

# Flexible Services

## CrossMedia D1.7

### Service-independent user profiles as the source of personalization: Experiences

<b>Document Name</b>	CrossMedia D1.7 Service-independent user profiles as the source of personalization: Experiences
<b>Project/WP Title:</b>	CrossMedia WP1 - Next Generation Editorial Systems
<b>Document Type, Security</b>	Public

<b>Document Title:</b>	Service-independent user profiles as the source of personalization: Experiences
<b>Agreed date of delivery</b>	30.09.2010
<b>Actual date of delivery</b>	13.10.2010
<b>Editor</b>	Henri Kivinen
<b>Version</b>	1.0
<b>Date Last Change</b>	13.10.2010
<b>File:</b>	CrossMedia D1.7 Service-independent user profiles as the source of personalization - Experiences.doc

Participants	Name	e-mail
VTT	Asta Bäck	asta.back@vtt.fi

### Executive Summary

This report describes the results of testing the service independent profile in two cases: recommending events and magazine articles. Event recommendations were tested by KSF Media and VTT, and magazine article recommendations by Sanoma Magazines Finland (SMF) and VTT. The service-independent user profile service was previously developed by VTT and described in the deliverable 1.5 of the CrossMedia project.

Users created their profiles by entering tags into a profile creation form. When creating their profiles users were able to select tags from semantic vocabularies or to write keywords freely. User's tags were converted into a user profile, which was then matched to annotated articles or events. Finally, the articles or events that corresponded to the profile were recommended to the user.

**Table of contents**

1 Introduction ..... 5

2 The concept of profile service and analyses of profiles ..... 6

    2.1 Test users..... 6

    2.2 Profile service and profile creation ..... 6

    2.3 The use of tags in profiles ..... 10

    2.4 Importing and exporting interests and profile ..... 18

3 Conclusions ..... 19

## List of Acronyms and Abbreviations

KOKO	The Finnish Collaborative Holistic Ontology KOKO is a collection of Finnish core ontology, which have been merged together
YSO	The Finnish General Upper Ontology YSO is intended to be the main ontology in Finland, interlinking domain and instance ontology

## 1 Introduction

This report presents a case study of utilizing service independent semantic user profiles for recommending events and articles. Both cases are described separately, and this report only presents the results relating to the concept of service independent profile and it looks at what kind of tags users made into their profiles. The report is based on the material from both tests.

The event recommendation case was made by utilizing event data in KSF Media's Evenemax event database. For the test, KSF Media together with its partners added a lot of event data into the database. At VTT, full text index was created out of the event descriptions and event titles. The tests were carried out in a laboratory where the test users were invited to come. A web page was created for profile creation, and a paper questionnaire was made to gather feedback from the users.

The preparations for the article recommendation case study together with Sanoma Magazines consisted of several steps: selecting and annotating articles, choosing questions for profile creation, preparing the profile service, creating an online questionnaire form, recruiting test users, and analyzing the user profiles and questionnaire entries.

The tests were made for several purposes. First, to get users' opinions and experiences of the concept of service independent user profile and creating such profile with semantic support, and to test what users thought about recommendation based services as well as to test how well event/ article recommendations could be made based on the profile. The results of the profile recommendations are presented in the case specific reports.

The event recommendation test was made at VTT premises so test users were informed and guided through the test in person, and they could get immediate answers and help if something was unclear or difficult. The article recommendation case was made online, which relied on people to read the information and instructions about the test and what to do by themselves.

## 2 The concept of profile service and analyses of profiles

### 2.1 Test users

The test groups were of different sizes. There were only 15 test users in the KSF Media test (from now on KSF test), and 119 or 337 test users in the Sanoma Magazine (from now on SM test) test. In the SM test, 337 test users made a profile, but only 119 of these users completed the final questionnaire.

There was also a clear difference in the demographics of these two tests: in the SM test there were very few men (only 3 out of 119 were men), whereas in the KSF test, one third of the respondents (5 out of 15) were young men. The higher proportion of men in the KSF test was visible in the tags which included many kinds of different sports. Also the age distribution was different: in the KSF test group, 13 out of the 15 respondents were only 23 years or younger, whereas in the SM test group only 8 respondents out of 119 were 24 years old or younger.

### 2.2 Profile service and profile creation

The users were asked to rate various aspects relating to the profile service and profile creation on a scale from -2 to +2 in the KSF test and on a scale from 1 to 5 in the SM test. In this report, the SM test values have been converted to the scale from -2 to +2. The value +2 indicates the strongest positive opinion and -2 the strongest negative value.

The test persons had positive attitudes towards the concept of the service independent profile. The KSF test users gave in average 1,1 for the interestingness of the profile service concept, and the SM test users gave an even a bit higher value of 1,6. The usefulness of the concept were evaluated to be at about the same level: 1,1 by the KSF test users and 1,5 by the SM test users. The likelihood of using the service got a bit lower marks: 0,7 from the KSF test users and 1,3 from the SM test users.

Table 1 summarizes the average values of answers given to the questions relating to the profile creation. We can see that there was a greater variety of average values in the KSF answers. The KSF test users were of the opinion that the profile service functioned technically, and naming one's interests was not difficult. The KSF test users gave the lowest average points to finding the intended term and the usefulness of the tag proposals. The SM test users gave slightly positive average values for all these questions.

**Table 1. The average value of responses to questions relating to created the profile. The answers were coded so that +2 is the highest positive value and -2 the lowest, negative value; 0 is a neutral point.**

Question	KSF average (n=14 tai 15)	SM average (N=103...119)
Technical functionality (Miten vaikealta profiilin luominen tuntui teknisen toteutuksen osalta?)	1,6	0,4
Naming one's interests (Kuinka vaikealta profiilin muodostaminen tuntui omien kiinnostuskohteiden nimeämisen osalta?)	1,3	0,5
Ease of using tag suggestions (Miten arvioit automaattisten ehdotusten käyttämistä kiinnostuksen kohteidesi kuvaamisessa)	0,5	0,4
Finding the intended term (Löytyikö valmiista ehdotuksista, mitä tarkoittit?)	-0,1	0,6
The usefulness of the tag proposals as a whole (Valmiiden ehdotusten hyödyllisyys)	0,2	0,5

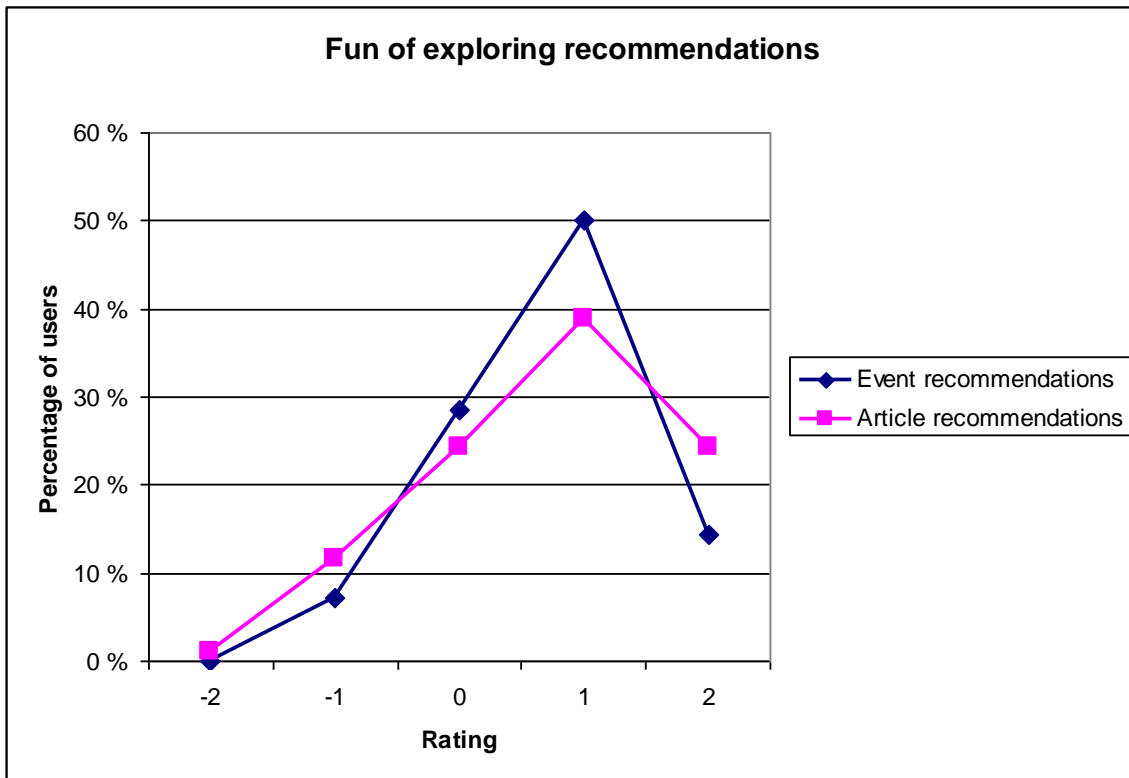
In both tests, the users were also asked whether they liked exploring event or article recommendations created based on their profiles. The users gave positive feedback on this question. The average values of the responses was the same in both test groups, and this value is 0,7. (Figure 1)

Figure 2 and Figure 3 show how users experienced using the profile and how likely they think they are to use the service. As can be expected, the evaluations for the interestingness of the concept of service independent profile service are higher than the likelihood of the user starting to use it.

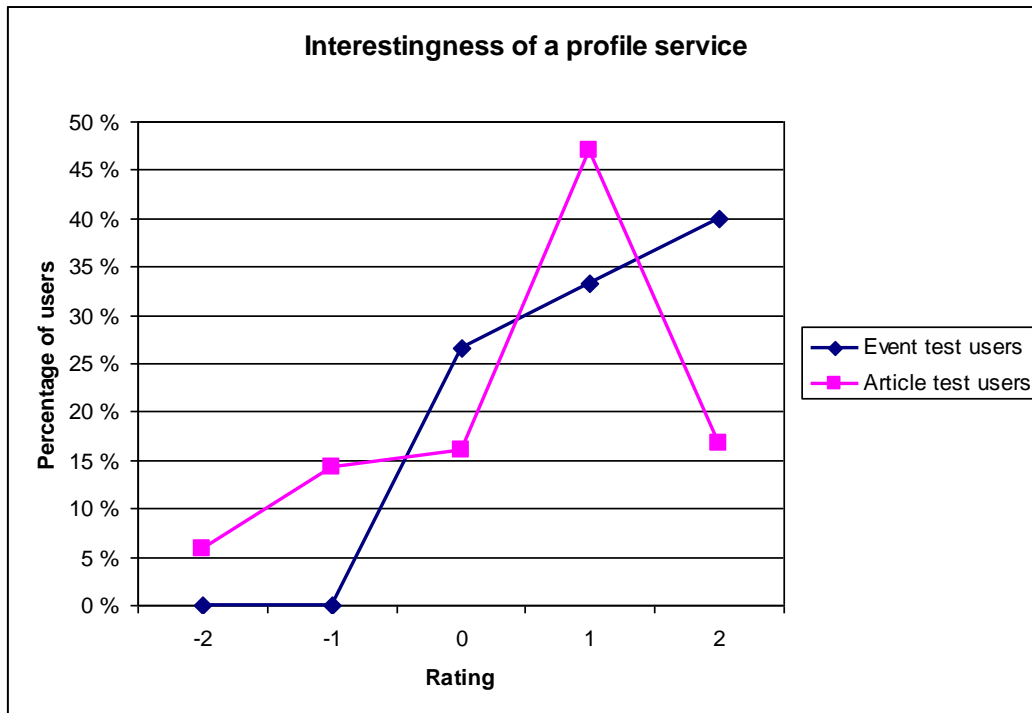
We can see that over 70 % of the KSF test users found the service concept very or somewhat interesting and the rest were neutral towards. Not a single negative rating was given by the KSF test users. Among the SM test, the interestingness was on average a bit lower, with the option +1 (somewhat interesting) begin the most

popular answer with nearly 40 % of the answers. 20 % of the test persons at the SM test gave a negative evaluation of the interestingness.

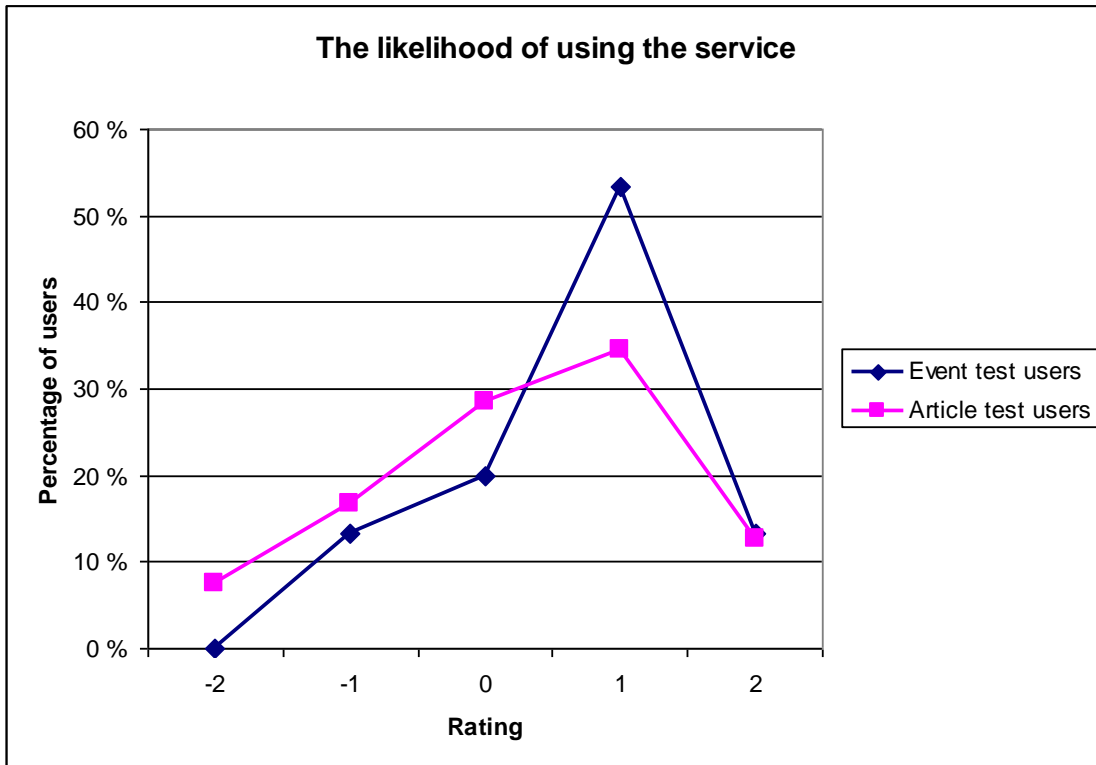
For the likelihood of using the service, the most common answer was +1 (somewhat likely) with 53 % of the KSF test users and 34 % of SM test users giving that answer. In both groups 13 % of the users indicated that they were very likely to use such a service.



**Figure 1. How fun it was to exploring events and articles based on one’s personal profile (2 = very fun; -2 = not fun at all).**



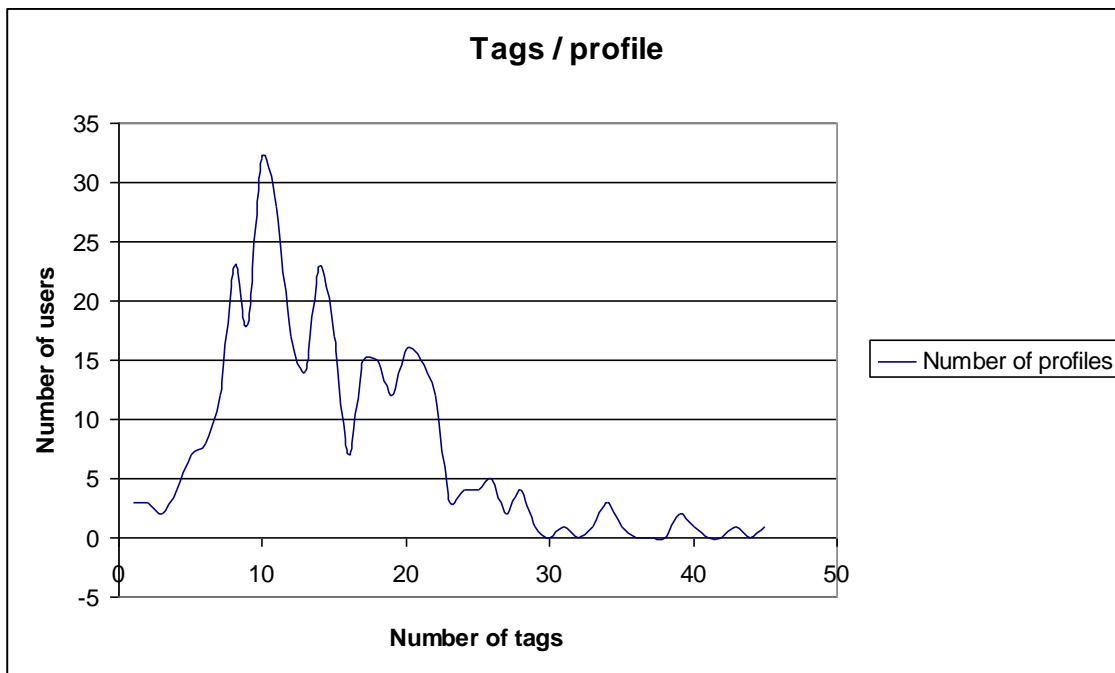
**Figure 2. The interestingness of the concept of a service independent profile service (2 = very interesting; -2 = not interesting at all).**



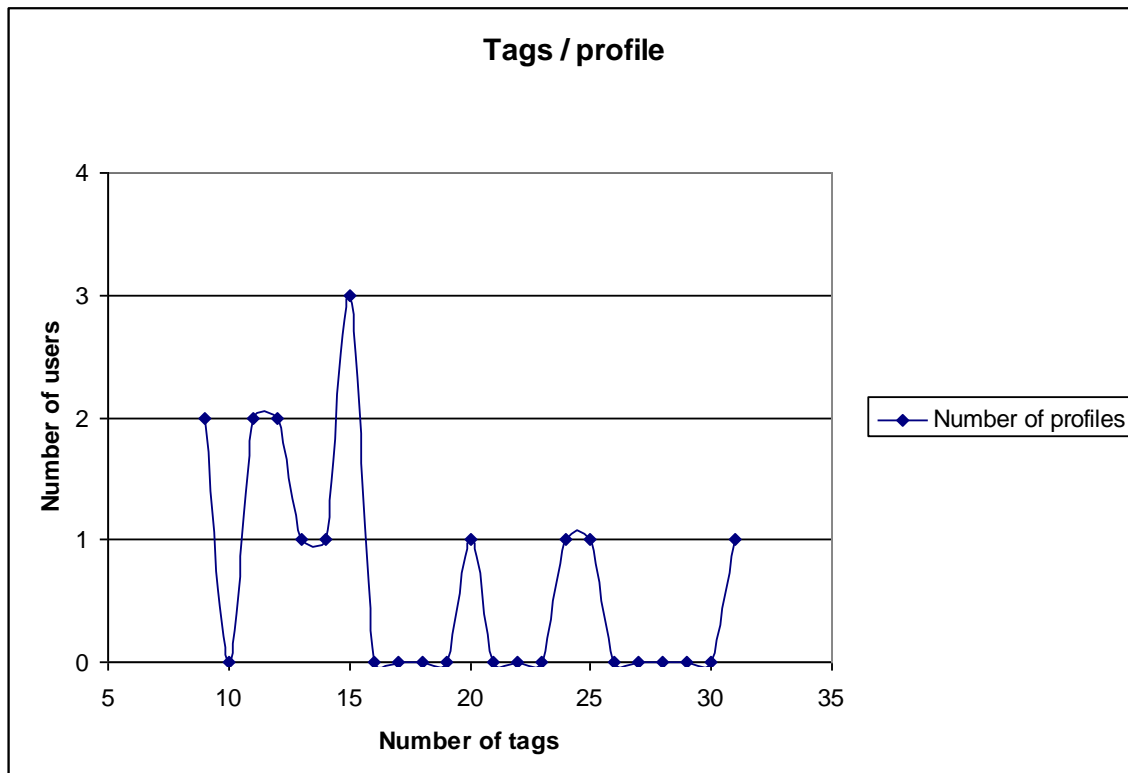
**Figure 3. The likelihood of using the service independent profile service (2 = very likely; -2 = not likely at all)..**

### 2.3 The use of tags in profiles

Users made on average 14,8 tags for themselves in the SM test (Figure 4) and 15,7 tags in the KSF test (Figure 5). In the SM test, the lowest quarter made 9 tags at the most, half did at least 13 tags, and the upmost quarter made at least 19 tags. The highest number of tags was 45 in the SM test and 31 in the KSF test.



**Figure 4. Number of tags users created for their profile in the Sanoma Magazine (SM) test.**



**Figure 5. Number of tags users created for their profile in the KSF media (KSF) test.**

To see how much variety and originality there was in the tags, we had to decide which tags to regard as being the same and which not. The following rules were used in separating tags, when counting the number of different tags:

- Singular and plural versions of one word were counted as being one and the same; also the use of capital letters was ignored.
- A tag that was linked to the semantic database was regarded as the same as a freely written word, if the label was the same.
- Other different variations of one concept were regarded as different tags.
- If the user had not known how to separate different concepts by using comma (,), several words were combined into one concept. This kind of a combination was regarded as one tag, and we did not separate the words into different tags. In the SM data set, there were 71 (or 1,5 %) such combined tags. In the KSF dataset there were only two such tags.

In the SM test, users made 1159 different tags. There are 337 users who made a profile, so on average each user created 3,4 new tags or concepts. In total, there were 4892 tags, which gives the average of 14,8 tags / user.

In the KSF test, users created much more new tags per user. The KFS dataset contained 190 different tags, and only 15 users, so in average, each users created 12,7 new tags. Most of the tags had been used only by one users, and on average, each tag was used 1,24 times. This difference to the SM dataset may partly be explained by the small number of users in this test. Also the differences in categories of the profile creation page may have contributed to the higher variety of tags in the KSF dataset: the KSF profile creation form included separate fields for favourite music, films and books, and people typically gave very specific information into these fields. The most popular tags among the KSF test users were comedy, dans, musik, and pop music.

Table 2 and Table 3 give some more statistics of the use of tags in these two datasets. In the SM test, the tags that had been given as examples in the profile creation page (“Example tags”) were used very frequently whereas they were hardly used at all by the KSF test users. The examples in the SM case were chosen so that they would be realistic and many people could identify them, but we must also assume that some people chose these ones just to be able to make the profile quickly and easily. Some of the example tags were however used in a different category than were it was given as an example, which indicates that people were not just using the examples as such. In the SM dataset 47 % of the 4892 tags were given as examples. On average these example tags were used 57 times, whereas other tags only 2,4 times.

**Table 2. The use of tags in the SM dataset, which contained 1159 different tags and 4892 tags in total made by 337 test users.**

Type of tags	Number of tags	How many times used in average	Total number / Percentage of the total amount of tags
"Example tags" visible in the profile creation page	40	57	2280 / 46,6 %
Tags incidentally created when the user did not know how to separate them with the comma.	70	1	70 / 1,5 %
Tags used only once (Excluding "example tags")	720	1	720 / 14,7 %
User created tags that had been used at least twice	328	5,5	1821 / 37,2 %

**Table 3. The use of tags in the KSF dataset, which contained 190 different tags and 236 tags in total made by 15 test users.**

Type of tags	Number of tags	How many times used in average	Total number / Percentage of the total amount of tags
"Example tags" visible in the profile creation page	11 example tags, out of which 6 had been used	2 / used tags or 1,1 / offered tag	12 / 236 = 5 %
Tags incidentally created when the user did not know how to separate them with the comma.	2	1	0,1 %
Tags used only once (Excluding "example tags")	159 (+ 4 example tags)	1	159 eli 67 %
User created tags that had been used at least twice	23	2,7	63 eli 27 %

In reality, the number of different concepts was not as large as these figures indicate, because there were many very closely related tags and concepts. For example the following words very given by users and from recommendation point of view, they have the same meaning:

- auto, autot, autoilu, auton ajaminen, autoilijat, autokauppa, Automobile, henkilöauto;
- mies, aviomies, puoliso, avomies;
- ero, avioero, avoero;
- ratsastus, hevonen;

People had also used a large variety of terms of describe watching TV, and additionally given names and types TV programs they like to watch. Also the names of certain TV programs were written in many ways adding to variety. It is important to know that a certain word or concept refers to a TV program, because it could lead to very wrong conclusions; e.g. The show called Dr. House was often written only as House, and this could be interpreted referring to a building if there is no indication of it being the name of a TV program.

Some questions in the SM profile creation form had lead people to write small sentences or at least to use many words to describe what they meant; e.g. “lapsen vaikeudet koulussa” (the child’s difficulties at school), or “rahaa voisi olla enemmän” (I would not mind having more money). Also relating to dreams people described the dream with additional details like going to a holiday to a very distant destination together with family.

There are also many concepts that are hard if not impossible to describe with one tag; e.g. “enemmän aikaa itselle” (more time for myself); “työn ja vapaa-ajan tasapaino” (balance between work and leisure time).

Some people had given very general term as their interests like “a new language” or “hobbies”.

We could also see the typical problems in freely written tags, like typos and mistakes in writing compound words correctly. Relating to picking terms from the ontology, we could also see that some users had probably picked a wrong term, and not the one they had intended.

Almost two thirds of the tags (64 %) were linked to ontology. Also the category where a tag has been given adds some semantics to the tag. For example, the concept money was given in different categories, and the category gives an additional perspective to the tag; e.g. somebody had given the tag money in the category “unpleasant things in my life”, which probably means that she has too little money or

other serious money related problems, whereas putting money into the category dreams, does not necessary imply to any big financial problems.

The popularity of the different tag categories is shown in Table 4 for the SM and Table 5 for the KM dataset. We can see that the first and the most general category of interests was the most popular one. All the categories were used to some extent. The free comments in the SM case indicated that some people thought that it was required to use all the categories.

**Table 4. The number of tags in the different tag categories in the Sanoma Magazine dataset.**

Category	Number of tags	Percentage of tags
I'm interested in (Minua kiinnostaa...)	846	17 %
I spend a lot of time with... (Asia johon käytän paljon aikaa)	541	11 %
Important to me (Minulle ovat tärkeitä)	660	13 %
My long term project (Kestoprojektini)	396	8 %
I would like to try (Haluaisin kokeilla seuraavia juttuja)	288	6 %
I would like to learn (Uusia asioita, joita haluaisin oppia)	287	6 %
I dream about (Haaveiluni kohteita)	439	9 %
I have (Minulla on)	461	9 %
I want to buy (Hankintalistallani)	338	7 %
My favourite TV shows (Mieliohjelmia TV:ssä)	421	9 %
Unpleasant things in my life (Ikäviä asioita elämässäni)	213	4 %

**Table 5. The number of tags in the different tag categories in the KSF Media dataset.**

Category	Number of tags	Percentage of tags
Interests (Intressen)	91	39 %
Favourite music (Favoritmusik)	64	27 %
Favourite films (Favoritfilmer)	46	19 %
Favourite books (Favoritböcker)	35	15 %

The profiles were created in a similar manner in both use cases by letting user fill in a form either by using their own words or by selecting suggested terms from semantic databases based on the first characters that the user types in. There were differences between the forms in how many fields were shown to the users and what the titles of the fields were.

In the KSF test only four different fields were used, and they were interests and favorite music, films and books. In the Sanoma Magazine case, there were 11 different fields including interests, but then also many more focused fields (the categories can be seen in Table 4).

We compared whether similar items were created by users in the different tests. We chose the tags that were filled in the interests field. The entries were prepared for the comparison by removing all duplicates and only using one concept for related terms. Finally, all terms were translated into Finnish in order to make it possible to spot identical terms.

After this clean-up there were 251 tags left from the SM case and 61 from the KSF case. We could identify 26 same tags in the both samples (Table 6). Common to these terms is that they are fairly high level concepts. Some more common terms would probably have been found if all tags had been included in the comparison.

**Table 6. Interests that users had indicated in both the KSF and SM tests.**

The name of the interest in Finnish	The name of the interest in English
Elokuvat	Films
Eläimet	Animals
Internet	Internet

Kesä	Summer
Kielet	Languages
Kirjallisuus	Literature
Kuntoilu	Gym, exercise
Leivonta	Baking
Lukeminen	Reading
Luonto	Nature
Matkat	Traveling
Muoti	Fashion
Musiikki	Music
Pilates	Pilates
Psykologia	Psychology
Ratsastus	Riding
Ruuanlaitto	Making food
Sisustus	Furnishing
Stand-up komiikka	Stand-up comedy
Taide	Art
Tanssi	Dancing
Teatteri	Theatre
Uinti	Swimming
Urheilu	Sports
Valokuvaus	Photography
Zumba	Zumba

### **2.4 Importing and exporting interests and profile**

Users were also asked whether they would like to import their interests from various social media services. In the SM test, 35 % gave a definite no as their answer, 34 % said maybe and only 7 % said yes. It is interesting to note that a quarter of the test persons did not know what to answer to this question.

In the KSF test, three persons or 21 % said that they would like to export information from other services, five (36 %) answered maybe and six (43 %) gave a definite no as their answer. Facebook, being the most popular social networking application, was mentioned several times in relation to exporting information from an external service. Some thought it could be a useful and acceptable source for creating a profile but others thought that there is too much activity and data about them in Facebook, and using it would not create a useful profile for personalising purposes. Issues like knowing who sees the profile and how its information is utilised were mentioned by several test persons are being critical issues and having impact on whether they were willing to create and share a profile. In the free comments of the SM test many respondents expressed concerns over how the profile would be utilised and were afraid that they would be bombarded with different kinds of advertisements and marketing mails and efforts.

In the KSF test, people were asked to rate to what kind of services they would like to take their personal profile to for personalisation. An events related service, as well as news and magazine type of services got mostly positive answers (yes or maybe). Out of the predefined choices, social networking services got the lowest mark. This answer raises questions whether the respondents understood the question correctly, because there already is a lot of interest profile related information about each user at social networking sites. Another interpretation for this is that people trust social networking services less than for example news or magazine related web sites.

### 3 Conclusions

Based on the users' ratings and observations in the test, we can see that users could get the idea of how the profile creation functioned technically. However, the ease of use did not get as high points as the functioning technically, and the usefulness of the proposed terms got an even lower overall mark. Making the profile creation easy to use is important so that people are encouraged to create a rich profile of themselves. In these tests people created on average 15 tags to describe their interests which can be regarded as a fairly high number for one session. If the profile is created to support long term use and use in different services, and this way the user sees real benefit of the profile, we can expect much larger and richer profiles.

Naming one's interests was evaluated as being relatively easy, but in the free text comments there were suggestions as to how to make naming interests easier. Also when we look at the tags that people created, we see that it would be useful to guide people into giving precise or detailed information about their interests. When the tags are connected to semantic ontology, it is easy to gather more information about tags. It is particularly easy and reliable to find more general terms relating to a tag, so therefore, it would be useful to encourage people to give rather more specific than more general tags.

Another aspect needing attention in further development is making it possible to express more complex issues as interests than only single tags or concepts.

Some users also expressed concerns about how much work creating and maintaining the profile would require.

People were hesitant about offering information about their web usage to be used as the basis for their personal profiles. In practice, web users' activities are very much utilised that way, but in many cases, there is no explicit connection to a certain person, and many users are unaware of this because it is done in ways that the user is unaware of.

Regarding the concept of user controllable and service independent user profile, it is important to offer clear information of who will be able to see the profile, and how the profile will be utilised. In other words, the user control should include not only what is included in the profile but also how it is utilised.