Proceedings Paper:
Thakker, D, Lau, L, Denaux, R et al. (3 more authors) (2014) Using DBpedia as a knowledge source for culture-related user modelling questionnaires. In: Dimitrova, V, Kuflik, T, Chin, D, Ricci, F, Dolog, P and Houben, G-J, (eds.) User Modeling, Adaptation, and Personalization 22nd International Conference, UMAP 2014, Proceedings. 22nd International Conference, UMAP 2014, 07-11 Jul 2014, Aalborg, Denmark. Lecture Notes in Computer Science . Springer , 207 - 218 (10). ISBN 978-3-319-08785-6

https://doi.org/10.1007/978-3-319-08786-3_18

Using DBpedia as a Knowledge Source for Culture-related User Modelling Questionnaires

Dhavalkumar Thakker[1], Lydia Lau[1], Ronald Denaux[2], Vania Dimitrova[1], Paul Brna[1], Christina Steiner[3]

1 School of Computing, University of Leeds, United Kingdom.
2 iSOCO, Madrid, Spain.
3 Knowledge Technologies Institute, Graz University of Technology, Austria.
{d.thakker, l.m.s.lau, v.g.dimitrova}@leeds.ac.uk; rdenaux@gmail.com; paulbrna@mac.com; christina.steiner@tugraz.at

Abstract. In the culture domain, questionnaires are often used to obtain profiles of users for adaptation. Creating questionnaires requires subject matter experts and diverse content, and often does not scale to a variety of cultures and situations. This paper presents a novel approach that is inspired by crowdwisdom and takes advantage of freely available structured linked data. It presents a mechanism for extracting culturally-related facts from DBpedia, utilised as a knowledge source in an interactive user modelling system. A user study, which examines the system usability and the accuracy of the resulting user model, demonstrates the potential of using DBpedia for generating culture-related user modelling questionnaires and points at issues for further investigation.

Keywords: Culture-related user model, linked data, questionnaire generation

1 Introduction

Today's globalising world requires a new set of skills and competences, among which culture takes a prominent role. Subsequently, a new breed of culturally-aware intelligent learning environments that address challenges when accommodating culture have emerged[1]. The application of the work presented here is set within the framework of the European project ImREAL[2] which considered user-adaptive situational simulations for interpersonal communication with cultural variations. Such simulation environments aim at developing intercultural competences and provide user-adaptive virtual learning experience by taking into account the learner's knowledge of other cultures. The example use cases range from medical interviews, business events (first meeting, business dinner) and the buddying of international students (meeting upon arrival, attending social events). Across the ImREAL use cases, dealing with cultural variations was an important common theme. Culture by nationality (country) was

[1] http://cats-ws.org/previous-cats/
[2] http://www.imreal-project.eu/

adfa, p. 1, 2011.

chosen as the prime focus, following findings in business and management indicating that nationality and countries are reliable indicators for tackling cultural diversity [1].

The key challenge for user-adaptive cultural simulations is to derive a model of a user's knowledge of cultural dimensions relevant to the simulated situations; this is the well-known cold start problem. In the culture domain, questionnaires are often used to obtain profiles of users. This relies on availability of subject matter experts and creation of diverse content including cultural dimensions relevant to the application context [15]. A major challenge is scaling up questionnaire-based user modelling to address cultural diversity and to include engaging examples [2]. Furthermore, a flexible and extendable way of creating and utilising knowledge sources is needed.

To address this challenge a novel approach is proposed here inspired by crowdwisdom and taking advantage of freely available structured linked data. The paper presents an interactive way of deriving a model of a user's knowledge of selected cultural aspects by utilising semantic datasets from Linked Data[3] (in this case DBpedia [3]) to serve as the knowledge base for culture-related facts. The approach provides ontology-based knowledge probing, implemented as an interactive agent called Perico, which builds an overlay user model (UM) of knowledge on selected aspects related to culture by nationality. In the context of user-adaptive systems, Perico can provide an engaging way to derive an initial UM prior interacting with the system, or can be invoked within the system to extend/verify the existing user model.

Perico[4] was presented elsewhere [8], together with an initial validation in a CrowdFlower[5] study which indicated that the interaction was fairly intuitive, but did not give in-depth knowledge of the challenges faced while interacting with Percio (very little qualitative data was provided by the users). A study with two experts inspecting the performance of the system pointed at possible issues with the user model accuracy and utility of DBpedia facts. The findings lacked quantitative backing and were missing the perspective of a real user. In this paper, a controlled user study is reported involving representative users of Perico - adults who wish to extend their knowledge on certain cultural aspects that they may need in everyday intercultural encounters, e.g. visit to a country for business or tourism. Adding to [8], this paper specifically focuses on the DBpedia knowledge extraction mechanism, providing detail of its implementation and utilisation for knowledge probing in user modelling.

The key contribution to user-adaptive systems is a novel, flexible and extendable way to construct culture-related user modelling questionnaires from DBpedia which is validated in a user study. Section 2 outlines how DBpedia has been used as a knowledge source for user modeling. Section 3 presents the user study, and the results are discussed in Section 4. We conclude by positioning in relevant literature (Section 5) and drawing lessons learnt for culture-related UM (Section 6).

[3] http://linkeddata.org/

[4] Perico is available online from http://imash.leeds.ac.uk:8080/perico/

[5] http://crowdflower.com/

2 Using DBpedia as a Knowledge Pool for User Modelling

In order to probe a user's knowledge in a domain, a user modelling system requires access to a knowledge base with domain facts. In the case of culture, key requirements for selecting the knowledge source include diversity and intuitiveness. The knowledge base must contain facts about a wide variety of cultural groups to increase the chance that it contains facts which are relevant to the user's own cultural group and to other cultural groups. Having a range of examples and authentic terms used in the specific cultural settings can increase the user's engagement with the question-like assessment format [2]. To meet these requirements, one of the largest multi-domain semantic dataset that currently exists, DBpedia [3], is chosen as the knowledge source. DBpedia is a community effort to extract structured information from Wikipedia and to make this information available on the Web. Over the last year, DBpedia has become a central interlinking hub for the emerging Web of Data [4]. DBpedia contains lots of instances, represents real community agreement and automatically evolves as Wikipedia changes [4].Extracting domain-related facts from DBpedia requires a set of seed topics and a strategy on how to extract relevant assertions as presented in next two sections.

Selection of Topics. The specific application domain in our case is cultural variations in interpersonal communication (the application focus of the ImREAL project). The relevant concepts in this domain were defined in an Activity Model Ontology (AMOn) underpinned by Activity Theory [5], including concepts like: `Subject`, `Object`, `Tools`, `Motivation`, `Outcome`, `Community`, etc. For example, interpersonal communication `Tools` are expanded to include `Mental Tools` (e.g. `Verbal Communication`, `Nonverbal Communication` and `Body Language`) and Physical Tools (e.g. `Clothing`). AMOn was further extended to AMOn+ by indicating the key interpersonal communication concepts that can possibly have cultural variations [6]. Both AMOn and AMOn+ are presented in earlier publications [5,6]. While these two ontologies provide structure for the important domain aspects (i.e. the abstract facts), the broad range of instantiations were missing (e.g. the variety of gestures, different clothing items or cuisine in different countries). DBpedia is used as a source for such instantiations. A set of seed concepts from AMOn+ is selected for extracting cultural-related facts from DBpedia (see Figure 1).

Seven domain topics, grouped into two categories, were selected as entry points for extracting cultural-related facts. The first category includes three topics used to extract cultural facts related to the ImREAL use cases: `gestures` – a prominent element in non-verbal communication, `clothing` – a key element in dressing norms, linked to social and cultural conventions, and `food` - specifically related to interpersonal communication in informal settings. Socio-political facts about a country give useful knowledge in interpersonal communication situations; the following were selected: `language`, `currency`, `human development index (HDI)`, and `generalised inequality index (GNI)`.

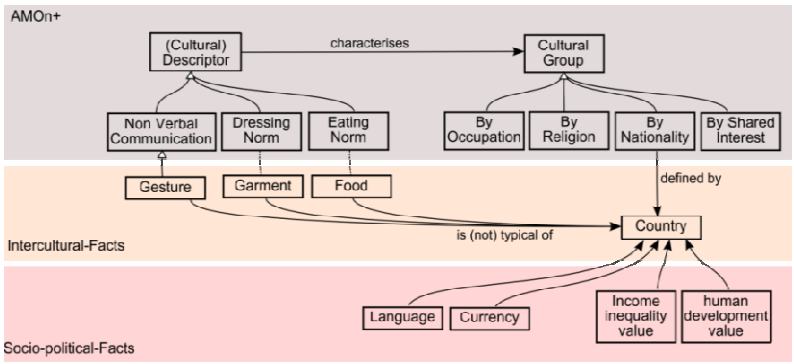

Fig. 1. Selected cultural-related topics following AMOn+. The top section shows a segment of AMOn+ that relates `cultural descriptors` relate to `cultural groups`. The middle section shows the types of intercultural facts extracted from DBpedia and their relation to AMOn+ concepts. The bottom section shows additional socio-political facts about countries.

DBpedia Facts Extraction Strategy. Knowledge pool of facts related to the selected topics is extracted from DBpedia by: (i) identifying DBPedia categories that relate to the selected concepts (for example, for the topic `Clothing`, the matching category is `category:Clothing`); (ii) traversing the DBpedia category network to find narrower pages (using `skos:narrower`) for the identified categories, i.e. searching for pages with a specific category as well as subcategories (for example, for the `category:Clothing`, `dbpedia:Loden_cape` is one of the narrower pages); (iii) traversing the DBpedia category network to find broader categories that are shared between the page to be extracted and the country linked to them; for example, traversing the category network for broader categories of `dbpedia:Loden_cape`, the categories `German_Culture` and `Austrian_Culture` and their respective super categories `Germany` and `Austria`, connected to `dbpedia:Germany` and `dbpedia:Austria` respectively, instances of a `Country`; (iv) inferring/adding new OWL axioms (basis statements an OWL ontology expresses)[6], such as: a class assertion axiom linking the DBpedia page with an OWL class that is (relevant to) a concept from the Cultural Variations module; an object property assertions linking the DBpedia page with one or more countries where this Cultural Variations concept occurs; and copying relevant literal data, such as labels and depictions. For example, from the extracted facts, assertions "`Loden_cape is a Clothing`" and "`Loden_cape occursIn Austria`" and "`Loden_cape occursIn Germany`" are inferred.

The resulting knowledge pool, available from the AMOn+ website[7], includes around 40K facts (OWL logical axioms) about 270 countries, 565 items of clothing, 4282 items of food, 88 gestures, 159 currencies and 288 languages. The ontology contains some 20K facts containing human-readable labels and depictions. The

[6] http://www.w3.org/TR/2009/WD-owl2-primer-20090421/
[7] http://imash.leeds.ac.uk/ontologies/amon/

DBpedia knowledge pool is used as the knowledge source for Perico's knowledge probing, output generation and user input interpretation which constructs the UM.

Knowledge Probing. A knowledge probing strategy is developed to select assertions from the knowledge pool and convert them into questions to be posed to the learner. Knowledge probing strategy takes a tuple <*P*, *Fi*, *G*, *T*, *Fo*, *A*> as an input, and returns an OWL axiom. The input includes: *P* - a pool of facts represented as a set of OWL axioms; *Fi* - a set of focus items, i.e. OWL entities that the selected axioms must contain (in Perico, these are `dbpedia:Country` individuals); *G* - a goal condition determining whether the strategy should keep selecting more axioms (in Perico, a goal is defined as a configurable number of facts that will be probed for each focus item); *T* - a function that assigns a topic to each OWL axiom in *P* using a selected set of topics, OWL entities, that specify the scope of the dialogue (in Perico, the topics include `gestures`, `food`, `clothing`, `language`, `currency`, `HDI`, and `GNI`); *Fo* - a function that assigns an axiom form to each axiom (currently, Perico includes two axiom forms - normal assertion, facts inferred from DBpedia, and negations, generated from inferred facts); *A* - a set of already probed axioms (Perico uses the list of probed axioms to avoid repetition of the facts the user is presented with).

The knowledge probing process ends either when the goal *G* has been met for all focus items, or when the fact pool does not provide enough axioms to meet the goal. When the knowledge probing mechanism returns a selected axiom, this is used as a basis for generating a *knowledge probing dialogue game*. Sentence openers are added to the informative assertions in order to indicate the communicative function of propositional-test-questions, in the form: <Sentence opener> <axiom rendering> <?>. Example sentence openers are: "*Is it true that*", "*Do you think that*", "*Is it likely that*", "*Did you find that*" or "*Did you experience that*".

User Profile Creation. The user's answers to the probing questions are used as evidence of his/her knowledge of the relevant cultural aspect about the country (focus item). Perico suggests pre-defined answers to the knowledge probing questions, such as: agreement, disagreement, inform-ignorance and inform-incorrect-question. Input interpretation includes recognising these pre-defined answers to the knowledge probing questions and annotating each answer with the appropriate discourse-related annotations. Once the answer is interpreted, the UM is changed accordingly. In particular, the UM contains: (i) scores for each of selected topics, plus an explanation containing the probed OWL axioms (e.g. *The user correctly disagreed with the assertion 'Moutza occurs in Spain'.*); (ii) an aggregated score for each of the focus items (countries which have been discussed), calculated as the average score for all the answers related to the country; and (iii) an overall score based on all probed countries. At the end of the dialogue, the aggregated scores and the overall score are presented to the user in a dialogue conclusion game.

3 User Study

A user study was conducted to address the following research questions:

RQ1: Is Perico usable and intuitive for the intended users; and what are the possible limitations of the interaction with Perico?

RQ2: Is the user model produced by Perico accurate against the user's perception of his/her knowledge in the selected cultural aspects?

Participants. The intended users of Perico are adults who can have everyday intercultural encounters, e.g. visit to a country for business or tourism; this relates to the ImREAL use cases - cultural encounters in interpersonal communication (Section 1). 22 participants (age 18-50, mean=28), living in the UK, were recruited on voluntary basis varying in their cultural exposure – British (11), Bulgarian(3), German(1), Greek(1), Indian(1), Jordanian(1), Malaysian(1), Maltese(1), Nepalese(1) and Polish(1). The cultural exposure of the participants was examined based on the 10 country cultural clusters developed in the GLOBE project [1]. The participants were asked to state their familiarity with the countries in each cluster as (i) *none* (no encounter with the national culture); (ii) *low* (short visits to the country, limited contacts with people from this culture); (iii) *medium* (living in the country for a short period, sequence of regular short visits, relationships with people from this nationality); or (iv) *high* (living in the country for a while; strong relationships with people from this nationality). Based in the top country score for each GLOBE cluster and the number of clusters for which the top country score is *high* or *medium*, the participants were divided into two groups: *Group1 – Narrow Cultural Exposure* (the participants' exposure as high or medium was to one or two GLOBE clusters only – usually the UK and the country in which they were born); *Group 2 – Broad Cultural Exposure* (the participants had medium or high exposure to three or more clusters).

Method. The sessions were conducted individually via a given URL to access Perico[8] and to provide feedback before, during, and after the interaction, as follows:

Pre-study questionnaire included questions on basic demographic data and cultural exposure based on the GLOBE clusters (see above). This was followed by the Cultural Intelligence Scale[9] questionnaire (CQS): CQ-strategy, CQ-knowledge (extended with questions about gestures, food and clothes), CQ-motivation, and CQ-behaviour.

Interaction session with Perico (30-45 min) covered four countries selected by the user - one country for each level of familiarity: *none*, *low*, *medium* and *high*. A session included a total of 92 questions – for each country, five probing questions for each of the topics: `gestures`, `food` and `clothing`, and two for each of `language`, `currency`, `HDI` and `GNI`. At the end of the dialogue about a country, Perico showed the aggregated UM for each topic for that country: *not-good* (the user did not answer correctly any question related to the topic), *need-improvement* (less than 50% correct answers), ok (correct answers 50-70%), *very good* (more than 70% correct answers). The participant was then asked to rate the accuracy of their UM for the selected country and topic as: *accurate* (agrees with Perico's diagnosis), *underestimated* (Perico's assessment was lower than the user's personal judgement) or *overestimated* (Perico's assessement was higher than the user's personal judgement). Also,

[8] The study URL is disabled; Perico can be accessed from http://imash.leeds.ac.uk:8080/perico/

[9] http://www.linnvandyne.com/fourfac.html

user comments on the UM scores and the session with Perico were collected. Overall, the dialogue sessions covered 36 different countries across all GLOBE Clusters. There were a total of 2024 questions – 440 for each of `gestures`, `food` and `clothing`, and 176 for each of `language`, `currency`, `HDI`, `GNI`.

Post-study questionnaire comprised of the CQ-knowledge part of the CQS questionnaire (see above), followed by the System Usability Scale (SUS)[10] questionnaire adapted for Perico - the first ten questions were unchanged; the last three questions were tailored to Perico's interaction: (SUS11) "The questions asked during the dialogue were easy to understand"; (SUS12) "The instructions provided during the dialogue were clear"; and (SUS13) "The assessment made by the dialogue was correct".

4 Results

Usability Scores. The overall usability of Perico based on the SUS scores (see Table 1) was very good. SUS4 and SUS10 indicate that the system was easy to learn and did not require additional support. Given that the participants had to answer 92 questions in 30-45 minutes, the mean dialogue-score (Table 2) indicates good quality. The results on user model accuracy (see below) shed light on the scores for SUS13 (correctness of Perico's assessment). The score for SUS1 (frequent use) can be explained with the lack of usage context in the evaluation instructions.

Table 1. SUS scores for general usability (scale: 0-4, the higher the number, the better).

SUS1	SUS2	SUS3	SUS4	SUS5	SUS6	SUS7	SUS8	SUS9	SUS10
1.9	2.7	3.0	3.8	2.3	2.3	3.4	2.8	3.1	3.7

Table 2. SUS scores for the dialogue in Perico (scale: 0-4, the higher the number, the better).

SUS11	SUS12	SUS13	Mean dialogue-score
3.0	3.3	2.5	3.0

Interaction Feedback. The relatively low scores on SUS5 (integration) and SUS6 (consistency) relate to deficiencies of Perico's interaction, which were highlighted in the users' comments, as summarised below.

Inadequate assertions: The users pointed at errors based on the DBpedia knowledge pool, e.g. '*Spain has 132 human development*', '*People in Cyprus use a garment called Icknield High School*'. Some facts were seen as 'historic', e.g. referring to clothes not used any more, such as '*People in Germany use a Garment called Altdeutsche Tracht*', or making statements that are not true, such as '*Frank is currency used in Germany*'. A user noted that the knowledge pool did not take globalisation into account – food, clothing, gestures have become common in countries which they did not originate from. Inadequate assertions are hard to detect automatically. Allowing the users to indicate that something is wrong with the question enables further filtering or extending of the extracted DBpedia fact pool.

[10] http://www.measuringusability.com/sus.php

Limited content: Some users commented that the gesture questions they were asked were mainly for USA; or that for some countries, e.g. Jordan, the dialogue presented mainly facts related to other countries. These cases relate to the use of negation forms – while useful for generating questions, the negation forms are less indicative for cultural assessment, which should reflect the resultant UM. Most users had concerns about the HDI and GNI questions - finding them confusing or superficial. Additional aspects to include in the dialogue when discussing a country, such as capital, population, climate, religion, festivals, popular sports, points of interest, were suggested.

Lacking coherence: Some users felt that the interaction was jumping from question to question and lacked structure (which was due to the random selection from the pool of possible axioms). A way to add structure could be to follow the GLOBE clusters, including strategies for deepening, i.e. probing the knowledge on countries in the same cultural cluster, broadening, i.e. exploring countries from different clusters, or comparing, i.e. relating countries by cultural topic (e.g. a participant suggested questions like '*How does Italy's income inequality compare to the UK – higher/lower?*').

Misleading sentence openers: Several users commented that the sentence openers had influence on the answers - 'experience', 'think' or 'know' about something provokes different responses, e.g. 'think' is more likely to elicit a guess even if the user does not know. To deal with this, users suggested asking for an explanation or adding an option for indicating that the answer was given by guessing (in addition to 'I don't know' as at the moment).

Cultural Intelligence Scores. The earlier Crowdflower study [8] showed a statistically significant decrease in the user's CQ-knowledge scores as a result of the interaction with Perico. As crowdsourcing scores could be unreliable, in this study we also analysed the CQS changes comparing the pre- and post-test self-assessment scores. The average scores for all CQS questions did not change much (4.18 in the pre-test and 4.05 in the post-test; marks 1-7, where 7 is highest confidence). The average values for all users on the relevant CQ-knowledge scores were lower in the post-test (4.15 in the pre-test and 3.86 in the post-test) but this difference was not statistically significant (Man-Whitney, $p=0.29$). However, considering only the 12 participants who were quite confident (CQ-knowledge scores in pre-test >4; included users from both groups), there was statistically significant decrease in their post-test CQ-knowledge scores (5.13 in pre-test, 4.5 in post-test; Man-Whitney, $p<0.0001$). The study results confirm that the interaction with Perico has an effect on the user's confidence when self-assessing their CQ-knowledge on gestures, clothing and food, especially for users who have high confidence scores before the interaction. Participants who lowered their scores were further interviewed - the main reason for lowering the CQ-knowledge confidence was the exposure to a diversity of instances of the selected topics; this made them realise that their knowledge was not as high as they thought before interacting with Perico.

UM Accuracy based on Topics and Cultural Exposure. The participants were asked to assess the accuracy of the user model for each country and selected topic. Zooming into the cultural exposure values per country sheds light into the reliability of the selected cultural topics, pointing at the usefulness of the questions generated from DBpedia on each topic. Based on all individual assessments, the percentage of

accuracy perceived by users was 81%, 10% of all cases Perico overestimated the users and 9% of all cases were underestimates. When the users had *none* or *low* exposure to a country, they found Perico's UM overestimated in 17% of the cases and underestimated in 7% of the cases (76% were accurate). In contrast, when the users had *medium* or *high* exposure to a country, they felt their user model was accurate 86%, where 3% was overstimated and 12% was underestimated (Figure 2 gives details).

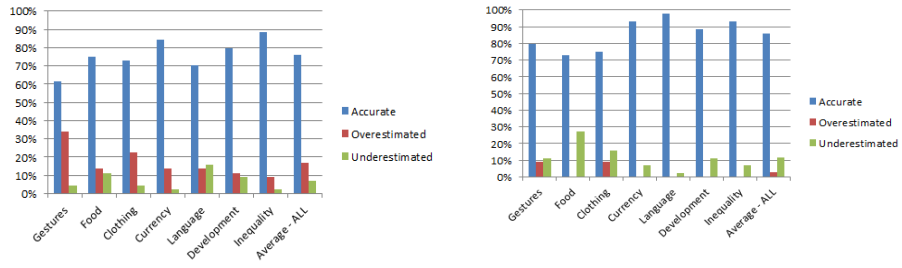

Fig. 2. User model accuracy based on selected cultural topics, compared against user exposure to the country – left (countries with none or low exposure) and right (medium or high).

Feedback on UM. The user comments in the cases when Perico overestimated or underestimated their UM provided useful feedback on the system's performance.

Answer indicated in the question: This feedback referred to cases when Perico overestimated the UM. The name of a country was given as part of the question and the correct answer was obvious. This happened mainly for currency, e.g. '*Indian Rupee*', '*Japanese Yen*', '*Bulgarian Lev*', '*Polish Zloty*', language, e.g. '*German is spoken in Germany.*' or gestures, e.g. '*Thai greeting*'. Such questions were seen as redundant, as they were not helpful for diagnosing the user's knowledge. Some possible strategies to avoid using the name directly could be: (i) identify countries which use the language, e.g. using `dbprop:regional` can be inferred that German is a regional language in Poland, and generate a question like '*Do Germany and Poland have a common language?*'; (ii) use `rdfs:label` to include a name without the country, e.g. use '*Złoty*' instead of '*Polish Zloty*', or `dbpprop:nickname` to include the nickname, e.g. use '*kint*' instead of '*Bulgarian lev*', or `dbpprop:subunitName` to use a subunit, e.g. '*Paisa*' instead of '*Indian rupee*'.

Answer given via knowledge elimination: Users felt that Perico evaluated their performance higher than what should be in the case when they had no knowledge of the countries they were evaluated for. They felt that they were able to answer the questions by knowing the facts (e.g. gesture or food facts) about other countries they knew rather than the focus item country for which they were diagnosed. For example, while evaluating a user on gestures from *Canada*, they knew the gesture in the question presented to them was a *Chinese gesture*, with which they had some exposure to. This enabled them to rule out the gesture's association with Canada and answer correctly. As commented above, a way to overcome the issue of user guessing is to (i) explicitly ask if the user knew or guessed the answer; and (ii) ask for additional justification or explanation. It should be noted that such questions are valuable for assessment of

culture-related knowledge, as the correct answers require knowledge of cultural aspects for other countries. This should be taken into account in the UM update.

Answer given using a clue in the question. This refers to questions with pictures – users felt that their cultural knowledge was overestimated as they could answer based on the picture presented in the question. For example, one user reported that: "*I could often tell if gestures were used in Hong Kong by looking at the picture - Most of which were obviously not taken in Hong Kong. Without the pictures I would probably have made more mistakes.*" Although the pictures give clue, they also make the questions more engaging and authentic and should not be disregarded. As above, a way to address this is by asking for justification of the answer or checking for a user's guess.

Answers were assessed wrongly. A main reason for the participants' statements that Perico underestimated their knowledge was that they believed certain facts were incorrect, which was observed exclusively for `gestures` and `food`. For example, "*High five is definitely used in Poland. Sign of the cross is as well*", "*Gestures for UK and US are very similar - Hook 'em bears is definitely used here.*" Perico uses DBpedia as a closed world source, i.e. if a certain fact is not in Wikipedia, it is false.

Being a crowd-sourced knowledgebase, Wikipedia does not always contain all the possible countries where a particular gesture is practiced or a particular food is part of the cuisine. This is linked to the issue of globalisation and points at the need to include a way of collecting facts from the users while interacting with Perico to accommodate richer crowdsourced knowledge of cultural aspects.

5 Related Work

Linked Data in general, and DBpedia in particular (a community effort to extract structured information from Wikipedia and make it available for free use [3]), has been a productive and popular source in user modeling and personalisation approaches. Considerable work has been done on enriching and semantically annotating social web content using linked data to improve adaptation and recommendation for content retrieval [9]; or semantically enrich and classify social tags to profile users [10,11]. Our work contributes to this growing trend to utilise Linked Data to address user modelling challenges[19], in this case we use DBpedia as a source of common sense knowledge for interactive user modelling in a domain requiring diverse content.

Due to the time and effort necessary to create assessment items (test questions) in e-assessment, automatic or semi-automatic item generation has gained attention over the last few years [12]. Linked Data is seen as a useful source for the generation of assessment items, offering models of factual knowledge and structured datasets for the generation of item model variables [13]. Several question answering systems for RDF data have been proposed, which in an essence translate questions into triples that are matched against the RDF data to retrieve an answer [14]. We use DBpedia on a similar premise. Our contribution to existing work in question-answering is the adaptation of this approach for interactive user modelling in the domain of culture.

Nationality-based cultural dimensions have utilised for culturally-aware user interfaces, e.g. [17, 18]. The prominent work in [16] brought in the topic of culture-based UM and Adaptation, presenting a way of automated customisation of the user inter-

face following a user's cultural model based on nationality. While the user's nationality is a useful source for adapting the interface, this is insufficient for user-adaptive learning environments when the focus is on developing cultural awareness skills. In such environments, the user's cultural exposure and awareness of cultural dimensions of other countries is crucial. To assess this, questionnaires are being used. However, questionnaires can become boring and user responses may get superficial [2]. While situational judgment tests, which assess a learner's recognition and understanding of cultural aspects by asking him/her to take decisions in carefully designed situations, can be effective, they require extensive design time by experienced subject matter experts [15]. Moreover, to be engaging, the test content has to include a range of examples, use images and other media [2]. Our approach paves a new avenue in culturally-aware user adaptive systems where freely available crowdsourced knowledge from Linked Data is utilised as a source of diversity for deriving culture-related UM.

6 Conclusions

Being an ill-defined domain, culture brings in an abundance of challenges for user modelling; given the rising importance of culture more attention will be paid at adding cultural dimensions in UM. The work presented here is an initial step in this direction. It extends conventional questionnaire approaches for deriving culture-related user profiles and addresses key limitations - dealing with diversity, enabling flexibility and extensibility. Cultural facts are derived from DBpedia and used for knowledge probing in an interactive user modelling system called Perico. It makes culture-related user modelling questionnaires more engaging by offering a diverse set of authentic examples and pictorial information in an intuitive and interactive way. In essence, Perico 'gamifies' questionnaires – a new strand of work which is seen as promising in a range of domains, including to deal with variations in cultural contexts [2].

The user study reported here shows that the approach has certain potential for user modelling. The users found Perico intuitive and, despite engaging in a 'questionnaire-like' interaction for more than 30 min, they gave positive usability scores. The users tend to agree with Perico's assessment in their UM. The systematic way to extract a portion of DBpedia facts, starting with seed concepts related to some aspects in interpersonal communication that have cultural variations, can be utilised to extend the knowledge pool with facts about country geography, religion, festivals, and tourism (suggested in the study). Having an underlying knowledge structure allows examining the UM accuracy based on categories; and hence the evaluation method can be repeated for an improved version of Perico with an extended knowledge pool.

Several remaining challenges require further investigation. Following a questionnaire style where examples are deliberately shown in a random-like way was seen as a lack of coherence in Perico's interaction. Ways for making Perico more conversational and open by allowing the user to provide justifications and suggest facts that are missing are worth investigating. The user modelling mechanism should be extended to take into account knowledge of other countries which is embedded in the questions. The utility of the questions can be further improved inferring more difficult facts (e.g. not making the answer obvious by giving the name of the country).

Acknowledgements. The research leading to these results has received funding from the European Union Seventh Framework Programme, grant ICT 257831 (ImREAL).

References

1. Gupta, V., Hanges, P.J., Dorfman, P. (2002). Review Cultural clusters: methodology and findings. *Journal of World Business*, 37(2), 11-15.
2. Puleston, J., Rintoul, D. (2012). Can survey gaming techniques cross continents? Examining cross-cultural reactions to creative questioning techniques,*ESOMAR Asia Pacific Conf.*
3. Bizer, C., Lehmann,J., Kobilarov,G., Auer,S., Becker,C., Cyganiak,R., Helmann,S. (2009). Dbpedia – a cristalization point for the Web of Data. *Web Semantics: Science, Services and Agents on the World Wide Web*, 154-165.
4. Heath, T., & Bizer, C. (2011). Linked data: Evolving the web into a global data space. *Synthesis lectures on the semantic web: theory and technology*, *1*(1), 1-136.
5. Karanasios, S., Thakker, D., Lau, L., Allen, D., Dimitrova, V., & Norman, A. (2013). Making sense of digital traces: An activity theory driven ontological approach. *Journal of the American Society for Information Science and Technology*, *64*(12), 2452-2467.
6. Blanchard, E.G., Karanasios, S., Dimitrova, V. (2013). A conceptual model of intercultural communication: Challenges, development method and achievements, In Proc. of 4th Int. Workshop on Culturally-Aware Tutoring Systems CATS2013 held at AIED2013.
7. Denaux, R., Dolbear, C., Hart, G., Dimitrova, V., & Cohn, A. G. (2011). Supporting domain experts to construct conceptual ontologies: A holistic approach. *Web Semantics: Science, Services and Agents on the World Wide Web*, *9*(2), 113-127.
8. Denaux, R., Dimitrova, V., Lau, L., Brna, P., Thakker. D., Steiner, C**.** (2014). Employing Linked Data and Dialogue for Modelling Cultural Awareness of a User. In *Proc. of IUI'14*.
9. Abel, F., Gao, Q., Houben, G. J., & Tao, K. (2011). Analyzing user modeling on twitter for personalized news recommendations. In *Proc. of UMAP2011*, 1-12, Springer
10. Abel, F., Herder, E., Houben, G. J., Henze, N., & Krause, D. (2013). Cross-system user modeling and personalization on the social web. *Journal of UMUAI*, 23(2-3), 169-209.
11. Meo, P. D., Ferrara, E., Abel, F., Aroyo, L., & Houben, G. J. (2013). Analyzing user behavior across social sharing environments. *ACM TIST*, 5(1), 14.
12. Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772-775.
13. Foulonneau, M. (2012). Generating educational assessment items from linked open data: the case of DBpedia. In *The Semantic Web: ESWC 2011 Workshops,* pp. 16-27, Springer.
14. Unger, C., Bühmann, L., Lehmann, J., Ngonga Ngomo, A. C., Gerber, D., & Cimiano, P. (2012). Template-based question answering over RDF data. In *Proc. of the 21st international conference on World Wide Web WWW2012*, 639-648, ACM.
15. Hays, M. J., Ogan, A., Lane, C. H. (2010) The Evolution of Assessment: Learning about Culture from a Serious Game. In *Proc. of the Workshop on Intelligent Tutoring Technologies for Ill-Defined Problems and Ill-Defined Domains* held at ITS2010, 37-44.
16. Reinecke, K., & Bernstein, A. (2009). Tell me where you've lived, and i'll tell you what you like: Adapting interfaces to cultural preferences. In *UMAP2009*, 185-196.
17. Marcus, A., & Gould, E. W. (2000). Crosscurrents: cultural dimensions and global Web user-interface design. *Interactions*, *7*(4), 32-46.
18. Dormann, C., & Chisalita, C. (2002). Cultural values in web site design. In *ECCE11 Proc.*
19. Herder, E., Dietze, S., d'Aquin, M., LinkedUp – Linking Web Data for Adaptive Education. (2013). UMAP2013 Workshops.