

What can we learn from users of avatars in net environments?

Brigitte Krenn

Austrian Research Institute for
Artificial Intelligence (ÖFAI)
Schottengasse 3
A-1010 Vienna, Austria
+43 1 53246212

brigitte@ai.univie.ac.at

Erich Gstrein,

Barbara Neumayr
sysis interactive simulations ag
Hasnerstrasse 123, A-1160 Vienna,
Austria
+43 1 95555 400

{eg,bn}@sysis.at

Martine Grice

Institute of Phonetics
Saarland University, P 1150
D-66041 Saarbrücken,
Germany
+49 681 302 4696

mgrice@coli.uni-sb.de

ABSTRACT

In this paper we describe a commercial application of a *net environment*, and present the user data we have collected so far from three launches of this application. A net environment in our definition is a virtual space inhabited by avatars which have been created and are subsequently visited and instructed by users via the internet. Net environments are a useful means for studying user behaviour in general, and they are particularly well suited for presentation of multimedia content and systematic gathering of user responses on the appropriateness or effectiveness of the different presentations.

1. Introduction

The focus of this paper is on collecting user profile, behaviour and preference data in animated web spaces, or *net environments*, in which agent representatives of users interact with each other in a virtual location. The first part of the paper will describe a commercial application of a net environment and present the user data we have collected so far from three launches of this application. As well as actual results, especially details of how the users designed their avatars and which actions they took in which circumstances, we are interested in the general question as to which information can be accessed and how. On the basis of this first study, we suggest ways of collecting user data in more sophisticated net environment applications which are currently under development in the NECA project (IST-2000-28580, <http://www.ai.univie.ac.at/NECA>).

In the net environment applications we consider in this paper, users have a degree of control over the design of their avatars and have different ways of communicating via their avatars. Such applications are available on the internet, and can thus be easily accessed by a large number of users and by a broad range of user groups. They are a useful means for studying user behaviour in general, and they are particularly well suited for presentation of multimedia content and systematic gathering of user responses on

the appropriateness or effectiveness of the different presentations.

2. Net Environments

We coin the term *net environment* to denote a multi-user application for the internet

- where the users are represented by avatars which are situated in a virtual location, engage in social relations and fulfil specific tasks depending on what is required of them in a given application;
- where the user is able to design her/his avatar with respect to its graphical representation, its personality traits and emotional disposition, as well as its interests. The amount of freedom the user has for defining her/his avatar is application-specific;
- where the agents are autonomous after creation by the user; the user may influence the agents by giving them advice, which the agents may or may not take into account, depending on the agents personality and mood, or on parameters set within the application.

The SAFE (Social Avatar Format Engine) platform developed by *sysis* is an implementation of such a net environment. The motivation of any action performed by the SAFE agents¹ is the wish/necessity to satisfy needs which arise from a multi-dimensional 'system of needs' inside the agent, covering aspects like hunger, thirst, arousal and curiosity. The environment or habitat in which the agents are situated offers all resources to satisfy the needs of the agents. The simulation of a SAFE application is 24 hours a day, 7 days a week, although there are reserved time slots where agents normally rest (i.e. from 02:00 to 07:00). A further characteristic of SAFE applications developed by *sysis* is their game-like character. Three such games are

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

AAMAS '02 workshop "Embodied conversational agents – let's specify and evaluate them!", July 16, 2002, Bologna, Italy.

¹ In the following we refer to SAFE agents as avatars or agents: avatars because they are comparable to templates filled by the user; agents because each avatar, after creation, has its autonomous existence in the net environment. The terms avatar and agent have to be differentiated from the term Embodied Conversational Agent, as the latter focuses on the verbal and nonverbal communicative potential of an agent whereas this is not the case for our notion of avatar or agent.

Austropolis, a virtual democratic state where agents run for president, **Cool School**, an environment centering on school life, where the schoolchildren agents' main aims are to enjoy themselves and to gain popularity amongst their peers, and **Flirtboat**, a cruise liner on which agents socialise and flirt with each other. The user data collected in this paper is taken from three different launches of Flirtboat, (two in German for the Austrian market and one in English for the UK). Current launches of Flirtboat can be found under flirtboat.icircle.com for the English version and www.flirtboat.at for the Austrian version.

SAFE applications are good community building tools as they are designed for particular interest and user groups. At the same time, they are flexible enough to allow the user to actively take part in the environment. They are a promising interface for e-commerce as they can offer information and customer support in a discursive and jocular manner. At a further stage of sophistication when the animation of the characters has improved, natural language processing and expressive speech have been incorporated, and verbal and nonverbal aspects of communication have been integrated, SAFE applications will also be a valuable testbed for socio-cultural modelling and learning. At the current stage, they are already a useful tool for various kinds of data collection, as questions can be integrated into the games themselves, thus increasing the chances of collecting responses. In the first launch of the German version of Flirtboat, for instance, 80% of the questions were answered by the users, even though these questions were presented in the form of online questionnaires integrated in the application. The availability of Embodied Conversational Agents will enable us to present the user with the questions in a more appealing way. We expect that this will further positively influence the willingness of the user to answer these questions. Moreover, presenting agents in action and asking questions about these actions concurrently allows us to improve the quality and reliability of user responses.

Flirtboat is a SAFE application where agent representatives of the users interact to find a partner. The user selects her/his agent from a number of characters (graphical representations) available in the application, and also defines personality, emotional disposition and interests of the avatar following a multiple-choice questionnaire. The situatedness of agents in Flirtboat is that they are on holiday on board a cruise liner. It is a leisure-pleasure situation, formal encounters are not relevant or expected. There are such situations as meeting at the pool bar, being on deck, at the cinema or in the disco.

After creation, the user can only further influence the agents by giving them advice, in particular by answering questions generated by the system. As in other SAFE environments, the agents are autonomous; they may accept or reject the advice given, depending on their personality traits and their current emotional state. The agents also simulate physiological needs such as hunger, and respond to situational requirements within the Flirtboat application such as the number of dates which need to be made to meet a given target.

In the current implementation, the animation of the agents is rudimentary and no natural language or speech processing is available. Instead, text is selected from a database of fixed chunks and presented in written form. These restrictions were necessary to enable *sysis* to make the application widely available on the

internet, given the state of the art in internet technology, and given bandwidth considerations.

Currently, the disposition of an agent can only be expressed in the emotionality of a text, accompanied by a simple FLASH movie of the agent. These communicational restrictions are compensated for by an e-mail facility where the users can send mails to their own avatar's partners or friends.

3. Data Collected

With the existence of interface characters, a demand has arisen for studies on the effects of these characters on users. Existing studies cover investigations of the influence of animated characters on the comprehension and recall of information, see for instance [3], investigations of the trustworthiness of life-like agents or the affective relationship between the user and the agent, as reported in [1]. These studies typically concentrate on user-agent interaction.

Another type of application is presented in [2]. Here the user is presented with two characters (Agneta and Frida) that are present on the screen. The characters have their personal attitudes and preferences, and their own background story which they gradually reveal to the user while commenting on user actions. The Agneta and Frida setting is closest to the new generation of net environments being developed in the NECA project.

In this paper we provide a brief survey of an evaluation conducted for the Flirtboat application. Flirtboat is available in German and English versions. The former was initially designed for the Austrian market. The latter is an adaptation where the graphical representations have been kept identical and the texts have been translated by a native English texter. The survey covers data from two Austrian launches (henceforth AUT Flirtboat 1, AUT Flirtboat 2) and a UK launch (UK Flirtboat 2).² Thus multi-lingual and multi-cultural data are available as well as data from subsequent launches of the application in one country.

Two classes of user data are analysed: (1) user definition/selection of avatar characteristics such as gender, age and personality; (2) data related to user activities such as if and how the users answer questions from the system, and which actions are instigated when users or their avatars take up contact with another avatar.

As the avatars are representatives of the users in the net environment, we might assume that avatar design at least to some extent reflects user characteristics. However, it is still an open question as to how we can reliably assess how far this is the case. The following data should therefore be taken primarily as avatar and not user characteristics.

3.1 Analysis of Avatar Design

Avatars were designed as follows:

Age: Age is grouped into five broad classes by the system, 13-19, 20-29, 30-39, 40-49, over 50. The users assign one of these age classes to their avatars. Comparing the Flirtboat 2 versions, we find that the age distribution of avatars is fairly constant for the Austrian and the UK application. The majority of avatars belong to the group of under 30s, whereas the over 40s are strongly

² Note these are just naming conventions and do not refer to version number.

underrepresented, and there are even fewer avatars in the age group over 50. See figure 1.

Gender: Whereas in AUT and UK Flirtboat 2 female avatars significantly outnumber male ones in the age group 13-19, we have an inverse distribution in all other age groups. The predominance of male avatars is particularly interesting in the UK version of Flirtboat, as it is hosted on a women's channel. The Austrian data reflect the gender distribution of Austrian internet users as reported in the Austrian Internet Monitor (<http://www.integral.co.at>).

Personality traits: Avatar personality is modelled along four dimensions: extroversion – introversion, intuition – sensing, thinking – feeling, judging – perceiving which lead to 16 personality types. For the psychological background see <http://www.socionics.com/main/types.htm>.

Within the data related to the dimensions extroversion – introversion and thinking – feeling more than half of the avatars were created with introverted personality in both countries. Likewise, more than half of the avatars were created with a feeling type personality.

In AUT Flirtboat 2 the proportion of avatars designed as introverts is particularly high among those specified as aged above 40. While in the Austrian sample the proportion of introverts in the different age groups varies between 52% and 67%, in the UK the variation is small (ranging from 53% to 57%). See figure 2.

Differences between avatar genders along the thinking – feeling dimension are greater in the UK sample. Generally the differences are more pronounced than extroversion and introversion. The standard assumption, that human females are more likely to belong to the feeling type, seems to be reflected in the way users designed their avatars.

3.2 Analysis of User Activities

A number of activities were available for users in the applications. Two types of activity can be distinguished: (1) Responses to questions posed by the system concerning life style, attitude, personality and flirting behaviour. (2) Actions for communication with other avatars and users.

3.2.1 Responses to System Questions

In this section we investigate the willingness of the users to answer questions generated by the system concerning personality, lifestyle, attitude and flirting. The answers are used within the game to enhance the avatar personality. The questions are formulated in the first person in order to increase the identification of the user with her/his answers, so that the avatar ideally becomes an alter ego of the user.

Due to game dynamics, attitude questions were accessible over a period of a month after the user had set up an avatar. Several attitude questions could be answered at a time if the user wished to do so, whereas flirt tips were only offered on particular days. Only either one personality question or one lifestyle question was available per login. Table 1 provides an overview of the question types, including how many of each type are available, and an example for each question type.

The availability of the different types of questions per login is also reflected in the average number of questions per type

answered by the users, see table 2. As the table shows, users answered a high number of questions regarding attitude, with half of the Austrian users answering more than 43 questions and half of the English users answering more than 45. Generally the means in the UK sample are lower.

Interestingly those with male avatars are less likely to answer questions, regardless of question type and Flirtboat version, as can be seen in table 3.

An analysis of answering behaviour by age groups has shown that in both the Austrian and the UK samples the youngest and oldest clusters were less prepared to answer questions on personality, lifestyle or flirting behaviour than the groups between 20 and 49, whereas no clear picture could be found with respect to attitudes. Moreover, users who designed an introverted avatar had a slightly lower propensity to answer questions, regardless of nationality or type of question.

3.2.2 Communicative Actions Available to the Users

Three types of communicative actions are available to the users: (1) The users may choose from a list of system defined actions when her/his avatar takes up contact with another avatar. (2) The user may write e-mails on behalf of his/her avatar to the avatar's friends or partners. (3) The users may directly engage in communication with other users via the chat facility.

Actions: The user has the following possibilities for instigating actions. Actions can be positive, negative or neutral. For instance 'bringing flowers' is classed as a positive action, whereas 'showing one's date with itching powder' is considered to be a negative action, 'inviting the other avatar to the pool bar' is a neutral action. Each action can be accompanied by free text input.

In AUT Flirtboat 1 one out of ten meetings were accompanied by an action. 80% of these actions had a positive character, and 4% were negative actions. The picture changed in AUT Flirtboat 2, where the number of actions decreased (only 8 out of 100 meetings were accompanied with a user-defined action), and the percentage of negative actions increased to 28%, mainly replacing neutral actions, which decreased to 2%. In UK Flirtboat 2, 50% fewer actions are taken by the users than in AUT Flirtboat 2. However 60% of these actions are accompanied by text input, and 80% of all actions have a positive character. In all applications, utterances accompanying an action are typically quite short, i.e., 38 characters on average.

E-mail: While in AUT Flirtboat 1, 308,122 mails were sent, on average 19 mails per avatar, the average number of mails per avatar was almost halved in AUT Flirtboat 2, i.e., there was an average of 10 mails per user. For UK Flirtboat 2 there was only an average of 8 mails per user. This may be due to the chat facility available in UK Flirtboat 2 and AUT Flirtboat 2. The average length of a mail in AUT Flirtboat 2, however, considerably increased – from 191 characters in AUT Flirtboat 1 to 295 characters in AUT Flirtboat 2. Mails are even longer in UK Flirtboat 2, 374 characters on average.

Chat: The data from UK Flirtboat 2 and AUT Flirtboat 2 have shown that the frequency of chat usage was different on different week days with the most intensive use on Sunday in the UK and on Monday in Austria. The chart of logins per time of day also show clear differences with peak- and off-peak times largely corresponding in both countries. The only time with hardly any

chat traffic at all was between three o'clock and eight o'clock in the morning, the absolute high lies between 21:00 and 22:00. Surprisingly the youngest age group (13-19) in both countries had the lowest mean for chat login frequency. At the other extreme, in the Austrian sample the highest mean for chat logins was observed in the oldest age group (above 50). Generally, the chat facility was used more by English users, not only in total numbers, but also in relative numbers (logins per user), and UK males use the chat slightly more frequently than females. Chat content was not logged.

3.3 User Satisfaction

Data was also collected on user satisfaction in AUT Flirtboat 1 after it had gone offline. In total 83 users responded. These represent 0.9% of the active users³.

Questions were answered on a five point scale with

sehr gut	gut	geht	eher nicht	gar nicht
excellent	good	average	rather poor	poor

for all questions except the one addressing the correlation of the avatar personality with the personality which the user intended to model. For this question the following scale was used:

Volltreffe r	zum Teil	ein Wenig	kaum	gar nicht
perfectly	fairly	moderatel y	hardl y	not at all

Summarizing the results, of the users who responded:

- 78% found the application excellent or good.
- 82% found it easy to find their way through the application. The ratings in these cases were excellent or good.
- 50.6% were fairly satisfied with their avatar, 19% were completely satisfied, and another 19% were moderately satisfied. These judgments point to the suitability of the psychological model incorporated in the application.

Asked what they wished for a new release, 38.5% wished they had more information about the other avatars (users) in the environment, but only 10.8% requested more animation. What we have found here is a clear interest in social engagement, the wish to get to know the co-inhabitants in the environment they or their avatars are part of. A reason why the requirement for more animation has been quite low might be that character animation in general is very rudimentary in the current Flirtboat. The visual representation of the avatar functions much more as a mask behind which users can conceal/reveal themselves, than as an active means of communication. The communication aspect in AUT Flirtboat 1 is covered by the actions that can be instigated

³ Users were considered inactive if 35 days elapsed without a login.

by the user, free text input, and e-mail facilities, as well as the chat in the Flirtboat 2 versions.

We predict that the judgment about the importance of animation will change in the NECA applications when the characters become more animated and thus more expressive.

4. Future Work

Summing up, from the analysis of the data collected, we can find out who is attracted by a particular application, which means of interaction is preferred by which user group, and how personality traits influence user behaviour.

Since animation was rudimentary and there was no speech synthesis, the current system cannot be employed for studying the effects of integrated verbal and nonverbal communication. In the final part of our presentation we shall outline the new generation of net environments which are currently being implemented within the NECA project. NECA will provide more sophisticated animation together with an integrated expressive speech synthesis component.

In this new system users will be presented with movies of agents interacting with each other and performing clearly defined tasks (e.g. making friends or rejecting an unwanted offer). User reactions to these clips will be collected by allowing the users to give their avatars feedback and advice in a number of ways.

Users can tutor their avatars after watching the movie, telling them what they did right and where they went wrong, ideally with hints on how to perform better the next time a similar situation arises. Such feedback provides the developers with information as to what users believe to be socially acceptable behaviour in a given situation. Even if the technology used in the application is not advanced enough to simulate this behaviour, it provides a goal for future research.

Furthermore, users will be able to give a global appropriateness rating of the agent's performance in the clip. This second type of information gathering is different from the first as it also accesses users' feelings about the appropriateness of the avatar's behaviour which the user might not be aware of. For example, they are doubtless aware that a scowl and a friendly smile are not equally appropriate actions when making friends or rejecting an offer. They might, however, have difficulties in pinpointing, or even verbalising, how the type of intonation or voice quality used plays a role in signalling the avatars intentions.

Once the target behaviour is established across a group of users in a given situation, the same method can be used for more diagnostic testing of the effectiveness of different aspects (facial expression, gesture, posture, intonation, voice quality) of avatar communication.

5. ACKNOWLEDGMENTS

This research is supported by the EC Project NECA IST-2000-28580.

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.

6. REFERENCES

- [1] Justine Cassell, Timothy Bickmore: Negotiated Collusion: Modelling Social Language and its Relationship Effects in Intelligent Agents. In: *User Modelling and Adaptive Interfaces*. in press.
- [2] Kristina Höök, Per Persson, Marie Sjölander. Evaluating Users' Experience of a Character-Enhanced Information Space. In *Journal of AI Communications*. Vol 13, No 3, 2000.
- [3] Susanne van Mulken, Elisabeth Andre, Jochen Müller. The Persona Effect: How Substantial Is It? In: *Proceedings of HCI-98*, Sheffield, UK.

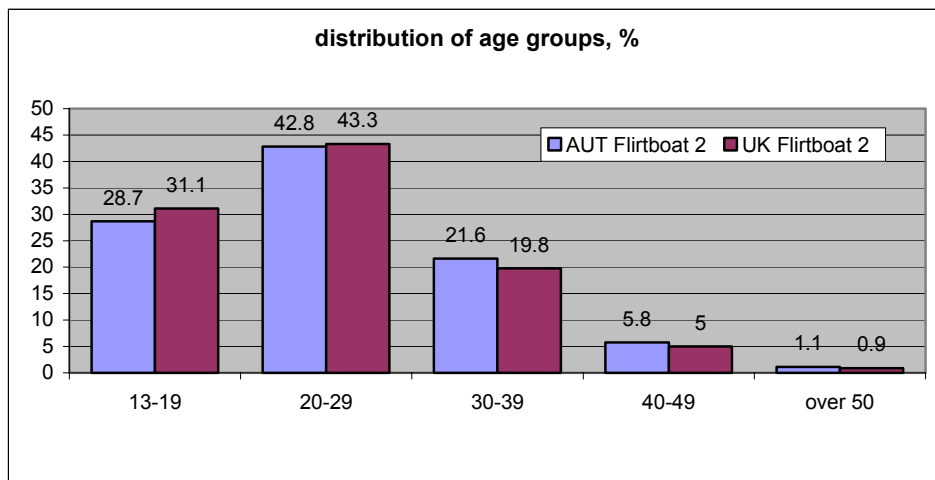


Figure 1: Distribution of age groups in Flirtboat 2

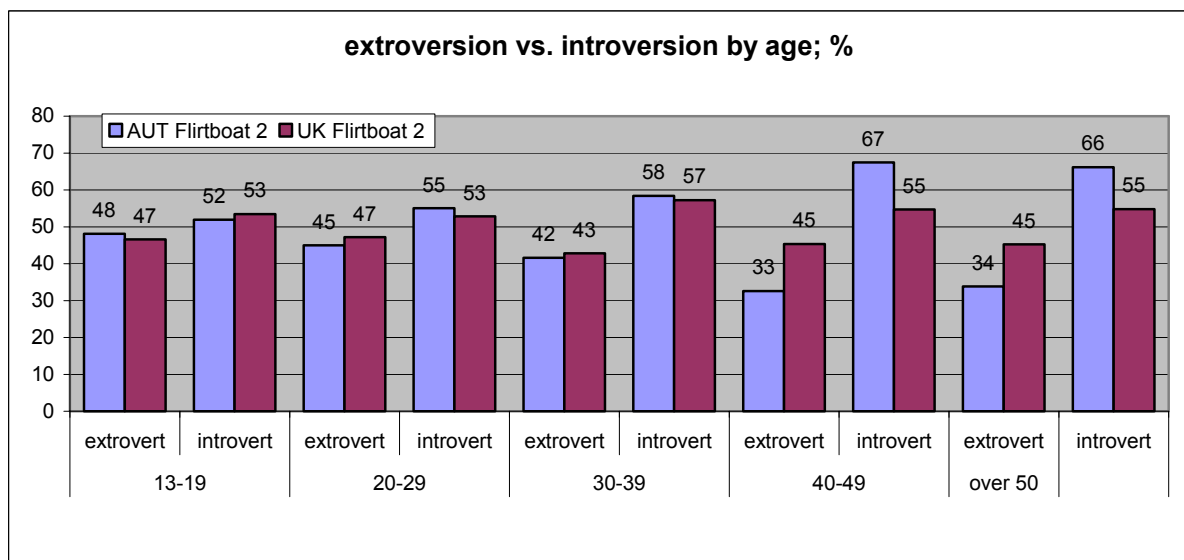


Figure 2: Distribution of extroverted and introverted avatars relative to the age group selected

Table 1: Examples of system questions

Question type	Number of questions	Example
Personality	32	Got up late, missed the bus, left my shoes at home. Do I go berserk or do I keep calm?
Lifestyle	21	Tonight I'll have... Sushi baked beans
Attitude	115	Good to see you. It's been a long time since my last holiday. Nothing but work. I started off in a small way with no one to rely on. But I've achieved everything I dreamed of. I hope you are ambitious too and want to make a career for yourself? want to go to the top not interested in careers
Flirt tips	89	I just popped into the Romero. Wow you should have seen the place! It's full of absolutely mind-blowing girls and boys. Imagine seeing your perfect match all alone at the bar. How do you get their attention? <ol style="list-style-type: none"> 1. I look them up and down insistently and devour them with my eyes. 2. I look deep into their eyes and smile <p>Feedback to 1: Be careful! Most people don't like being treated as a mere sex object. This could go horribly wrong. You'd better try a nice smile.</p> <p>Feedback to 2: Oh yes! Mucho romantico! And you're on the safe side as you'll see straight away whether your date is interested.</p>

Table 2: Number of questions answered per type

Flirtboat-Version	number of questions answered	personality	lifestyle	attitudes	flirt tips
AUT Flirtboat 2	Mean	8,50	6,42	46,58	8,68
	50% Quartile	3.00	2.00	43.00	3.00
UK Flirtboat 2	Mean	5,54	3,49	42,61	5,23
	50% Quartile	2.00	1.00	45.00	2.00

Table 3: Number of questions answered per type and gender

Flirtboat-Version	gender	total numbers of questions answered, means			flirt tips
		personality	lifestyle	attitudes	
AUT Flirtboat 2	female	8,67	7,21	48,35	8,98
	male	8,38	5,86	45,30	8,47
UK Flirtboat 2	female	6,06	3,80	43,07	5,75
	male	5,00	3,16	42,12	4,68