

# Taxonomy of personalisation for Generating personalised content in Technical Support Forums

**Solomon Gizaw**  
University of Limerick  
Localisation Research Center,  
Solomon.gizaw@ul.ie

**Jim Buckley**  
University of Limerick  
Localisation Research Center,  
Jim.buckley@ul.ie

## Abstract

There are two traditional approaches to meeting international users' requirements during content generation. The first is localisation which adapts a content to meet the language, cultural and other requirements of a specific locale. The second is personalisation which involves delivering relevant content and presenting information according to individual users' preferences.

The fundamental question that must be answered in generating personalised content is: what are the relevant attributes for personalising content? Work to date in personalisation has been based on several logic-based standards and frameworks that have been proposed. However these different standards have led to a degree of inconsistency in the field and are open to accusations of validity.

This research aims to empirically identify the relevant attributes for personalising content. It focuses on data obtained from technical support forums on the web, a growth area in terms of customer support. It uses a grounded-theory based approach to analyse the conversations on these forums in order to identify the personalisation attributes of the support provided. This paper reports on a preliminary study in this work, which analyses data from a number of technical support forums, and presents an initial taxonomy of empirically derived personalisation issues for this domain.

## 1 Introduction

With the growth of the WWW (World Wide Web) and the Internet, the necessity to address the needs of global markets with different cultural and individual requirements has also increased (Capstick et al. 1998). There are two traditional approaches to meeting such user requirements. The first is localisation which adapts

a content to meet the language, cultural and other requirements of a specific target market "locale"(Microsoft Press 1999). The term "locale" in this context refers to some specific regions or country. The second is personalisation that is primarily concerned with ensuring that information can be delivered to different end users in a format which reflects their specific needs (Gibb and Matthaiakis 2007).

In this context, localisation considers culture as a collective behavior or characters of a community who is living in some specific region or country. The information generated from the localisation process is expected to represent the cultural needs of that community. However the uniqueness of individual interests is not necessarily addressed in the current localisation process (H. Sun 2004). The same author, Sun, also stated that current localisation practices suffer from a narrow and static vision of culture resulting in usability problems for IT product and design. Localisation can therefore be seen as an intermediate stage before full personalisation (F. Gibb and I. Matthaiakis 2006).

On the other hand, personalisation involves delivering relevant content and presenting information according to individual users' preferences. These preferences are gathered explicitly or implicitly from the users to build a user model. Unlike localisation, the consideration is not bounded by locale rather; it goes beyond community interest and incorporates individual preferences.

There has been a plethora of research in the area of personalisation (J. Vesanen 2007, Miceli et al 2007, P. Ralph and J. Parsons 2006, A. Tuzhilin and G. Adomavicius 2001, K. Instone 2004, A. F. Smeaton and J. Callan 2005). This research

has proposed systems for personalisation of material and proposed different sets of attributes of personalisation, upon which these systems should be built. However, these proposed attributes are all theory-based and thus are open to accusations of being inconsistent and misguided. Thus it is the aim of this research to define personalisation attributes empirically and to rank these attributes in terms of order of importance for specific domains. The insights we ultimately derive will provide an empirical foundation for performing personalisation on content and constructing user models. Hence, this work attempts to develop an empirically-derived taxonomy of personalisation, to complement and enhance the existing theory-based taxonomies.

This paper reports on a preliminary study in that context where we use aspects of Grounded theory to identify the attributes exhibited in a web-mediated technical support domain. Technical support, as defined by (Das 2003), is a post sales service provided to customers of technology products to help them incorporate a given product into their work environment. It not only serves to improve the users' needs with respect to the product but can also provide a source of income for companies that provide the support services.

## 2 Literature Review

Personalisation is the process where customers receive different treatments based on their past behavior. These treatments are concerned with what is good for the business, serving the customers, and improving the quality of their resulting experience (Instone 2004). Thus personalisation describes the problem of how to use customer's information to optimize a business's relationship with its customers (Sahai and Kuno 2002).

The three core motivations for personalisation, from a user's perspective are (Blom 2000): to access information effectively, to accomplish a user's work goal, and to accommodate individual differences. To accomplish these core motivations, the implementation of personalisation has three dimensions: what to personalise (content, interface, service, modality), to whom to personalise (individuals or categories of individuals) and who does the personalisation (is it implicit or explicit) (F. Haiyan and P. M. Scott 2006).

With regard to this 2nd dimension, personalisation requires user information in order to create a user profile which can be used to identify, classify, store, filter, organise and present content which matches that individual's needs (F. Gibb and I. Matthaikakis 2006). Various personalisation applications can contain different types of attributes about individual users. However, in many applications, this attributes generally can be classified into two basic types – demographic and transactional, where demographic describes who the user is and transactional describes what the user does (A. Tuzhilin and G. Adomavicius 2001).

Different standards have been defined in the literature by different standard bodies, to identify and classify personalization attributes. Table 1 shows the categories of classification of personalization attributes as defined by some of these standard bodies. These attributes in turn are used as the basis for personalisation of content by many researchers in the field. Table 2 show some examples, detailing the personalisation attributes used by different researchers.

However, this literature basically is based on what researchers and service providers think are the user personalization issues rather than what the users actually want and which attributes really matter in the process of generating personalised content. Even though there has been an amount of work in implementation of personalisation, individual researchers have based their approaches on achieving personalisation goals that, while intuitively correct, have never been empirically evaluated as the core or even the correct, personalisation goals. This work attempts to address this by empirically deriving personalization attributes of relevance to the user.

## 3 Motivation

Many organizations have moved their customer care technical support from the product manual to the World Wide Web. Traditionally they have adopted a one-size-fits-all approach by delivering a FAQ and simple knowledge base search engine (Steichen and Wade 2010).

No	Standard Body	Name	Purpose	Attributes
1	W3C	P3P	standard for profile security	Demographic attributes (Identity, Age, Revenue) Professional attributes (Employer, Job category, Expertise) Behaviour attributes (Trace of previous queries, Time spent at each navigation link)
2	Telematics Information Engineering Project Number 8011	User profiles	For digital libraries	Personal data (Identity) Collected data (Content, Structure and origin of accessed documents) Delivery data (Time and support of delivery) Behavioural data (Trace of user-system interactions)
3	IEEE	Learning Object Meta-Data (LOM)	Educational Purpose	Educational Difficulty (Hard, Easy) Interactive Type (Active, Mixed, Expositive) Interactive level (low, medium, high) General Life Cycle
4	Dublin Core Metadata Initiative	Dublin Core Metadata Elements	Core Metadata standards	Title Subject Description Creator Publisher

Table 1 - Personalization Standards: Bodies and attributes

Attributes	Instances	References	Related to No
Product state	Product installation state Configuration state Pro-active actions Re-active actions	Steichen and Wade 2010	3
Knowledge state	Novice Expertise Procedural Specialised	Alba and Hutchinson 1987	3
User Values	Privacy Security Trust Brand Price	Cranor et al. 2002 Bart et al. 2005 Wind and Rangaswamy 2001	2
Orientation	Goal Oriented Utilitarian Hedonic	Celsi and Olson 1988 Hoffman and Novak 1996 Wolfinbarger & Gilly, 2001	1
Behaviour	Beliefs Interest E-Joyalty Involvement	Sun 2004 Reichheld and Scherer 2000 Srinivasan et al. 2002 Hoffman & Novak, 1996 Novak, Hoffman, & Yu-Fai, 2000 Anderson and Narus 1984	2
Demographic make ups	Gender Age Income	Sun 2004 Bouzaghoub and Kostadinov 2006 Tuzhilin and Adomavicius 2001	1
Content preference	Procedures content Consult Explanations Overview first	Steichen and Wade 2010	4
Process Type	Activity : Task Concept :- Narrative Text, Table, Image, Summary	Novak, Hoffman, & Dubachek, 2003	
Educational	Difficulty (Hard, Easy) Interactive Type (Active, Mixed, Expositive) Interactive level (low, medium, high)	Steichen and Wade 2010	3

Table 2:- Categories of Personalisation Attributes from the Literature

In contrast, the internet-based support systems that will thrive in the next generation will have to overcome the existing language and cultural barriers, particularly in applications operating in an international business environment (Schütz 1996). In addition, in order to meet user requirements to provide effective service for individual users, tailored to their preferences, service providers not only need to localise the content, but also to personalise it.

One problem with moving to this personalized support environment is that the research work in this area has been based on achieving personalisation goals that, while intuitively correct, have never been empirically evaluated as the core or even the correct, personalisation goals. This study will move towards addressing this gap in the literature by conducting an empirical study to identify the core attributes of personalisation in customer support.

## 4 The empirical Study

### 4.1 The Research question

The fundamental question that must be answered in generating personalised content is: (1) how can we generate content beyond localisation to satisfy user requirements? This, in turn is based on: (2) what are the relevant attributes for personalising content, and presentation of information of interest to users? These questions have not been addressed consistently in many areas, including the customer care technical support area. There is a need to have an empirical evidence of the core personalisation issues and thus to have a clear definition and measurable goal as a guide for generating personalised content.

### 4.2 The Design

This experiment has four phases: Selection of unit of analysis, data analysis method, coding process and interpretation and categorising process.

#### 4.2.1 Unit of Analysis

Considering the above limitations with regard to personalised content development, the paper will attempt to empirically identify relevant personalisation issues that arise in community support forum conversations and rank these issues towards characterising and generating personalised content in a customer care scenario. Community support forums are used because research has shown (Oxton 2010, Steichen and

Wade 2010) that customers are abandoning official technical support facilities provided by companies and increasingly migrating to community forums for their technical support, suggesting that these forums are giving customers the individual care they require.

Before we continue further let's define some of the terms which are used in this paper: A thread is a question posted in the forum with its responses from the community. There can be many threads in a forum. A message is the response of a question poster or other community participants to respond for the question. There can be one or more messages in one thread. We decided to include technical support forums which are the most popular ones from the Google search ranked list since the Google search ranks according to most visited forums. We have looked at their characteristics in terms of number of responses and time-delay of responses. At last we performed aspects of grounded theory on the individual messages in threads in these community forums.

The empirical first phase analysis is conducted on selected 7 IT technical support forums looking at a total of 31 threads which are categorised as: General Hardware, Networking, Security and virus, and Windows as shown in the figure 1.

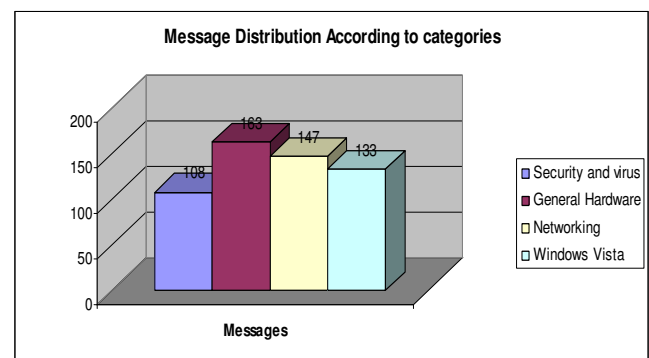


Figure 1: Message distribution According to Categories

The messages between the community forum member in each forum ranges between 3 and 54 with a total of 551 messages. The average messages per each thread are 17.8. The response time for each technical support request post ranges between 5 minutes to 5 Days. However 65% of the posts have responses within one or two hours.

## 4.2.2 Data analysis method

This study incorporates issues including the users' emotional and social drives and perspectives, their motivations, expectations, trust, identity, social and cultural norms. Consequently, we will use an inductive approach, driven from rich data, and employ the appropriate qualitative empirical analysis.

In these forums, we intend to employ statements as our data analysis unit of choice. A member can raise a question and post it so that other forum members can read and respond. Statements will initially be analysed for instances when users' signal that their needs are not addressed by the current response, as such statements strongly signal that an individual information need is not being met. The collected data set will be analysed to identify users' interests. One way of doing so is using the Emergent/Open Coding approach. The emergent coding approach is one of the techniques used to analyse text content in a qualitative research. It first examines a subset of the data independently and each data set develops a list of key coding categories based on interpretation of the data (Lazar et al 2010).

This research scenario also demands a non-predefined coding schema, so the method of Open Coding is suitable in this context. Open coding is the part of the analysis concerned with identifying, categorizing and describing phenomena found in the text. It analyses and identifies any interesting phenomena in the data (Lazar et al 2010). Each derived phenomenon will be coded. At last, similar instances are combined together in new ways after open coding by making connections using Axial Coding.

## 5 Results

### 5.1 Characteristics of community forums

After selecting the community forums for analysis, the coding of the messages is performed. This is done using open coding and in vivo coding to sort concepts into categories, so that in each category the concepts are both as similar as possible to each other, and as different as possible from concepts in every other category identified. Table 3 shows the characteristics of the forums categorised and their instances, examples and number of counts for each category.

## 6 Conclusion

Generally the number of responses on each thread on average and the time of responses for each post show the community forums are making efforts to deliver relevant and personalised information for their users and this shows that it is the right place to find characteristics and behaviors of different users.

A preliminary grounded finding of the messages shows users are primarily concerned with experience, trust, user-values, emotions and constraints respectively. User experience is main important issue to consider delivering personalised information. Characteristics of Emotion are OK, but it sometimes needs moderation. This shows that, the level of novice and expertise must be defined and categorised according to domain specific. Personalisation issues must have a system to identify the users' domain specific knowledge related to his questions. The taxonomy of personalisation already present doesn't put the priority of personalisation attributes into consideration.

The intention of responses intended to deliver only the relevant knowledge which doesn't consider other personalisation attributes. Even if there is a lot of discussion that are performed sometimes it ends without any solution for the user because of many misleading speculation and suggestions. However, sometimes good community forum participants try to understand the situation and try to answer accordingly in a way that the user can understand and use the information.

In the future much more analysis needs to be done with wider directed samples for more theory building and saturation.

### Acknowledgments

This research is based upon works supported by Science Foundation Ireland (Grant Number: 07/CE/I1142) as part of the Centre for Next Generation Localization ([www.cngl.ie](http://www.cngl.ie)).

Category	Instances	Example	Number of occurrence	Coding Type
Experience	Novice	<ul style="list-style-type: none"> <li>➤ I'm a total novice</li> <li>➤ have little knowledge</li> <li>➤ have no clue</li> </ul>	22	INVIVO
	Expertise	<ul style="list-style-type: none"> <li>➤ Im an IT Technician</li> <li>➤ Ive got a pretty good idea what im doing</li> </ul>	14	Open Coding
Trust	Distrust	<ul style="list-style-type: none"> <li>➤ I'm not about to put my credit card info</li> <li>➤ I'm afraid to delete it</li> </ul>	10	Open Coding
	Degree of Trust	<ul style="list-style-type: none"> <li>➤ Tend to agree</li> <li>➤ You are probably tell the truth</li> </ul>	6	Open Coding
User Values	Price	<ul style="list-style-type: none"> <li>➤ Cheaper Price</li> <li>➤ No cost</li> <li>➤ No need to pay</li> </ul>	10	IN VIVO
	Brand	<ul style="list-style-type: none"> <li>➤ Blame Your vendor</li> <li>➤ Your vendors don't want to support</li> </ul>	5	Open Coding
Emotional	Anger	<ul style="list-style-type: none"> <li>➤ It's hard to soar like an Eagle when you are flying with Turkeys</li> <li>➤ Would you cut off your legs while running a race?</li> <li>➤ You don't know what you are talking about</li> <li>➤</li> </ul>	11	Open Coding
	Emphasis	<ul style="list-style-type: none"> <li>➤ really handy thing called SEARCH</li> <li>➤ QUITE helpful</li> <li>➤ Very important</li> <li>➤ Disabled your Firewall</li> </ul>	8	Open Coding
	Frustration	<ul style="list-style-type: none"> <li>➤ It's hard to soar like an Eagle when you are flying with Turkeys</li> <li>➤ Would you cut off your legs while running a race?</li> <li>➤ You don't know what you are talking about</li> </ul>	4	Open Coding
	Stress	<ul style="list-style-type: none"> <li>➤ Oops sorry for the wrong doing, Shows how stressed i was</li> </ul>	3	Open Coding
Constraints	Moderate User	<ul style="list-style-type: none"> <li>➤ I am a volunteer here with a job and family so I ask that you be patient when waiting for replies.</li> </ul>	9	Open Coding
	Moderate Answer	<ul style="list-style-type: none"> <li>➤ Just FYI, first one must ascertain what is wrong before attempting to fix the problem. Just trying different fixes willy-nilly in hopes of resolving the problem is a waste of time and energy and more likely to make things worse than better.</li> </ul>	4	Open Coding

Table 2:- Categories of Personalisation Attributes from the Literature

## References

- Adomavicius G. and Tuzhilin A. 1999. *User Profiling in Personalization Applications through Rule Discovery and Validation*, KDD-99
- Alba J. W. and Hutchinson J. W. 1987. *Dimensions of Consumer Expertise*, Journal of Consumer Research
- Blom J. 2000. *Personalization: a taxonomy*, extended abstracts on Human factors in computing systems, The Hague
- Bouzeghoub M. and Kostadinov D. 2006. *Data personalization: a taxonomy of user profiles knowledge and a profile management tool*, Technical report ACI APMD, project MD-33
- Capstick J., Diagne A. K., Erbach G., Uszkoreit H., Cagno F., Gadaleta G., Hernandez J. A., Korte R., Leisenberg A., Leisenberg M. & Christ O. 1998. *MULINEX: Multilingual web search and navigation*, In Proceedings of Natural Language Processing and Industrial Applications
- Celsi R. L. and Olson J. C. 1998. *The Role of Involvement in Attention and Comprehension Processes*, Journal of Consumer Research
- Cranor L., Langheinrich M. and Marchiori M. 2002. *A P3P Preference Exchange Language 1.0*, W3C Working Draft <http://www.w3.org/TR/P3P-preferences/>
- Das A. 2003. *Knowledge and Productivity in Technical Support Work*, Management Science
- Gibb F. and Matthaikakis I. 2007. *A framework for assessing web site localisation*, Electronic Library, Vol. 25 Iss: 6, pp.664 – 678
- Hoffman D. L. and Novak T. P. 1996. *Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations*, Journal of Marketing
- Instone K. 2004. *An Information Architecture Perspective on Personalization*, Designing Personalized User Experiences in eCommerce, pp. 75-93
- Kirsh D. 2000. *A Few Thoughts on Cognitive Overload*, Intellectica, 30(1):19–51
- Lazar J., Feng J. H. and Hochhesier H. 2010. *Research Method in Human–Computer Interaction*, John Wiley & Sons Ltd publisher
- Microsoft Press 1999. *Computer Dictionary*
- Oxton G. 2010. *The power and value of on-line communities*, Consortium for Service Innovation in keynote, April CNGL meeting
- Reichheld F. F. and Schefter P. 2000. *E-loyalty*, Harvard Business Review, 78, pp 105–113
- Sahai A. and Kuno H. 2002. *My Agent Wants to Talk to Your Service: Personalizing Web Services through Agents*, in proceedings of the first international workshop on challenges in open agent systems, pp. 25--31.
- Schütz J. 1996. *Combining Language Technology and Web: Technology to Streamline an Automotive Hotline Support Service*, In: Proceedings of AMTA-96, Montreal, Canada.
- Sun H. 2004. *Expanding the Scope of Localisation: A Cultural Usability Perspective on Mobile Text Messaging Use in American and Chinese Contexts*, Rensselaer Polytechnic Institute Troy, New York
- Steichen B. and Wade V. 2010. *Adaptive Retrieval and Composition of Socio-Semantic Content for Personalised Customer Care*, Centre for Next Generation Localisation
- Wind J. and Rangaswamy A. 2001. *Customerization: The Next Revolution in Mass Customization*, Journal of Interactive Marketing, 15(1), pp13–32