allow for dialogues in the wide medical domain, the dialogue manager should be able to handle this wide domain as well. In fact, these domain aspects have major implications on the requirements. To cover the wide medical domain, an external knowledge based is needed and as the domain is quite complex, the knowledge base should be organised as an ontology.

Furthermore, means of adding personal information about the user should be possible, e.g., using a separate user model. This information needs also to be mapped to ontological concepts but should be stored separately. As the user information may change over time, the knowledge base needs to be dynamically extensible.

To handle inquiries by the KRISTINA users, the topic of the inquiry needs to be identified. As the domain size is very big, reasoning techniques should be used to identify the user's intent and the information which is needed by the system to fulfil the query.



Finally, for allowing conversational dialogue, i.e., not aimed at querying information from the system, a conversational module is required which needs to take account for user characteristics, e.g., gender, age, culture, mood.

3.5 List of Requirements

The above presented requirements for dialogue management in all four categories are summarised within the following list. Hence, the dialogue manager is required

- to provide mechanisms for adapting the dialogue strategy to the user's emotion with regard to their cultural background,
- to contain an emotion generation module which enhances the system actions with emotional context,
- to provide a set of system actions which reflect all culture-dependent actions which will reflect the necessary aspects of cultural communication,
- to maintain a dialogue state which enables the DM to select appropriate culture-dependent system actions,
- to be connected to an external knowledge base which will enable the DM to handle the complex domain, and
- to offer mechanisms which will enable idle or listening behaviour.



4 CONCLUSIONS

The interaction between the users and the system defined by the KRISTINA use cases have certain characteristics which render them as being different compared to conventional task-oriented human-machine communication. Due to the sensitive medical subject of the conversation and the user group containing members of different cultures, we have analysed the interaction to identify cultural and emotional aspects which may have a potential influence on the KRISTINA interaction and thus may be taken account of within the dialogue management.

For this analysis, we have not only relied on the literature but have also taken into account audio recordings of sample dialogues. However, oftentimes, the recordings could not reveal any noticeable insights. Hence, to some extent, we are only able to define a set of potential aspects of the interaction which may be relevant for KRISTINA. Here, cultural-dependent usage of emotions as well as the rhetorical style have been found to be of high importance.

Based on the analysis and the set of potential interaction aspects, we have identified several potential requirements for the dialogue manager. Most prominently, the dialogue manager of the KRISTINA agent will need mechanisms which are able to handle the cultural differences, e.g., by generating emotional system acts for cultures which require a high degree of emotional interaction. In general, the dialogue manager should be aware of the culture to adapt the strategy accordingly.

As emotion generation itself presents another important requirement for dialogue management, the system needs to be equipped with an emotion generation module.

Finally, not only cultural and emotional requirements have been identified: also the domain scope is relevant for the dialogue manager. For the KRISTINA interactions, the very huge medical domain will be covered. As this is hard to be modelled within a dialogue manager directly, all domain-related actions may be delegated to an external knowledge base.



5 REFERENCES

- [1] G. Bertrand, F. Nothdurft, W. Minker, H. Traue and S. Walter, "Adapting dialogue to user emotion-a wizard-of-oz study for adaptation strategies," in *Proceedings of the Paralinguistic Information and its Integration in Spoken Dialogue Systems Workshop*, Springer, 2011, pp. 285-294.
- [2] N. Jaksic, P. Branco, P. Stephenson and L. M. Encarnação, "The effectiveness of social agents in reducing user frustration," in *CHI'06 extended abstracts on Human factors in computing systems*, ACM, 2006, pp. 917-922.
- [3] T. Partala and V. Surakka, "The effects of affective interventions in human-computer interaction," *Interacting with computers*, vol. 16, no. 2, pp. 295-309, 2004.
- [4] E. A. Butler, B. Egloff, F. H. Wilhelm, N. C. Smith, E. A. Erickson and J. J. Gross, "The social consequences of expressive suppression," *Emotion*, vol. 3, no. 1, p. 48, 2003.
- [5] R. Plutchik, *Emotion: A Psychoevolutionary Synthesis*, Harper & Row, 1980.
- [6] R. Cornelius, *The science of emotion : research and tradition in the psychology of emotions*, London: Prentice Hall, 1996.
- [7] A. Mehrabian, „Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament," *Current Psychology*, pp. 261-292, December 1996.
- [8] E. André, „Preparing emotional agents for intercultural communication," *The Oxford Handbook of Affective Computing*, pp. 309-320, 2014.
- [9] C. Elliott, R. J. Adams and S. Sockalingam, "Multicultural Toolkit. (Toolkit for cross-cultural collaboration)," *Awesome library-US office of Minority Affairs, USA*, 1999.
- [10] G. Hofstede, *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*, Thousand Oaks, CA: Sage Publications, 2001.
- [11] R. B. Kaplan, "Cultural thought patterns in inter-cultural education," *Language learning*, vol. 16, no. 1-2, pp. 1-20, 1966.
- [12] R. S. Zaharna, „Understanding cultural preferences of Arab communication patterns," *Public Relations Review*, Bd. 21, Nr. 3, pp. 241-255, 1995.
- [13] E. Feghali, "Arab cultural communication patterns," *International Journal of Intercultural Relations*, vol. 21, no. 3, pp. 345-378, 1997.
- [14] "Cultural Information - Turkey | Centre for Intercultural Learning," [Online]. Available: <http://www.intercultures.ca/cil-cai/ci-ic-eng.asp?iso=tr>. [Accessed 19 August 2015].
- [15] "Cross Cultural Communication , Know about turkey," [Online]. Available:



<http://de.slideshare.net/Lalitsharma81/cross-cultural-communication-know-about-turkey>. [Accessed 19 August 2015].

- [16] H. Enginarlar, "A contrastive analysis of writing in Turkish and English of Turkish high school students," *Unpublished Ph. D. dissertation. Ankara, Turkey: Hacettepe University*, 1990.
- [17] J. Pittermann und A. Pittermann, „A data-oriented approach to integrate emotions in adaptive dialogue management,“ in *Proceedings of the 12th international conference on Intelligent user interfaces*, ACM, 2007, pp. 270-273.
- [18] J. a. P. A. Pittermann, „Integrating emotion recognition into an adaptive spoken language dialogue system,“ *IE 06. 2nd IET International Conference on Intelligent Environments*, Bd. 1, pp. 197-202, 2006.
- [19] E. André, M. Rehm, W. Minker und D. Bühler, „Endowing spoken language dialogue systems with emotional intelligence,“ in *Affective Dialogue Systems*, Berlin Heidelberg, Springer, 2004, pp. 178-187.
- [20] M. Gnjatović und D. Rösner, „Adaptive dialogue management in the nimitex prototype system,“ in *Perception in Multimodal Dialogue Systems*, Berlin Heidelberg, Springer, 2008, pp. 14-25.
- [21] J. Gratch und S. Marsella, „A domain-independent framework for modeling emotion,“ *Cognitive Systems Research*, Bd. 5, Nr. 4, pp. 269-306, 2004.
- [22] J. Dias und A. Paiva, „Feeling and reasoning: A computational model for emotional characters,“ in *Progress in artificial intelligence*, Berlin Heidelberg, Springer, 2005, pp. 127-140.
- [23] R. Aylett und A. Paiva, „Computational Modelling of Culture and Affect,“ *Emotion Review*, Bd. 4, Nr. 3, pp. 253-263, 2012.
- [24] D. Jan, D. Herrera, B. Martinovski, D. Novick und D. Traum, „A computational model of culture-specific conversational behavior,“ in *Intelligent virtual agents*, Berlin Heidelberg, Springer, 2007, pp. 45-56.
- [25] S. Mascarenhas, R. Prada, A. Paiva und G. J. Hofstede, „Social importance dynamics: A model for culturally-adaptive agents,“ in *Intelligent virtual agents*, Berlin Heidelberg, Springer, 2013, pp. 325-338.
- [26] R. Aylett und A. Paiva, „Computational modelling of culture and affect,“ *Emotion Review*, Bd. 4, Nr. 3, pp. 253-263, 2012.
- [27] M. B. Arnold, „Emotion and Personality,“ *Psychological Aspects*, Bd. 1, pp. 37-54, 1960.



APPENDIX 1

A.1 Template

Language

Try to detect general trends across a culture. Do not assess individual dialogues.

In what way do participants signal that it is their partner's turn?

Key words

Nonverbal signals (eg. Gaze)

Pause

Do interruptions occur frequently? (form 1 - 'never' to 5 - 'permanently')

Where is the attention of the participants while talking? (their partner, surroundings, themselves)

Where is the attention of the participants while listening? (their partner, surroundings, themselves)

In what way is understanding signaled? (e.g. nodding, "I see", ...)

Which partner tends to be more active? (User, KRISTINA)

Which partner tends to ask more questions? (User, KRISTINA)

Which partner tends to talk more? (User, KRISTINA)

Do participants convey their intention directly (e.g. "Take aspirin") or indirectly (e.g. "Aspirin is often used to cure headaches")?

Do participants express their intention verbosely or concisely? (from 1 - 'verbosely' to 5 - 'concisely')

How frequently do the participants use filler words? (from 1 - 'never' to 5 - 'permanently')

Which filler words are being used?

Do participants use logical arguments to persuade their partner of an idea (e.g. to take medicine)?

Do participants focus on good presentation of an idea when trying to persuade their partner?

How frequently do the participants use gestures? (from 1 - 'never' to 5 - 'permanently')

Which gestures are being used?

Is there anything else you noticed?