



Asta Bäck, Magnus Melin, Pirjo Näkki,
Sari Vainikainen, Risto Sarvas, Lassi Seppälä &
Sami Vihavainen

Tags and tagging

| Creating meanings, organizing, and socializing
| with metadata

Tags and tagging

Creating meanings, organising, and socialising with metadata

Report on the <täky> project

Asta Bäck, Magnus Melin, Pirjo Näkki & Sari Vainikainen
VTT

Risto Sarvas, Lassi Seppälä & Sami Vihavainen
HIIT, TKK



ISBN 978-951-38-7238-0 (soft back ed.)

ISSN 1455-0865 (soft back ed.)

Copyright © VTT 2008

JULKAISIJA – UTGIVARE – PUBLISHER

VTT, Vuorimiehentie 5, PL 1000, 02044 VTT
puh. vaihde 020 722 111, faksi 020 722 7001

VTT, Bergsmansvägen 5, PB 1000, 02044 VTT
tel. växel 020 722 111, fax 020 722 7001

VTT Technical Research Centre of Finland, Vuorimiehentie 5, P.O. Box 1000, FI-02044 VTT, Finland
phone internat. +358 20 722 111, fax +358 20 722 7001

VTT, Vuorimiehentie 3, PL 1000, 02044 VTT
puh. vaihde 020 722 111, faksi 020 722 7052

VTT, Bergsmansvägen 3, PB 1000, 02044 VTT
tel. växel 020 722 111, fax 020 722 7052

VTT Technical Research Centre of Finland, Vuorimiehentie 3, P.O. Box 1000, FI-02044 VTT, Finland
phone internat. +358 20 722 111, fax +358 20 722 7052

Technical editing Leena Ukaskoski

Bäck, Asta, Melin, Magnus, Näkki, Pirjo, Vainikainen, Sari, Sarvas, Risto, Seppälä, Lassi & Vihavainen, Sami. Tags and tagging: Creating meanings, organizing, and socializing with metadata. Report on the Täky project. 2008. VTT Tiedotteita – Research Notes 2449. 86 p. + app. 4 p.

Keywords tags, tagging, social bookmarking, social media, mobile, presence, location, applications, web application, semantic web, semantics

Abstract

This report summarises the results of the <täky> project (Käyttäjien tuottama metatieto merkitysten luojana ja osana käyttäjäkokemusta/ User created metadata as meaning indicator and part of the user experience). The project wanted to bring understanding on the role of tags and tagging as a part of user experience and value to companies when developing future media products and services. The project also wanted to study how freely defined tags and more formal ontologies can be connected to each others.

The project looked at tagging both in mobile and web applications. In mobile applications, the key question was what kind of a role does the user interface, application and visibility of tags play. In web applications, the aim was to study user-friendly ways of creating and utilising tags.

Tags were defined in the project as the automation of keyword search. This definition refers to that fact that a tag typically creates a dynamic link to other resources described with the same tag. Tags are more than keywords: they offer functionality. Another aspect differentiating tags from keywords is the cumulative effect of many people adding tags. In current applications, tags are not a single, well-defined concept. The applications and their user-interfaces as well as the purpose of tagging affects what kinds of tags people create. When adding the tagging feature to an application, it is important to consider all these aspects carefully in order to be able to provide a valuable feature to users.

A large number of user tests were made during the project, both utilising existing applications, such as Jaiku and del.icio.us, and applications that were developed in the project. Tilkut application was developed in the project for testing semantic support in tagging, and CamQ for playful mobile tagging. User test results are presented throughout the report. Ontologies for describing tags and tagging are presented with examples.

Opportunities and challenges in adding tags to a commercial service were studied with the help of interviews, questionnaires and by utilising a framework for evaluating and comparing social media features in existing applications and by comparing current and potential future features.

Preface

This publication contains the results of the project <täky> – User generated metadata as meaning indicator and part of the user experience (2006–2008). The project was nationally funded by the Finnish Funding Agency for Technology and Innovation (Tekes) and was carried out by VTT Technical Research Centre of Finland and Helsinki Institute for Information Technology HIIT, TKK. The goals of the project were to understand the role of tags and tagging as a part of user experience and to bring value to companies for offering media products and services.

Beside Tekes, we would like to thank AinaCom Oy, Nokia Oyj, Profium Oy and SanomaWSOY Oyj for supporting the project financially and by giving expertise and tools for the project. Additionally, we want to express our special thanks to Yahoo! Research Berkeley, and School of Information at UC Berkeley for their co-operation with the case studies, as well as all the users who participated in the user studies.

The project group members were Asta Bäck, Timo Järvinen, Magnus Melin, Pirjo Näkki and Sari Vainikainen from VTT, and Risto Sarvas, Lassi Seppälä and Sami Vihavainen from HIIT.

Executive summary

In this project we studied the role of tags and tagging in applications from an end-user's perspective, and how could that knowledge be used in building tagging functionality into products and services. Our results obtained from user studies can be categorized into the following three guiding sentences, which are in boldface.

Do not think that tags and tagging is a single concept. Tags and tagging are ambiguous concepts, which can not be thought as one single technology or phenomenon. Tags are used and created in a variety of contexts, for a variety of reasons, with a variety of meanings, and in a variety of very different systems. For example, tags in the photo management system Photoshop Elements are very different from tags in del.icio.us. Therefore, talking about tags and tagging can be misleading, and one should not think that they are always the same thing.

Key factors that shape tags and tagging are the social context of tagging (private, shared, or public), the user interface of the system (what can be done, what kind of use is implicitly suggested), and the openness of the system (can tags in one system be used in another). One implication of these is that one should design tagging so that the users have opportunities to control how and when to use the tags.

Nobody tags for the sake of tagging. Tagging can not be the aim in itself of a system, it is a tool for enabling something that people find useful or entertaining. Making the tagging process entertaining or game-like is one solution to motivate tagging. However, building an entertaining game or an application is not trivial. Leveraging interaction among people (*i.e.*, sociality) is another approach. However, then the social motivations within people should be understood, and those be the key drivers of development.

It is good to bear in mind that tags and tagging are an unfamiliar concept in general. It might be an everyday concept among IT professionals, but not among regular people who have to figure out what tags are in a particular system.

There are opportunities in building tags into products and services. Increasing the semantics of tags opens up new uses for them: better content analysis and new ways of aggregating and searching content, and also, the users themselves. This can be used, for example, in targeting advertisements. The semantics of tags can be increased by using existing semantic knowledge bases and vocabularies.

The following pages of this document dig deeper into the rationale behind these three takeaways and most certainly give insight and ideas for new products and services.

Contents

Abstract.....	3
Preface	4
Executive summary	5
1. Introduction.....	8
2. Defining a tag.....	10
2.1 About tags and tagging applications	10
2.2 Tag ontologies	11
2.3 Defining a tag in Täky project.....	16
2.4 References	19
3. Presence and tagging: two systems, three trials.....	21
3.1 Introduction	21
3.2 Materials and methods.....	22
3.2.1 The course of the user study.....	22
3.3 Studied location aware mobile tagging systems.....	23
3.3.1 Jaiku	23
3.3.2 Zonetag/Flickr	26
3.3.3 User groups	28
3.3.3.1 High school students (Finland)	28
3.3.3.2 Interest group of birdwatchers (Finland)	28
3.3.3.3 Group of urban adults (USA).....	28
3.4 Results	29
3.4.1 Adoption of Jaiku among the user groups.....	29
3.4.2 Location tagging in Jaiku	33
3.4.3 Privacy in location awareness	34
3.4.4 Specificity of location tagging	36
3.4.5 Adoption of ZoneTag among the user groups.....	37
3.4.5.1 Tagging	38
3.4.5.2 Specificity of location tags.....	40
3.5 User generated location tags vs. automatic geographical labels	41
3.6 Cross service location tagging.....	42
3.7 Discussion and conclusions.....	43
4. Game as a metadata / tag creation application.....	47
4.1 CamQ – A mobile tagging game.....	49

4.2	Conclusions	54
4.3	References	55
5.	Utilising tags in an online encyclopedia	57
5.1	Introduction	57
5.2	Background study	58
5.3	Service provider view	59
5.3.1	Workshop	59
5.3.2	Editor interviews	60
5.4	User view	60
5.4.1	Questionnaire	61
5.4.2	User interviews	61
5.4.3	User tests	62
5.5	General framework for social media features	64
5.5.1	Introduction to the framework	64
5.5.1.1	Concept and system	64
5.5.1.2	Content and user	65
5.5.2	Case: Wikipedia and the commercial online encyclopedia	66
5.6	Conclusions	68
5.7	References	69
6.	Intelligent Tagging – Case Tilkut	71
6.1	Introduction	71
6.2	The aim of the Tilkut service	71
6.3	Implementation	73
6.4	User tests	75
6.4.1	Users	75
6.4.2	Methods	76
6.4.3	Results	76
6.4.3.1	Background of users	76
6.4.3.2	Use statistics	77
6.4.3.3	Use experiences	77
6.4.4	Analysis of results	79
6.5	Semantic Tilkut	80
6.6	References	83
7.	Conclusion and discussion	84

Appendices

Appendix A: Example RDFs of Semantic Tilkut

Appendix B: Example SPARQL-queries of Semantic Tilkut

1. Introduction

When people became active as content creators in various web applications, it became necessary to offer some ways to manage this user-created content. A practical solution turned out to be tags, any words that people could freely attach to their resources, be they bookmarks, photos, music, places, or even people. Tags can be seen as a derivative of the old concept of metadata and index words. They have been used to manage resources like books. Traditionally metadata has been assigned only by professionals, and often by using predefined vocabularies. Tagging changed these both aspects.

Storing resources and their tags given by different people in an open internet application has created new additional value and opportunities. When different people describe a resource with their own words, we get cumulative information about what people find interesting or relevant in the resource. This is also descriptive metadata of the resource. People may find other people who are interested in same issues via common tags. It has been possible to create successful applications where people benefit directly from creating tags, but there is a also cumulative benefit from the tags created by all users. From media companies' point of view, tags raise many questions. Most of all, could tags and tagging be utilised to create metadata (*i.e.*, structured keywords) for the web resources, and how can tagging be utilised to enhance the user experience.

This report summarises the results of <täky> project (Käyttäjien tuottama metatieto merkitysten luojana ja osana käyttäjäkokemusta/ User created metadata as meaning indicator and part of the user experience). The goals of the project were to understand the role of tags and tagging as a part of user experience and to bring value to companies for offering media products and services.

The project looked at tagging both in mobile and web applications. In mobile applications, the key question was what kind of a role does the user interface, application and the visibility of tags play. In web applications, the aim was to study user-friendly ways of creating and utilising tags; this was done with the help of user tests of both existing applications and new ideas and applications. As the third research question, the project wanted to study how tags can be connected to ontologies. There are two possible directions here. There seems to be potential to utilise tags to add information into ontologies – here the challenges relate mainly to the assessing the reliability of tags. Tags could potentially be linked to available ontologies and this way the intelligence of the applications could be improved with the aim of offering better user experience. Also microformats, which offer a light-weight solution to adding semantics, were addressed in the project. The viewpoint was particularly creating easy ways of generating microformats.

The research in mobile tagging was carried out by HIIT. There the approach was to do qualitative user studies on different types of technology that either explicitly or implicitly have tags and tagging built into them. Two separate mobile applications, ZoneTag by Yahoo! Research Berkeley and Jaiku by Jaiku, were given to three groups of people: a group of high school friends in Finland, a group of bird watchers in Finland, and a group of young adults (“hipsters”) in San Francisco. Each mobile application was given for use for two months. The study and its results are reported in Chapter 3. In addition to these user studies HIIT built a camera phone game, the objective of which was to make tagging pictures entertaining. CamQ game and test results are reported in Chapter 4.

VTT studied tagging in web applications and the opportunities in linking tags and ontologies. VTT carried out a number of user tests throughout the project. An online encyclopedia was used as a case in evaluating the opportunities of tagging in enhancing the user experience and adding more metadata into media company resources. The key results of the encyclopedia case are reported in Chapter 5.

VTT also made application development to see how tagging could be done in a more structured way. The developed application, which is called Tilkut, also gave the opportunity to study the user experience in linking tags to ontologies, and generating event data as microformat. This work and its results are reported in Chapter 6.

VTT and HIIT researchers also made a common effort to define the concept of tag. The results of this are presented in Chapter 2 together with the state of the art about tags and tagging, and an overview of tag ontologies: Chapter 2.1 was mainly written by Asta Bäck, 2.2 by Sari Vainikainen, and 2.3 by Risto Sarvas. Each of the Chapters 3–6 is an independent whole with conclusions based on the results presented in the chapter. Chapter 3 was written by Sami Vihavainen, Chapter 4 by Lassi Seppälä, Chapter 5 by Sari Vainikainen, Pirjo Näkki and Asta Bäck, and Chapter 6 by Pirjo Näkki and Sari Vainikainen. The conclusion and discussion in Chapter 7 were written collaboratively by the already named researchers from these two institutes. The final chapter summarises the key findings of all studies mentioned above and discusses their implications from media companies’ point of view.

2. Defining a tag

2.1 About tags and tagging applications

A lot of the success of tags and tagging can be attributed to the freedom of being able to use them in any way that the user finds useful and of choosing any words he or she wants. One of the first applications utilising tags was del.icio.us (<http://del.icio.us>) and there tags are used for managing bookmarks. Flickr is another of the first popular application where tags were used. There users may share and publish their photos and describe them with tags.

Tagging is typically flat: all tags are at the same level and any number (or a high number) of tags can be applied to a resource. This has some drawbacks for utilising them – even for users and more so for applications that aim at utilising this information automatically.

Other well-known and frequent challenges with tags are that people use different words to describe the same thing, or a word has several different meanings (polysemy). People may also describe things at various levels of detail – an expert in a subject will use more detailed and specific words, whereas others use more general words. Also different forms of the same word (singulars, plurals, typos). [1]

In addition to differences in vocabularies there are also differences between people in how they tag and why they tag. Also, applications have different restrictions and support to tags, which naturally affects the user behaviour. Some people are very diligent and use many tags, some only few. Also, a single user's tagging habits evolve over time as they become more used to tagging and understand which ways of tagging serve their personal interests best.

In relation to describing things and concepts that would need two or more words, existing applications have taken different approaches. In some applications like in del.icio.us a tag must be one word, and other applications like Flickr and Kaboodle support creating tags that consist of more than one word.

There are several research papers [1, 2, 3, 4] that report studies about the type of tags people use. In these papers the work has been based mostly on del.icio.us or Flickr tags. In [4], Flickr tags (tags describing photos) were analysed. The purpose of the study was to add semantics to tags and convert them into RDF descriptions. The researchers identified following tag categories: place, time, event, name, action and camera.

Based on research papers, where del.icio.us tags were analysed, the following summary can be made out of the identified tag categories: topic, type of referenced resource, proper name (person, company, product, event, and location), subjective tags (adjectives, ratings), self reference, toDo tags and time.

When the aim is to utilise tags, different types of tags give different opportunities. Topic (like travel, semanticweb, cat, cars) can be used for analysing user's interests as well as characteristics of tagged resource. Type of referenced resources (like blog, wiki, video, music) can be utilised as information of user's preferred resource types. Proper name (like person, company, product, event, location (Nokia, BBC, Oslo)) can be used as an indication of interests as well, and relating to them, additional information may be obtainable on the web. For example, from the Nokia company profile, we can harvest information about their business areas such as mobile devices and mobile networks, and this indicates interest in these topics as well. Subjective tags (like interesting, great, +++) as well as toDo tags (like toRead, toPrint) can be used together with other tags to confirm user's interest areas. Self reference type of tags (like Mystuff, mywork) can be used by analysing the linked resources, which give indication of the user's interests and activities.

Tagging behaviour and practices vary a lot between people, because tagging is made mostly for personal use, not for common good. Users' tagging habits can be characterised with the help of the number of different tags, the total number of tags, the number of tagged resources as well as the number of tags assigned to a resource in average. Also a lexical analysis may be made on tags: how many of the user's tags are nouns, adjectives etc. and how much there are recognised words compared to unrecognised words.

Concurrent usage of tags is an interesting aspect. The general habit seems to be to tag from general tags to more detailed tags [1]; some users use a lot of synonyms or different spelling variants (plural, singular, acronyms) of a word or concept. Users also have different ways of combining words depending on the opportunities and restrictions of different applications. For example in del.icio.us, where only one word can be used as tag, people have adopted different ways to tag concepts that consist of two or more words, such as connecting the word with the help of a hyphen, an underscore, a slash or using caps at the beginning of each word.

2.2 Tag ontologies

Several ontologies have been defined relating to tagging. They define the key concepts like tags, resources and their relations. The focus of different ontologies varies and there

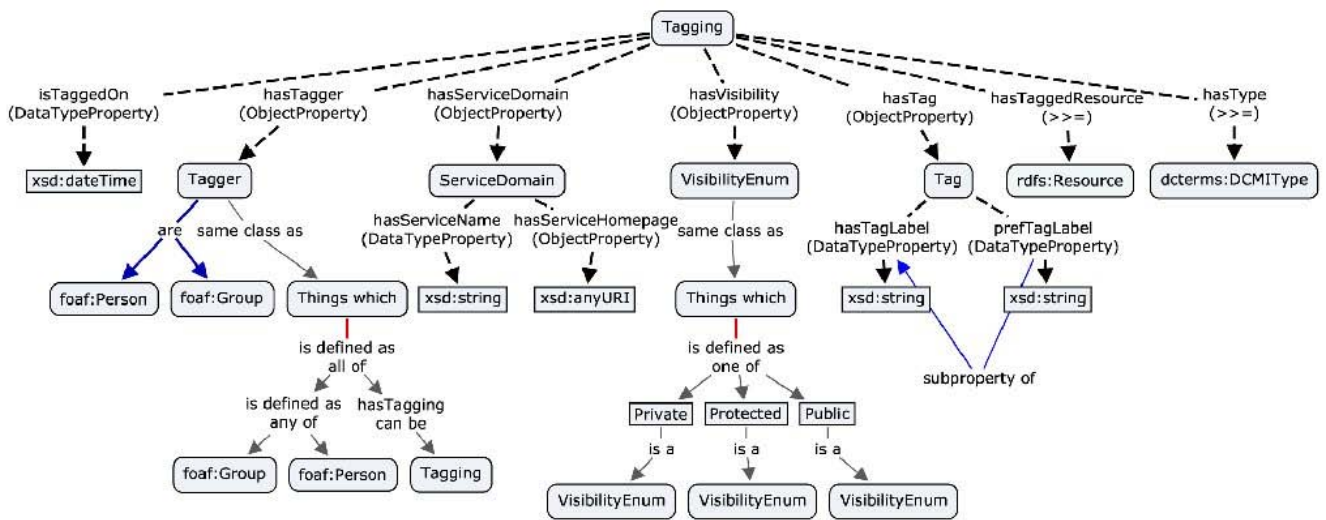


Figure 2. Visualisation of TagOnt ontology [9].

The SKOS (Simple Knowledge Organisation System) ontology is utilised to express the structure of vocabularies, thesauri, classification schemes, taxonomies and folksonomies. The main classes of SKOS ontology are Concept for describing the terms and ConceptScheme for describing a set of concepts. Concepts relates to the ConceptScheme with inScheme property. The relationship between the terms can be described with several different properties like narrower, broader and related. The list of classes and properties of the SKOS ontology are shown in Figure 3. The SKOS Specifications are currently published as W3C Working Drafts. [10] SKOS is quite widely used for describing thesauries and taxonomies semantically. In the Tag ontologies, the Tag class is determined as a subclass of SKOS's Concept, so the properties of SKOS can be utilised also for describing the tags.

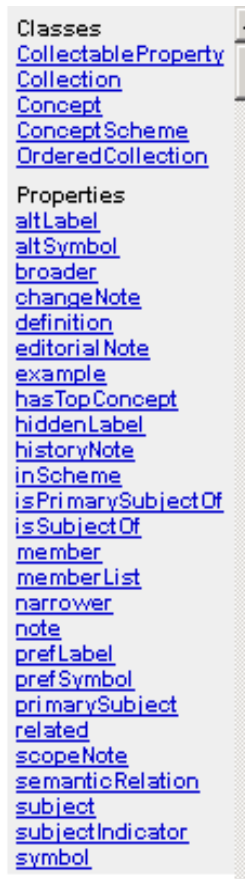


Figure 3. Classes and properties in the SKOS ontology.

The SCOT (Social Semantic Cloud Of Tags) ontology describes the tagging activities of the users as Tagcloud, which contains information about the tags and their usage frequencies and co-occurrences as well as the total number of tags and posts. The core concepts of SCOT are Tagcloud and Tag; TagCloud relates to Tag with the hasTag property. The SCOT ontology is linked to SIOC, FOAF, and SKOS-ontologies. TagCloud is a subclass of sioc:Container class and sioc:User relates to TagCloud class with hasScot property. Sioc:User is a subclass of foaf:OnlineAccount. [11]

One problem of tagging from information retrieval viewpoint, are the tags that have different meanings depending on the context. The MOAT (Meaning Of A Tag) ontology aims at solving this by providing a way to describe different meanings for the same tag. The meanings of the tags are described with help of semantic knowledge bases like DBpedia, Geonames or YSO (Yleinen Suomalainen Ontologia, General Finnish Ontology).

The core concepts of the MOAT are Meaning and Tag. The tag object extends the Tag class from Richard Newman's tag ontology. The link between a Tag and a Meaning object is described with the hasMeaning property. Each Meaning has exactly one

meaning described with help of URI (moat:meaningURI) and one person who has created this meaning (foaf:maker).

The tagMeaning property is used for linking Tagging and a meaning of a tag in a given context. The domain of Tagging is a tag:RestrictedTagging, which means that only one unique Tag is defined per Tagging. [12]

Annotea Bookmarks ontology is originally created for describing bookmarks but it could be utilised also for describing the tags (topics) relating to bookmarks [14, 15, 16]

Beside the tag ontologies, there is also need to be able to describe more in detail the resources that have been tagged, people and networks who have tagged and sites where the activity has taken place. There are several vocabularies and ontologies that can be used as building blocks when designing ontology for describing tagging and other user activities on the net.

SIOC (Semantically Interlinked Online-communities) is a framework and ontology aimed at connecting online community sites and internet-based discussions. SIOC provides rich, semantic data from social media sites and it can be used to store communication in social media services like bookmarks, blogs, photo sharing services etc. There are already exporters for applications like Wordpress blog software and social media services like Jaiku and Flickr.

FOAF (The Friend of a Friend) ontology describes people, links between them and things they have created and done. FOAF is widely used in connection to other ontologies for describing the Agent who has created for example a resource, tagging or comments. The agent can be a person, an organization, a group or other agent for automatically generated tags.

Combining the tag ontologies with SIOC and FOAF offers great opportunities to describe and combine user activities in different sites in a semantic way. The number of services that support above mentioned ontologies is increasing. Good examples of co-utilising SIOC, FOAF and SKOS as well as MOAT, SIOC and SKOS can be found in the article [13].

An example of utilising the ontologies is described in Chapter 6.5. Semantic Tilkut.

2.3 Defining a tag in Täky project

Research on tags and tagging activity brings forth a very simple issue: the different ways in which people use the word tag. Tags can be stickers with RFIDs, tags can be metadata on a photograph, tag can be a keyword, tag is called a label, and so on. Tagging and tags also suffer from being a buzzword associated with the Web 2.0 and social media phenomena, therefore, tags and tagging are used to brand a service or technology as “bleeding edge” or state-of-the-art.

From a research point of view, we have to be careful in using words which have ambiguous meanings, and be especially careful in using buzzwords loaded with hype. Especially, because there is a strong and established field of research in metadata and organization of information.

The same can be said from a business perspective: there is an established culture and invested technology in using metadata, so what is the relationship between, for example, current metadata systems in media companies and tags and tagging?

The process of defining what tags are should be re-thought. The problem is that rather than defining or specifying a certain technology the definition focuses on trying to capture all the uses and implementations of tags. This is fruitless, because tags are used in a myriad of ways in applications very different from each other. There is no one way of implementing tags and tagging, and there is no one way of using tags. Nevertheless, we try to find the common lowest common denominator in the ambiguity of the concept of tag and tagging.

As an example of the ambiguity of tags, Wikipedia defines a tag as follows:

“A tag is a (relevant) keyword or term associated with or assigned to a piece of information (a picture, a geographic map, a blog entry, a video clip etc.), thus describing the item and enabling keyword-based classification and search of information.”

On the other hand, Wikipedia defines an index term as follows:

“An index term or descriptor in Information Retrieval is a term that captures the essence of the topic of a document. It is used as keyword to retrieve documents in an information system, for instance a catalogue or a search engine. A popular form of keywords on the web are tags which are directly visible and can be assigned by non-experts also. Index terms can consist of a word, phrase, or alphanumerical term. They are created by analyzing the document either manually with subject indexing or automatically with automatic indexing or more sophisticated methods of keyword extraction. Index terms can either come from a controlled vocabulary or be freely assigned.”

Let's take these definitions as examples of the common use of these words in the information technology context, and let's take a closer look. The definition of a tag states that it is *a keyword or term associated with or assigned to a piece of information*. Then again in the definition of an index term it states: *A popular form of keywords on the web are tags which are directly visible and can be assigned by non-experts also*. According to these definitions, keywords and index words are tags and vice versa. The only exception might be that tags can be assigned by non-experts, and this implies that keywords can not.

So if tags and tagging are something new, they should be different from keywords and assigning keywords. But according to Wikipedia, there is no difference.

We see two ways how tags and keywords differ. The first one is in their use, and the second one is their technical characteristics.

It seems that the word tag is used when adding a keyword to a piece of information is informal. In other words, personal photographs or bookmarks are "tagged" rather than "annotated with keywords". Also, the definition of an index term implies that keywords are created by experts and tags are by non-experts. This suggests that if a piece of information has a keyword, the keyword is well thought and encapsulates an expert opinion.

As information and communication technology has become everyday household technology, such basic activities like adding keywords to personal information has also become a common task. Because these keywords are assigned often in a pastime and social context (in contrast to a more formal work and office context) they are less formal. Unlike in a work context, or official information organization context, keywords in a pastime context have no rules to follow and are, therefore, less formal. To put it simply: *tags are keywords used in an informal context*.

On the other hand, tags have also appeared at a certain phase in network technology: the domestication of the Internet. Point being, that perhaps someone came up with the idea of using keywords in an informal context already in the 1970s but the impact of the idea was small because there was no Internet. The Internet and the Web have made it possible to share, collect, and re-use tags in a network of computers and users. The social dimension of tagging and sharing tags would have not happened without the Web. Also, pioneering services like Flickr and del.icio.us have had an important role in shaping what is understood by tags.

Also, the use of tags emphasizes the communicative nature of keywords. Keywords traditionally are used to organize information and to communicate the contents of the information in few words. However, there are occasions when people use tags mainly to communicate, not to organize, and the communication is not necessarily linked to the

contents of the information, or the communication is implicit communication such as communicating distinction and togetherness by using words understandable only to a small group of friends.

Nevertheless, we see the big difference between keywords and tags in the *technology* associated with them. Keywords are passive alphanumerical strings that can be written by hand on a piece of paper. Tags, on the other hand, are alphanumerical strings that can be activated, for example, by clicking a mouse cursor on top of it.

Let's divide the brief history of the web roughly into two: Web 1.0 was the world of Netscape, home pages, and the Altavista search engine. Web 2.0 is the current web with YouTube, Google, web browsing on the phone, Facebook, and the Internet as a household concept. (As mentioned above, the term Web 2.0 is somewhat vague and ill defined, but a common enough term for this discussion).

Let us imagine the Flickr photo sharing and publishing service in the era of Web 1.0. It would be a collection of web pages with images under the web address www.flickr.com (in Web 1.0 no one dropped vowels). Next to the images the users can write keywords, for example, "cat". On the upper right corner the service had a search bar. If the user wrote the word "cat" on the search bar and clicked the go button, it would show the images that had the keyword "cat" next to them.

The Flickr service of Web 2.0 (the actual flickr.com) the user can write a tag next to an image. Then anyone can click on the tag "cat" and get all the images tagged with the word "cat". Therefore, a tag automates the task of writing a keyword on the search bar and pressing go. The end result of both Web 1.0 and 2.0 versions of the search for images with a keyword "cat" are the same. The tags are only faster because the user does not have to click on the text box (the search bar), type three letters, and click on the go button.

Therefore, we define tags as *an automation of a keyword search*.

So what does this definition imply? First, it means that tags are not words or anything static, tags are an operation. Tags always require a computer because they are like a script to facilitate users' searches. Keywords and index terms can be written by a pencil on a piece of paper – tags can not. However, keywords are an integral part of tags, because they are keywords with an operation built into them.

Second, tags are not the same as web links. Web links are, conventionally speaking, static connections between two resources on the web. Clicking a tag has the potential to produce different search results every time it is clicked – it is dynamic because it is a search operation.

Third, looking upon tags as operations brings forth an interesting finding: The automation of such a trivial and light task as writing a keyword in a search bar has generated numerous new uses and new ways of browsing the web. Popular services such as del.icio.us, Flickr, and Last.fm would not exist without tags and the easy browsing of web content that tags enable.

To summarize: Tags automate a keyword search into a single click. Tags are not keywords, tags are not links, tags are an operation like running a small script, and therefore, tags require a computer. Perhaps this sounds like stating the obvious, but in a world full of buzzwords and vague concepts this gives a new perspective on tags. It also brings tags into the sphere of existing research on information organization and metadata. Tags are not mysterious, they are keywords that have a search built into them. In the following writings in this report, we hope to shed light on the uses of keywords and tags and how they are used in ways perhaps not imaginable to the information organizers few decades back.

And we also suggest the following: be careful when using the word tag, you might be talking about keywords.

2.4 References

- [1] Golder S. A., and Huberman, B. A. The Structure of Collaborative Tagging Systems. <http://arxiv.org/ftp/cs/papers/0508/0508082.pdf>
- [2] Marlow, C., Naaman, M., Boyd, D., and Davis, M. Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead.
- [3] Xu, Z., Fu, Y., Mao, J., and Su, D. Towards the Semantic Web: Collaborative Tag Suggestions. http://www.ibiblio.org/www_tagging/2006/13.pdf
- [4] Maala, M. Z., Delteil, A., and Azough, A. A Conversion process from Flickr tags to RDF descriptions. <http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-245/paper5.pdf>
- [5] The TagCommons Working Group. <http://tagcommons.org/>
- [6] Richard Newman's blog entry about tags. <http://www.holygoat.co.uk/blog/entry/2005-03-23-2>

- [7] The schema of the Tag ontology.
<http://www.holygoat.co.uk/owl/redwood/0.1/tags/>
- [8] Knerr, T. Tagging Ontology – Towards a Common Ontology for Folksonomies.
<http://tagont.googlecode.com/files/TagOntPaper.pdf>
- [9] The schema of the TagOnt ontology. <http://code.google.com/p/tagont/>
- [10] SKOS Specifications & Documentation. <http://www.w3.org/2004/02/skos/specs>
- [11] SCOT Ontology Model. <http://scot-project.org/?p=5>
- [12] MOAT: Meaning Of A Tag. <http://moat-project.org/>
- [13] Bojārs, U., Passant, A., Cyganiak, R., and Breslin, J. (2008). Weaving SIOC into the Web of Linked Data. LDOW2008, April 22, 2008, Beijing, China.
<http://events.linkedata.org/ldow2008/papers/01-bojars-passant-weaving-sioc.pdf>
- [14] Annotea Bookmark schema.
<http://www.w3.org/2003/07/Annotea/BookmarkSchema-20030707>
- [15] Koivunen, M.-R. Web Tagging with Annotea Shared/Social. Bookmarks and Topics. <http://annotea.org/www2006/annotea.htm>
- [16] Annotea.org. <http://annotea.org/>

3. Presence and tagging: two systems, three trials

3.1 Introduction

Tagging is a phenomenon of the web 2.0 era. Tags are user generated and non hierarchical. User generated taxonomies are called folksonomies. Tags, being generated by the user means that the taxonomies are not created top-down but bottom-up. This means that no controlled vocabulary or predefined taxonomic structure exists, but everyone can have their own vocabulary without any cultural, social, political restrictions. They have a dualistic nature. Tags can be used in both for organizing media or for communication. In organizing user generated media, tags, are considered as a promising way to organize self made media in meaningful way especially to the user. The communication aspect of tags is related to presence or media sharing, depending on the used service.

Tagging is usually a side feature of a service. For example in presence sharing one can tag her location and share it with her friends. In media sharing one can tag her photographs for organizing purposes or to ‘advertise’ the media object to others. Even though there are advantages and different uses for tagging, it is not fully known how people really use them.

When we talk about mobile location tagging in many cases it means also location awareness. In this study two types of mobile tagging technologies were studied. A location based microblogging system Jaiku and a cameraphone photo uploading and tagging system ZoneTag. The object was to find out how so people use location tagging features in these systems. Especially,

- How does location based mobile tagging technologies adopt to people’s every day lives?
- What kind of privacy concerns users might have towards location awareness?
- How specifically are locations tagged?
- Would it be reasonable to use user generated location tags across the two services?

The report is structured as follows. Chapter 2 describes the background information about the course of the user study, the studied tagging technologies, and user groups recruited for the intervention study. Chapter 3 describes both the qualitative and quantitative results from the user studies. In chapter 4 the use of user generated location tags vs. automatic geographical labels is discussed. In chapter 5 the lessons learned from the user studies are used for discussion whether it would be reasonable to use user generated tags across several tagging systems.

3.2 Materials and methods

3.2.1 The course of the user study

For the user study three small social groups were recruited: Group of high school students (Finland), an interest group of birdwatchers (Finland), and a group of urban adults (USA).

All the participants were given Nokia N70 smartphones and free data plans and introduced to two social tagging services: Jaiku and Zonetag/Flickr. Every group used each service for two months. For the participants there was no other reward than being able to use phones with free data plan for the time of the use period.

The users were introduced to the services in group sessions. User accounts were made for every user in advance. Everyone was taught about the services individually so that it was sure that the services worked for everyone and everyone also knew how to use them. The users were also instructed to use the services like they would use them in their everyday lives. It was also emphasized that there was no obligation what so ever to use them if it did not feel natural. After the use period had been running for two weeks each user was contacted to check that everything is alright with them and there are no technical problems.

Most of the users were interviewed two times per service, before and after the use period. Because of scheduling problems some users were interviewed only once per service. The users were interviewed individually about their social networks, communication practices, and use of different communication channels with their social networks, how they used the studied services in their everyday life, and what kind of feelings they had about them. The interviews took place at users' homes or at their work place or at school. Each interview lasted about one hour. Interviews were audio recorded for further analysis.

Also quantitative data from the use of the both service was logged. Yahoo Research Berkeley logged the user data related ZoneTag use. This consisted photo uploading and tagging related data. Jaiku logged data user from participants Jaiku activity. This consisted written presence lines and written location tags. After the quantitative data had been gathered it was analyzed by first separating location names from ZoneTag tags, and categorizing the location tags from ZoneTag and Jaiku data based on their specificity. The categorization was made to every group separately. The location tags were categorized to groups 'larger than or equal to city', 'more specific than city but larger than or equal to neighborhood' and 'more specific than neighborhood'.

3.3 Studied location aware mobile tagging systems

3.3.1 Jaiku

Jaiku is a presence network service. It can also be called as microblogging service or a social networking service. With Jaiku user is able to share information about him/herself publicly to anyone or privately only to his/hers own Jaiku contacts. The information that can be shared consists of:

- Presence line
 - Presence line is a written message that a user can write to inform for example about his/hers presence, opinion, mood or location. Maximum length of a one presence line is 140 characters.
- Location tagging
 - Location name consists of country, city and neighborhood. Jaiku service positions the user based on his/hers cell phone's location and calculates the location using nearby cell towers. Jaiku fetches the country name automatically (e.g. Finland or USA) but cities and city neighborhoods are named by the user. When user gives a name to a city or a neighborhood s/he is in, Jaiku system saves that user created name to user's data base. When user comes to the same previously named area Jaiku automatically fetches the name of the location from the data base. The difference between city and neighborhood is that city covers a larger area than neighborhood. To user's Jaiku contacts his/hers location is show as [neighborhood], [city], [country]. User named locations are also shared to users' contacts. For example if user A names his home neighborhood as "homehood" the name is shared to user B. When B comes to visit user A Jaiku shows to all user B's contacts that B's location is "homehood".
- Rich presence, which shows presence line, location and phone context information
 - Current presence line and time from the last update
 - Last updated location and time from the last update
 - Profile indicator shows if contact's phone is in general, meeting or silent mode
 - Speakers voice level
 - Is phone's vibration on or off
 - Has the contact any current calendar events in his/hers phone
 - When did the contact use his/hers phone last time
 - How many Bluetooth devices are nearby
 - How many friends are nearby based on contacts' phone's bluetooth transmitters.

Location, calendar events and bluetooth information can be set hidden if wanted. Also Jaiku can be used only as normal contacts list by disconnecting it from the server. If Jaiku is disconnected no information about the user is transmitted to his/hers contacts but it also means that the user does not get any information updates from his/hers contacts.

Jaiku can be used from both web browser and mobile interface. In mobile interface Jaiku is used as a replace for the normal contact list. All users expect one used Jaiku version 0.32 in their phones. One test user had a phone which used Jaiku version 0.80 Beta. Main differences between the two are that in 0.80 Jaiku does not show as much phone's context information but user is able to access his/hers contacts' earlier presence lines from mobile phone. Version 0.32 does not have this history feature but history data is always accessible from web interface regardless the used mobile application version. There is also a Java version available for non Symbian phones. However, it is much plainer in features compared to Symbian version. Figure 4 shows Jaiku's mobile interface and its' basic features. Figure 5 shows Jaiku's web interface.

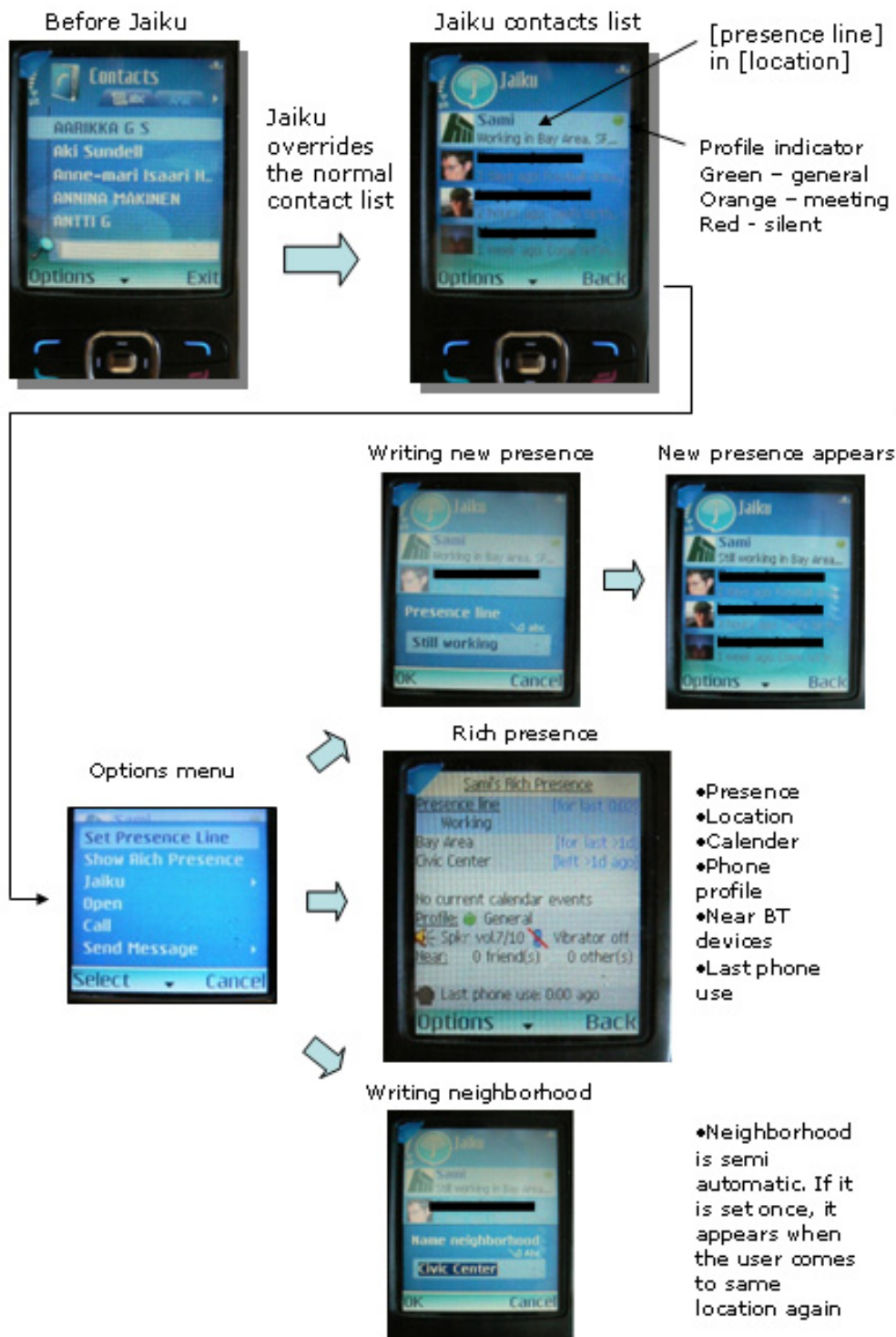


Figure 4. Jaiku's mobile interface and its basic features.

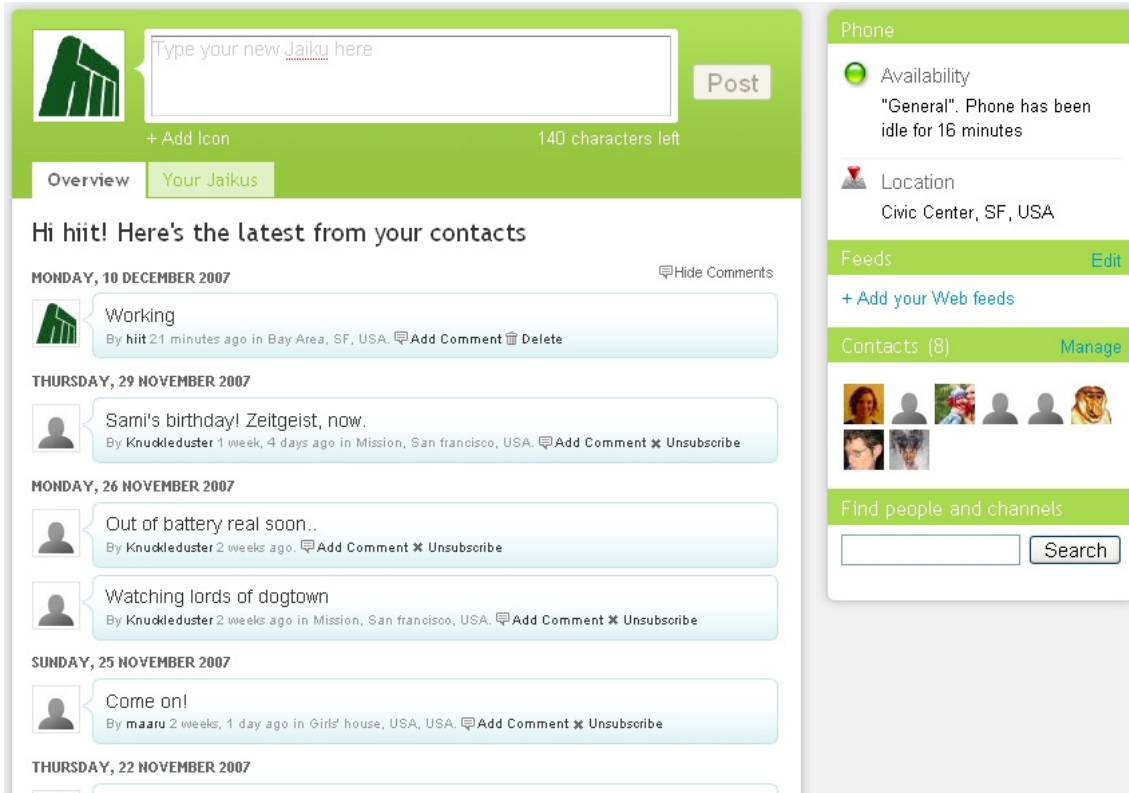


Figure 5. Jaiku web interface.

In Jaiku's web interface it is possible for a user to update his/hers presence line (write "jaikus"), see the history of own and contacts' current and historical presence lines and locations, comment contacts' presence lines, and see contacts' availability. The web interface gives the same information as the mobile interface. However, the difference is that in the web interface user can also view and comment earlier presence lines. Through the web interface user can also add web feeds like a feed for her Flickr site.

3.3.2 Zonetag/Flickr

ZoneTag is a photo uploading and annotation application for Symbian 60 phones. When a user takes a photo with her camera phone ZoneTag asks if she wants to upload the photo to her account in Flickr photo sharing service. If the user answers yes ZoneTag asks if she wants to add tags, a title or adjust privacy setting for the photo. If the user wants to add tags she can either write them manually herself or scroll suggested tags, which are tags that other users have written earlier. After putting tags, a title, and setting up the privacy settings, if any, the photo can be uploaded to Flickr by pressing Upload. Figures 6 and 7 show ZoneTag's interface and an example of a photo in Flickr.

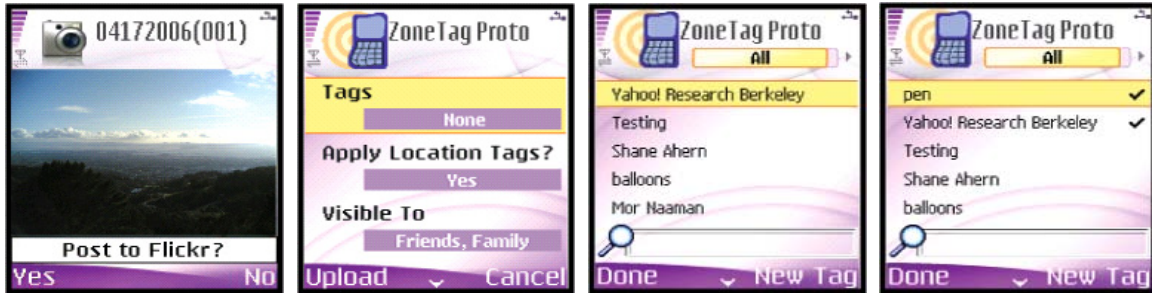


Figure 6. Different phases in ZoneTag interface.

Camera Phone Photo Saturday 4:11 pm
6/23/07 Oakland, California

ADD TO FAVORITES ADD NOTE BLOG THIS ALL SIZES

ZoneTag: [Photosphere](#) / [About](#). Owner only: [Fix Location](#) / [Add Tags](#) / [Settings](#)

Comments

[littlecountrygirl](#) says:
i don't believe this exists. too perfect, too sunny, too happy.
Posted 2 months ago. ([permlink](#))

[photophobia pro](#) says:
the palm trees out there always freak me out. oakland's looking good though.
Posted 2 months ago. ([permlink](#))

[electric heart pro](#) says:
So this new phone I'm using has something called ZoneTag on it... ZoneTag automatically uploads photos from your phone to Flickr. It figures out when and where the picture was taken and adds the "appropriate" tags to the photos. Hence the Oakland tag...I'll have to complain to ZoneTag. DP is still in SF!!! And it's still the place for fun times! And it misses you two!
Posted 2 months ago. ([permlink](#))

Uploaded on June 23, 2007
by [electric heart](#)

electric heart's photostream
367 photos
[browse](#)

This photo also belongs to:

Pink Saturday 2007 (Set)
You are at the first photo.
6 photos
[browse](#)

Tags

- zonetag
- cameraphone
- Oakland
- California
- USA
- celltagged

Show machine tags (1)

[Add a tag](#)

Additional information

© All rights reserved

- Taken with a Nokia N70-1. [More properties](#)
- Taken on June 23, 2007
- See [different sizes](#)
- Viewed 20 times

This photo is public.

[Flag this photo](#)

Figure 7. An example of a photo in the Flickr photo sharing service.

3.3.3 User groups

3.3.3.1 High school students (Finland)

The high school group contained a group of 10 friends, 5 males, and 5 females, age 18 to 20 years. One of the boys was a good friend of one of the girls and knew the rest of the group pretty much only through her. Otherwise the group was quite close-knit and they all were in their senior year in a same high school. They spend time together in the school and sometimes also on free time. During the ZoneTag use period they had several joint events, like a ferry trip to Stockholm. During Jaiku period most of them were studying for entrance exams to Universities. During that time they did not meet face to face very much but most of them went to the same preparation course. They all lived in Greater-Helsinki area, with their parents.

Participants were not particularly tech savvy but were accustomed to use today's largely diffused new communication technologies such as cell phones and internet. For communication with each other their main channels were cell phone calls, SMS messages and instant messaging. 8 out of 10 did not have a camera phone beforehand.

3.3.3.2 Interest group of birdwatchers (Finland)

Birdwatchers were a group of seven men and one woman recruited from the members of a bird watching association in Helsinki area. They were aged between 18 to 64 years. As a group the interest to bird watching and belonging to the same bird watching club was the only common nominator. Four of the participants knew each other quite well beforehand, but the other four had not met other participants before the study.

During the study many of them did a lot of travelling mainly in southern Finland. Some of the participants also took photos of birds for their profession. Most of the communication within the group was bird watching related. They also used actively a nation wide bird observation service (<http://www.santamargarita.fi/lintutiedotus>) which made it possible to receive instant information about the recent bird observations.

3.3.3.3 Group of urban adults (USA)

USA group was a group of nine friends, "urban adults", in their 30's living in Bay Area, California. Seven of them lived in San Francisco and two, a couple with a baby, in Santa Rosa, about a hundred kilometers from San Francisco. Four of them were men and five were women. They had various backgrounds and most of them were not originally from Bay Area but somewhere else from US. One of the men was originally

from Finland but had been living in San Francisco for four years. They were not particularly tech savvy except the Finnish man, who was a software engineer, and a man living in Santa Rosa, who was a radio journalist, movie director, and musician. However, all the participants were fluent in using new communication technologies such as cell phones, instant messaging, SMS, and internet. Many of them for example were very used to internet shopping and ordered things like clothes or household equipments from internet, for example from Ebay. Other occupations in the group were clothing designer, fine artist, graphical designer, elementary school teacher, social worker, and home mom.

Everyone from the group was quite close friends with each other. Three of the girls lived together and one of them was dating with the Finnish software engineer. They did not have any particular common interests but shared a same kind of ‘hipster’ life styles and spent time together from time to time. For example all of them were invited to spend Thanksgiving (which is an important family holiday) together to the three girls’ house. However, all of them had their independent, ‘adult’, lives. The couple living in Santa Rosa was the most separate from the group. Others saw them mainly in larger parties, such as the Thanksgiving.

3.4 Results

3.4.1 Adoption of Jaiku among the user groups

Jaiku was a new concept to everyone. There was a clear learning phase with all groups to find the proper ways to use the system. Even though the norms of use took some time to form, the basic concept of Jaiku, a service that enables one to communicate with his/hers friends what s/he is doing and also what is his/hers location, was still quite easy for all to understand. Thus even Jaiku was a totally new concept for everyone to use for communication they felt it was pretty straight forward. However there were issues related for example to battery consumption, privacy, and the accuracy of the location technology that reduced the use for some participants

Table 1 shows the Jaiku activity of the groups. During the use period 25 participants sent 1424 Jaiku messages and wrote 225 location tags. However there were large differences in usage activity between the groups. The group of high school friends was the most active group sending 821 messages and tagging 94 location names. The interest group of birdwatchers was also quite active sending 550 messages and tagging 122 locations. However the group of urban adults were very passive in comparison to the other groups and sent only 53 messages and tagged 9 locations.

Table 1. Jaiku activity among the groups.

	H'schoolers	Birdwatchers	Urban adults/USA	TOTAL
Number of users	9	8	8	25
Use period	Apr 07 – Jun 07	Apr 07 – Jul 07	Oct 07 – Dec 07	
Jaiku messages sent	821	550	53	1424
Locations named	94	122	9	225

High school friends

The high school group was overall quite keen in using Jaiku. During the first month they used it mostly for sending messages to the whole group. The messages were mostly insider comments, and telling to the others what the writer is currently doing or where he or she currently is. During the second month a norm to use Jaiku also for person to person communication was formed. That was mostly because they had become more familiar with the system and were exploring the limits of its proper use. Starting to use Jaiku for person to person communication also increased the overall activity. Figures 8 and 9 are examples of Jaiku message feeds from the high school students.

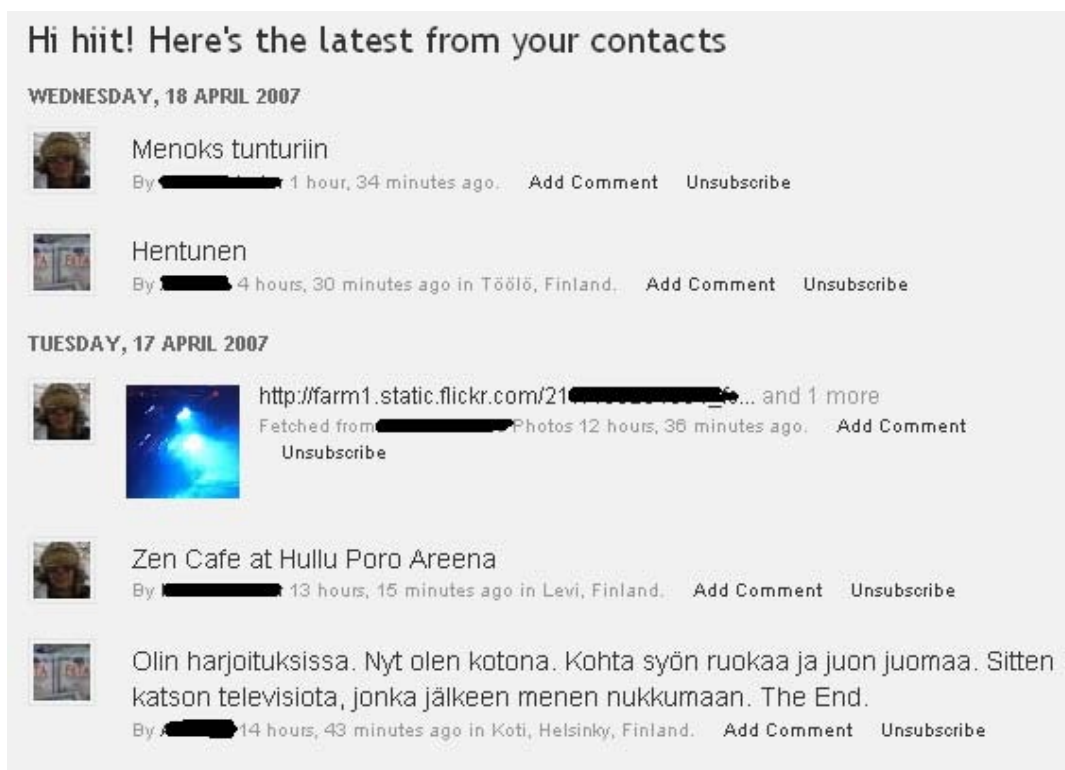


Figure 8. An example from the first month use of Jaiku among the high school group.



Figure 9. An example from the second month use of Jaiku among the high school group.

Birdwatchers

In the beginning of the study the bird watchers used Jaiku mainly for reporting the birds they had seen. However, later they started to write more and more of the bird topic issues. Overall Jaiku was not very useful for them. This was mostly because bird watching was the only common nominator between them and Jaiku was not very efficient for reporting bird sightings. It was not as efficient to that purpose as their existing SMS system for bird sightings. Figure 10 gives an example of the bird watcher's Jaiku feed.



Figure 10. An example from use of Jaiku among the bird watchers.

Urban adults / USA

Urban adults group was not very active in Jaiku. Especially in the beginning of the test period many of the participants had some concerns about the privacy in Jaiku. In the session where they were introduced to Jaiku some said that it was “creepy” that they are being automatically tracked all the time. Also some felt that it was unnecessary to be connected more than they did with their emailing, phone calls, and text messaging before Jaiku.

“The idea of always being connected annoys me” – F30

Over all Jaiku had little effect on their communication and it seemed that it did not change their well established ways to communicate with each other. That was mainly because their mutual, distant communication consisted mostly of organizing everyday things and inquiring after daily happenings. For them other channels, like SMS and

email, were more efficient than Jaiku for those tasks. One of the participants posted messages quite frequently and tried to get the others to be more active. He got one other to be more active but the others did not join the conversation. Many of them also had not used to charging their phones every night. Because Jaiku consumes quite a lot cell phone battery some of the participants did not keep it on all the time.

Also for the couple living in Santa Rosa one reason for their lack of use was that they were too separate from the rest of the group and also they got separated during the study. The woman from the couple also said “recently I’ve been trying to be much less connected to digital world than what I have been before” She told she had been using Myspace, Flickr and email almost daily and felt she was almost addicted to them. But now when her life changed because of the divorce it was kind of a new start. She said she wanted to spend more time in activities like reading a book and being in nature and not being connected to digital world all the time. Figure 11. shows an example of the urban adults’ Jaiku feed.



Figure 11. An example from use of Jaiku among the urban adults.

3.4.2 Location tagging in Jaiku

Many of the participants from the high school group and bird watchers group tagged location frequently. Especially in the high school group participants enjoyed watching other participants’ locations and by that felt obligated to tag their own locations too. They also had positive experiences related to location feature. For example once a person called the other group member after seeing, through Jaiku, that they were at the same location.

Bird watchers did not use the location feature for their benefit and there was not incidents where they would have needed to know who from the group is at some specific bird watchers location. However they did name locations quite frequently, 122 all together. Many of the names were bird watching places or names of little villages around Southern Finland. Some of them thought that the amount of users using the system should be much larger, ~50, in order to be useful for their interest group.

3.4.3 Privacy in location awareness

Some people reported privacy concerns and some did not. The ones that had some concerns were mostly from the urban adults group from US but when asked if one would share the location information with other people than inside the study group some participants also from high school friends and birdwatchers had some concerns.

In the US group especially some of the first reactions towards the feature that makes it possible to “follow your friends” were quite negative.

“I can’t lie anymore” – F29

“This feels creepy” – F30

“Initially, actually because I didn’t use it for first couple of weeks. I thought the whole concept was creepy.” – F29

However, for some participants after testing the Jaiku, the privacy issues didn’t bother that much anymore, and Jaiku felt “harmless”. They also realized that they do not have to be logged on all the time. *“if it did [bother] I would just log off”* – F29. Eventually in the end of the study the great majority of the participants from every group felt that it is alright to share one’s own location with a group of friends. However, four things came up when it was asked about sharing location with other people or instances.

1. In high school group there was a quite lot rejection towards the idea of sharing one’s location with parents.

“Using this [Jaiku] with my parents would not be good at all. When my parents are sleeping and it [Jaiku] says “Studio51” [a name of a night club] ...that would not be nice at all” – F18/high school group

“I would not use this [location feature] with my parents...or with my boyfriend. With friends I could use this” – F19/high school group

2. Sharing location with the general public or with all the contacts on the phone was not seen useful or appropriate.

“Someone could check when I’m not home and break in to my house” – F29/US group

“With these people it is OK, they are like friends. But of course if everyone had it [location feature] on their phones” – M18/ high school group

“The concept of it [Jaiku] seems creepy. Like that someone would want to know where I was all the time and have the ability to check that. But my own group of friends...I don’t feel that short of threat of invasion of privacy” – F29/US group

3. Temporal privacy dimension in location sharing. Some said how it’s alright to share location information in daytime but night time might be different. Also some people just turned off their phones for a night time as they had done before the Jaiku too.

“Before this study I have not been able to see where my friends are. It has been a little bit weird. It’s also a borderline case to see where your friends sleeps” – F18

“I may have switched it off sometime but I don’t have any places which would not be appropriate for others to know. It is like with a cell phone normally is, it will be switched off when one goes to sleep.” – M33/bird watchers

4. Concerns about giving context data to the government or to large corporations came up several times in the US group.

“I have sort of ethical concerns about its use by the company who owns it [Google]. They have accesses to this information about us, where we are, what are we doing. It makes you uncomfortable to be tracked. Especially if it’s going to be used for marketing purposes or federal government keeping track on their citizens” – F30/US group

“We have been coming a surveillance society in this country and the ties between military industrial complex, police state and corporate world are becoming tighter and tighter. I don’t know where Google stops and NSA starts. All that data is easily used in other purposes” – M33/US group

It was disturbing when Patriot Act went on” “The patriot act was when a lot of privacy things were compromised. And then it just goes further and further” – F30/US group.

3.4.4 Specificity of location tagging

It seemed that regardless of the technical boundaries in GSM based location technology people still want to use more specific location than the technology can currently provide and what the user interface promotes. In Jaiku the regions to name are ‘Neighborhood’ and ‘City’. Thus when one wants to name a location in Jaiku she selects either ‘name neighborhood’ or ‘name city’. Regardless of the technical boundaries (and privacy concerns) from all 225 location names that were named by the participant 31% were categorized to be more specific than neighborhood. These kind of tags were for example: ‘home’, ‘coffee shop’, ‘library’). Table 2 shows the specificity of written and used location tags in Jaiku

Table 2. Specificity of written and used location tags in Jaiku.

	H'schoolers	Urban adults/USA	Birdwatchers	TOTAL
Number of users	9	8	8	25
Use period	Apr 07 – Jun 07	Oct 07 – Dec 07	Apr 07 – Jul 07	
Jaiku messages sent	821	53	550	1424
Locations named	94	9	122	225
Level of location specificity [%]				
location < city	70 %	60 %	70 %	70 %
location < neighborhood	45 %	30 %	26 %	34 %
Level of Location specificity of sent Jaikus				
location < neighborhood	40 %	17 %	18 %	31 %

The previous suggests that there is a demand to name locations specifically. The same thing came out during the interviews. Many users did notice the restrictions of the location technology. For example 29 M from USA group told how he first named locations quite accurately. However, after noticing how Jaiku showed same location for him regardless of him being at work or eating lunch a couple of blocks from work he started to name locations using larger areas, like neighborhoods.

Bird watchers were often on the move around Southern Finland. They did name locations often but the names were not as accurate as with high school group. From all of their Jaiku messages only 18% had a location tag more specific than neighborhood

compared to 40% of the high school group. During the Jaiku study the high school participants wrote a lot of Jaiku messages from places like 'home', 'library'.

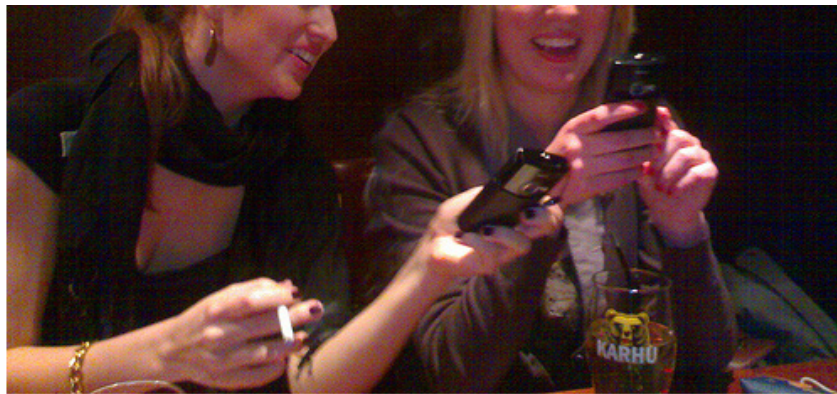
Overall the user's intuition was clearly to use more specific names than neighbourhood districts (which Jaiku's interface promotes) especially for places they spend much time.

3.4.5 Adoption of ZoneTag among the user groups

Like Jaiku also ZoneTag was given to participants for at least two months. Sending photos from a camera phone to Internet was a new concept for every participant. Most of the participants did not even have a cameraphone. Few participants from the birdwatchers group and US group did have a Flickr account before the trial and almost all had previous experience from sharing photographs in some Internet's social media service, for example in IRC-galleria or in MySpace.

Generally ZoneTag was loved among the participants. Participants liked the concept of uploading photos instantly from their cameraphones to their photo sharing account in Flickr. Many of them also continued using ZoneTag after the official ZoneTag period. However, especially in the bird watcher group some participants did take lot of photos but did not share them on Flickr simply because they had no need to share them with other bird watchers they did not even know that well. Also when asked if the participants would continue using ZoneTag after they have to pay the data plan fee themselves, great majority of the participants said no.

Some participants used Flickr and thus also ZoneTag more for sharing pictures to friends and some more as a depository for their cameraphone photos. Many said how ZoneTag made it easy to get photos out of the phone and how some photos would not even be taken at all if they wouldn't be upload to Flickr. However, for some more enthusiastic photographer the quality of the camera phone was not enough. For example one participant from the US group wanted to change her Flickr account to another so that the camera phone photos would not 'pollute' her more professional looking Flickr account. Figure 12 shows a photo that was uploaded from a user's camera phone to Flickr.



Tags

- team puhelin
- zonetag
- cameraphone

[Add a tag](#)

Additional Information

© All rights reserved

- Taken with a [Nokia N70-1](#).
[More properties](#)
- Taken on [March 26, 2007](#)
- See [different sizes](#)
- Viewed 22 times

Only friends & family see

[Flag this photo](#)

ZoneTag: [Photosphere](#) / [About](#). Owner only: [Add Location](#) / [Add Tags](#) / [Settings](#)

Comments



[\[redacted\]](#) **pro** says:

Ohhoh mites nyt tollai :s
Posted 5 months ago. ([permalink](#))



[\[redacted\]](#) **pro** says:

teil on aika viileet puhelimet.
Posted 5 months ago. ([permalink](#))



[\[redacted\]](#) **pro** says:

pilli-kaljal maanantaina :D
Posted 5 months ago. ([permalink](#))

Figure 12. A photo uploaded from the user's camera phone to Flickr.

3.4.5.1 Tagging

During the interviews most of the participants told that they do not write tags to their photos very often or at all. In the beginning of the study many of them were actually not even familiar what it means to tag one's digital photograph. Thus the concept of tagging was not clear. Even in the end of the study some of them did not know that it is possible to click a tag on Flickr and that way find other photos that have the same tag. So it was not very clear to many of them how the tags could be even used. Many said that title is enough. Also the concept of mobile tagging did no suit for many of them because there is no time to tag a photo when one is 'on the move' with her camera phone.

"Often it [a photo] goes [is uploaded to Flickr] without a tag. If I'm in a hurry there is no time, or it does not even come into my mind" – M20/high school

However, 247 tags were written from mobile phones and 1614 photos were uploaded. Also suggested tags (tags that someone had been written earlier) were used frequently. However, in many interviews it came up how suggested tags were used accidentally and

many cases in wrong photos. That is because if a user selects a tag and uploads a photo the same tag is selected to next photo too if not unselected by the user. Table 3 shows the overall ZoneTag activity of the groups.

Table 3. Photo uploading and tagging with ZoneTag.

	H'schoolers	Urban adults/USA	Birdwatchers	TOTAL
Number of users	10	9	7	26
Use period	Feb 07 – Mar 07	Jul 07 – Aug 07	Jul 07 – Aug 07	
Photos uploaded	452	839	323	1614
Mobile tags written	159	67	21	247
Suggested mobile tags used (written once)	206	714	83	1003
Web tags written	11	13	87	111
Total number of tags used	376	794	191	1361

Even though all the groups were pretty active in uploading photos the high school group was clearly the most active in writing tags. Some times tags were written just for humorous purposes, but mostly they were said to be put there so that other people outside of the study could stumble across the photo while searching photos in Flickr.

“What ever I see in a photo I put a tag. I know that people search them and that’s one way they might come across my photos” – M33/US group

Only one participant said that he tags photos for himself for the future use. However, he did it only from the Flickr web interface because mobile tagging was too inconvenient.

”I don’t usually tag individual photos. I do tag all the photos in a same batch [when uploading photos from web interface]. Normally it’s something like ‘San Francisco’, ‘Peru’. Just for the future use. If I want to find something” – M28/US group

One participant also mentioned how there is no problem going through her and her friends photos in Flickr because there are so few of the photos. However, she did note that if there were a lot of photos it would be good if they were somehow organized

“It might be different if there were several hundred photos in Flickr. Then it would be nice if they were some organized. So that you could find them by searching” F18/high school

3.4.5.2 Specificity of location tags

Like location tags in Jaiku also in ZoneTag location tags were often quite specific. 36% of all tags that were location names referred to a more specific area or place than neighborhood. With the high school group the number was even 65%. That was mostly because of the one tag ‘Ressu’. Ressu was the name of the high school, so it was a strong common nominator for the group and often a place where photos were taken. It was written several times again but also used very often (76 times) as a suggested tag. Table 4 shows the specificity of written and used location tags in ZoneTag/Flickr.

Table 4. Specificity of written and used location tags in ZoneTag/Flickr system.

	H'schoolers	Urban adults/USA	Birdwatchers	TOTAL
Number of users	10	9	7	26
Use period	Feb 07 – Mar 07	Jul 07 – Aug 07	Jul 07 – Aug 07	
Photos sent	452	839	323	1614
Mobile tags written	159	67	21	247
Suggested mobile tags used (written once)	206	714	83	1003
Web tags written	11	13	87	111
Total number of tags used	376	794	191	1361
Location names (tags) out of written tags	40 %	28 %	16 %	30 %
Level of location specificity				
location < neighborhood	52 %	26 %	35 %	41 %
Location tags out of all used tags	46 %	34 %	25 %	36 %
Level of location specificity in all used location tags				
location < neighborhood	65 %	18 %	30 %	36 %

3.5 User generated location tags vs. automatic geographical labels

Both technologies Jaiku and ZoneTag offer a user to use some automatically named locations using GSM location technology. In Jaiku the country is automatically named (e.g. Finland or USA) but more specific names as city or neighborhood are named by the user and are thus user generated. In ZoneTag the user can adjust ZoneTag to 'guess' the location where the photo is being uploaded. In that case ZoneTag adds that location as a tag to the uploaded photo. Most accurate guess ZoneTag is able to make is a ZIP code of location. However, that is not always accurate and the ZIP code can sometimes be wrong.

With these features the technologies attempt to reduce the user's burden of adding tags and naming locations herself. However, they are general geographical labels and don't take into account the personal or social context of the photo taking event or the location shared in location awareness service. Both have their pros and cons. Automatic location tags are unambiguous and therefore understandable for also for larger audience. User generated tags on the other hand are more ambiguous to general public but offer more information to the user who shares the information and to people in her social network who share common experience with the sharer.

Between the two services there were some differences users' habits of naming locations. In jaiku the locations, especially more specific ones were also a part of communication which told the others what the user is currently doing in that location. One example of that was a preparation course that many of participants from the high school group went. It was named by the users as 'Hentunen' which referred to name of the business. When seeing this location name as someone's location other knew that he or she is most probably busy and can take phone calls. However, if the location would have been a general address it would have been unambiguous but also might not have told anyone anything about the current presence of the user.

During the ZoneTag interviews it also came up several times how the participants wanted to add the photos tags that were meaningful also to others outside the group. Even though there was variation between the participants in location naming habits it did seem that photos were tagged with more unambiguous location names than what locations were tagged in Jaiku.

3.6 Cross service location tagging

People might be fairly lazy in tagging location names in both location awareness services and media annotation. However, those who do that might have various motives for that depending on the social context but also depending on technology. To ease the burden of tagging locations in location based systems one possibility might be to use user generated tags across the services. For example if a user tags her location in Jaiku the tag would automatically be added to her ZoneTag photos that are taken in the same location. However, if tags are used across two or more service there are at least following things to consider when designing the service.

- Different privacy needs in different systems
 - It's might be alright if someone sees the location of one's photo in Flickr site in some near future, but it might not be alright that at in that instant when one tags the location to her photo, that tag also appears as her location in Jaiku and thus appears to one's contacts' phones. It might be that the user has totally different social network on these two different services and thus different privacy needs. Also appropriate accuracy for a shared location might be different. For example if one tags a location to her photo it might be alright to share the location also in location awareness service only if the specificity of the location is for example city or a neighborhood, but not if the name refers to a specific building
 - With photographs the main thing to share with others is the photo. That means that the tags written into photos are there as metadata. Thus there might not be that big privacy concerns if a location tag is written to a photo (which itself usually tells a lot about the location and event) with application like ZoneTag than if it is written to a location awareness service like Jaiku which only shows the location tag
 - The concept of Zonetag and Jaiku also differs in a way that even though ZoneTag makes photo sharing more instant compared to traditional photography, it is still a discrete reproduction of one's life. At the same time Jaiku and especially Jaiku's location awareness feature, being automatic, is much more continuous. In practice this means that sharing one's location with Jaiku is more ubiquitous than sharing photos with ZoneTag. Thus even though location tagging is an important part of both services, there is a larger potential for privacy concerns and conflicts in Jaiku. Thus it might be more inappropriate to use ZoneTag tag automatically with Jaiku that vice versa.

However, if we would assume that the privacy issues in that sense would be solved so that the user's location in Jaiku would be shared only when it's appropriate for her, then would it be useful to combine the user generated location tags written in Jaiku and written in ZoneTag?

- The quantitative data analysis of this study indicated that when people tag locations both in Jaiku and ZoneTag, they often name the locations quite accurately. In both services the level of location specificity in all used location tags was over 30%. Thus it might sound logical to combine these two tag pools that have both be generated by a same user. Based on the results of this study, yes, but with at some challenging technical and usability restrictions. First, when writing a tag in Zonetag user should also label it as a location tag so it could be automatically reused in Jaiku. That would again be another extra step for user to do, and considering the already often too heavy burden of tagging photos especially from a camera phone it might be too much. Second, if we think about using Jaiku's location tags automatically in ZoneTag it might not be that simple. The current GSM based location technology being somewhat inaccurate would most probably lead to situations where for example a tag 'home', originally written in Jaiku, would be misplaced to a photo which has been taken only near home but not at home. This inaccuracy of GSM technology may still be alright in Jaiku, because, as the user studies show, in many cases it is enough to know if a friend is just near some place (in a neighborhood), and there is necessarily no need for accurate location. However, if a location tag is put into a photo it is easy to imagine that users want the location to be right and not 'almost there'. For example, users would probably not want to automatically have a tag 'home' added automatically to a photo that is taken of a garbage can but is only near the user's home.

3.7 Discussion and conclusions

The main things that came up during the studies were:

1. People's various attitudes towards the location awareness technology
 - a. For example some thought it was quite useful to see where her friends were. However, some thought it is simply unnecessary or even annoying to "be always connected".

2. People's various uses for location awareness
 - a. For example coordinating activities or communicating about one's doings by sharing a strongly context dependent location name with friends.
3. Various attitudes towards privacy issues in location awareness
 - a. For example many teenagers would not want to disclose their location automatically for their parents but with good friends it was alright. Also there were some concerns about disclosing too much information to the federal government or large corporations.
4. Specificity of the user generated location names
 - a. Regardless that Jaiku supports and promotes naming (i.e. tagging) locations only in accuracy of a neighborhood; over 30% of named locations referred to a more specific location than that. Also location tags put to ZoneTag photos had about the same level of accuracy than locations tagged in Jaiku
5. Problems of using location tags across the services
 - a. Although the location accuracy of user tagged locations was quite similar with both systems the nature of their use is different in a way that at least automatic use of tags across systems might be problematic. For example in Jaiku it is enough to see that a person is approximately 'at home', but a tag home would not most probably be wanted to a photo that presents something totally different than person's home and is only near home.

All in all our results showed that the location tagging features in Jaiku and ZoneTag were useful in close social groups, such as friends, as means of sharing experiences. With Jaiku the usefulness came up mostly when the users felt need to be connected to friends constantly.

However, it seems that currently there are still technical issues, privacy issues, and usability issues which affect on the larger diffusion of these technologies. Namely, issues with disclosing one's location information to friends, technical issues, such as battery consumption, and usability related issues, such as burden of learning how to use these new systems. Also to promote the diffusion of these kinds of technologies they should be easily accessible to users. For example it would help to get a larger user base if the technologies would be as standard equipment already when a phone is bought. However, that might eventually mean that every one was 'always connected'. Thus because the nature of these technologies is "being always connected" the technology would have to adapt to people's complex social surroundings. Currently user's location information is shared with all her Jaiku contacts. However, it's not same if one's phone shares the owner's location to her parents, co-workers, boyfriend, or just to close

friends. It is also not same if a phone shares the owner's location to her co-workers 24/7 or just during a work day. To bridge this socio technical gap the location based tagging technologies, especially continuous location tracking technologies, should be both very flexible and easy to use so that people would find them useful also in a larger context than only within a small specific group of good friends. For example there would be a need for systems where a user can easily define different groups to whom she discloses her location and when.

As with all technology there is always the burden of learning to use the technology. With location tagging it seemed that it was not necessarily easy for everyone to imagine the possible benefits there could be for example with tagging one's photographs. With user generated tagging, and especially with mobile tagging, there is also the burden annotation. Often the user of a phone is on the move and there is no time to stop to smell roses or tag photos. One possible solution would be to use tags across multiple services. However, that might bring some new privacy issues and is not self evident from technological point of view either.

Then again, we can raise a question what is enough. Has every photo and location be tagged and named? Not necessarily. However, our results show that it enhances the communication within friends. Though the users didn't generally keep tagging their current locations or tagging their photo photographs extremely important, they still did do it. Even with the current socio technical difficulties the total number of location related tags written from users' mobile phones was nearly 500. It tells that there is definitely potential in location tagging systems.

From research point of view the main results were:

1. The appropriation of communication technologies depends not only on the given technology, or the social surroundings, or persons' individual factors but is a complexity of those all. That is the case also with mobile, location tagging technologies. Based on our results when one shares her location information with the others, the most significant factor is that the level of need for sharing location information. i.e. the need has to be more significant than the burden of naming places and loose of privacy are. This means that the technologies are most useful when the social group is well knit. In that case privacy concerns come insignificant and use of collective location tagging most useful.
2. Users who tag locations to share their location information with their social networks tend to name locations very specifically or fairly roughly depending on the familiarity of the location and the communicative motivation for the location sharing.

From business point of view the main results of this study were:

1. When designing location tagging related services that include instant sharing of location information, the complexity of people's social surroundings and the burden of learning technology should be taken into account in both the design process and the marketing process. Location based technologies being a fairly new concept for large masses there is a knowledge burden on would-be adopters and also lack of social norms of use. As concrete solutions two factors came up from the study. First, location aware systems should be already build in to the cell phone when it is bought from a shop or be otherwise very easily implemented by the user. Second, the potential users should be well informed about privacy setting and other privacy related issues. The latter is important because especially privacy concerns, whether reasonable or not, might have a crucial negative effect on the first impression and thus on willingness to use these technologies. It is also good to take into account that especially the privacy issues might be culturally dependent.

4. Game as a metadata / tag creation application

Metadata creation can only be automated to some extent; for example, computers cannot yet determine the contents of an image as easily as the contents of a book. Thus, in most cases, human involvement is mandatory for the creation of adequate metadata, but tagging and keywording can be a tedious and time-consuming task. For example, writing individual keywords for hundreds of vacation photos after transferring them from camera to computer can prove to be too much to handle. The same goes for coming up with proper tags and typing them in when uploading even a part of one's photo library to a photo sharing service, that is, if the photos have not been previously tagged. In both these cases the act of creating metadata for the photos probably requires more time and thought than the uploading or transfer process itself – maybe even more than taking the photos did. In addition, creating metadata for personal use, for example for organizational purposes, does not usually yield immediate benefits and therefore does not encourage people to do it right then and there. When the most important motivators, i.e. organizational and social incentives [3], fail, could disguising the act of creating metadata or making it a side effect of some other activity be a solution?

The use of games or playful applications to overcome the lack of motivation is one possible solution to the problem. For example, The ESP Game [1] pairs two random players to guess what the other player sees in an image. Both players are presented with the same random image from the WWW and they start typing in words that describe the image. Once both players have typed in the same word(s) they are awarded points and presented with a new image. This way the game gathers relevant keywords (i.e. the words typed in by both players) that can be used as tags in other applications. The ESP Game claims to have collected nearly 35 million labels (or keywords) since its launch in October 5, 2003. Google has licensed the method of The ESP Game for its own game, Google Image Labeler [2], designed to improve Google's image search. The basic idea is to use the players to do image recognition tasks to define the contents of an image. Presumably this most often leads to players listing concrete objects they recognize in the images, for example the words car, street, bus, as seen in the 'Taboo Words' list in the screenshot below (Figure 13). The Taboo Words in The ESP Game are words that have been labeled enough times for the image in question and are not usable any more by the players. Slightly more abstract but easily recognizable concepts, like colors and the word traffic in the screenshot, are also common. The keywords gathered this way are very impersonal and best suited for search engines like Google Image Search and not for personal use or social photo sharing.

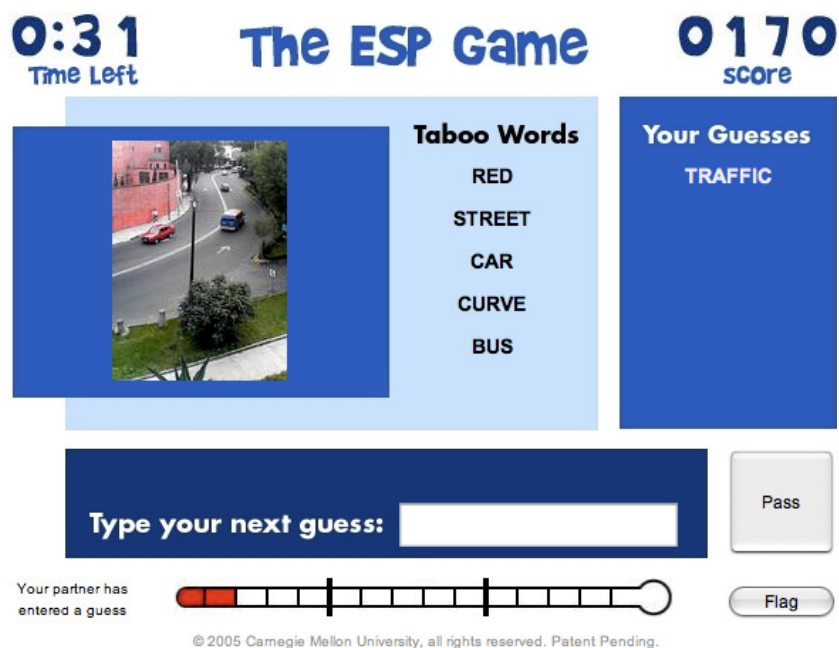


Figure 13. Screenshot from *The ESP Game* [1].

Whilst taking advantage of human computation and collaborative tagging to achieve something computers cannot achieve the aforementioned games also make an effort to make the task more appealing. Making tagging fun or disguising the act of tagging can help to overcome the lack of motivation; people playing *The ESP Game* or *Google Image Labeler* are just typing in keywords for random images shown to them, but they are doing it in a context of a game. These kinds of games or game like systems bring in competition and play (i.e. having fun) to complement or replace the organizational and social incentives. This, of course, is dependable on what kind of metadata creation the game is directed towards and how the metadata is going to be used. *Google Image Labeler* doubtfully invokes the organizational or social motivations in the average user at all as the keywords generated are going to be used for *Google's Image Search*. Instead the game is purely trying to appeal to the playful nature of its users, unless the users are also interested in helping *Google* to improve its service. If a game would allow the user to upload his/her own photos to be used in the game between his/her friends or family then it would invoke all the four incentives.

There are more examples of tagging related web-games that concentrate on creating metadata for images. *Peekaboom* [5] and *Phetch* [6] are two games from the creators of *The ESP Game*. *Peekaboom* pairs the players to identify and segment objects in images while *Phetch* is a multiplayer treasure hunt game designed to generate explanatory sentences for random images on the web [9, 10]. It is not a surprise that all of these games are about creating metadata for images, as it is the most common use for tagging. Other application areas and mediums are yet to find their use in tagging games. If GPS feature becomes more common in mobile phones, it would allow, for example, location-

based games similar to, for example, GeoCaching that could generate tags for map applications. In addition the mobile platform has not yet been utilized for tagging images using games. Mobile phone equipped with a camera allows a different approach to tagging photos: instead of random images from the Internet, photos taken by the user can be used as game elements immediately after they have been taken. A prototype of a mobile game that could be used as a tagging tool is presented in the following chapter.

4.1 CamQ – A mobile tagging game

CamQ is a mobile multiplayer game that utilizes the mobile phone's inbuilt camera. In CamQ the players compete against each other by creating and solving picture based quizzes. The quizzes are created by taking a picture using the phone's camera and giving the picture three descriptive words and an optional textual clue. Other players then try to solve the quiz by guessing the three words given by the quiz's creator. The objective for creating a quiz is to get an equal amount of correct and incorrect answers, so making obvious or impossible quizzes is not profitable.

CamQ's prototype was developed using the open-source MUPE (Multi-User Publishing Environment) platform [4]. The MUPE platform, which was originally created in Nokia Research Center, can be used for rapid development of context-aware multi-user applications, games and services that can be run on any mobile phone with Java support.

The game is persistent and asynchronous; i.e. players can access and play it anytime. There is no limitation on the number of players in the game. The only limitation is the server's capability to handle the connections. The gameplay and user interface are designed to comply with the limitations of the mobile platform.



Figure 14. Screenshots from CamQ.

CamQ was primarily designed as a game and secondarily for metadata creation. The prototype requires users to take pictures specifically for the game's purpose. For the game to be more useful for creating metadata for images it would need to be adapted so that users can use any picture from the phone's memory in the game. This way the quiz words given by the user and the answers gathered from other players could be embedded into the image file's EXIF/IPTC data and subsequently used as metadata when uploading pictures to a photo sharing service or transferred to user's own personal computer. Essentially this is a similar approach as in the The ESP Game and Google Image Labeler, but implemented for the mobile platform and to be used with user's own pictures instead of random images on the web. The metadata for the images can be gathered from the collected answers; only the most frequently entered answers can be used if there is a need to restrict the amount of gathered tags. Uploading the picture to the game's server could even be included in the basic photo taking functionality as an optional step.

The final prototype was tested using two different approaches. The first was a test run of one week with five players. The objective was to evaluate the game's appeal to players. Their reaction turned out to be mostly indifferent and uninterested – while the player's impression of the game was likely affected by technical problems, the test players were of the opinion that the game was “too simple”, “not enough gamelike” and “hard to grasp”. When asked about using the game as a tagging tool the players were worried about privacy issues, if used to tag personal material. The second test was an event-based game; the game was open for play over a weekend, but most of the playing was supposed to happen over a 7-hour time span on the day after the weekend. The test's objective was to gather sample data (photos and keywords) to evaluate how the game is used this kind of setting. In the second test 67 players registered into the game, 25 quizzes (one image, three words + optional clue per quiz) and 297 answer words were submitted. Some examples from the second test are presented in Images 1, 2 and 3. The quiz words and the clue given with the image are shown in the image's caption. Answers gathered for the quizzes during the test are shown in Tables 5, 6 and 7.



Image 1. Suunto / sport / computer. This is not just a clock.

Table 5. Answers to the quiz in Image 1.

Answer #1:	<i>altimeter</i>	<i>trainer</i>	<i>pulse</i>
Answer #2:	<i>suunto</i>	<i>pulse</i>	<i>meter</i>
Answer #3:	<i>multimedia computer</i>	<i>pulse meter</i>	<i>wearable computer</i>
Answer #4:	<i>polar</i>	<i>suunto</i>	<i>heart rate</i>
Answer #5:	<i>suunto</i>	<i>meter</i>	
Answer #6:	<i>gps</i>	<i>kompassi</i>	<i>compass</i>



Image 2. snow / bicycle / trapped. Hemmed in.

Table 6. Answers to the quiz in Image 2.

Answer #1:	<i>snow</i>	<i>helsinki</i>	
Answer #2:	<i>snow</i>	<i>bicycle</i>	
Answer #3:	<i>bicycle</i>		
Answer #4:	<i>tampere</i>		
Answer #5:	<i>snow</i>	<i>bicycle</i>	
Answer #6:	<i>snow</i>		
Answer #7:	<i>bike</i>		
Answer #8:	<i>winter</i>	<i>snow</i>	<i>street</i>



Image 3. culture / museum / trashfood. The sign and the reality.

Table 7. Answers to the quiz in Image 3.

Answer #1:	<i>jäätelö</i>	<i>pehmis</i>	<i>icecream</i>
Answer #2:	<i>mall</i>	<i>escalator</i>	<i>metro</i>
Answer #3:	<i>kamppi</i>	<i>escalator</i>	
Answer #4:	<i>helsinki</i>	<i>kamppi</i>	<i>shopping</i>
Answer #5:	<i>hard</i>	<i>cold</i>	
Answer #6:	<i>pehmis</i>		

The essential difference in the tests, in addition to the different size of the test groups, was that in the second test the users were colleagues and at least somewhat familiar with each other. This probably led to a more relaxed and open attitude towards creating the quizzes and answering them – joking, displays of opinions and playful attitude etc. was observed. The kind of social interaction noticeable in the second test is similar to what is commonly noticed around mobile phone pictures [7]. As with The ESP Game and Google Image Labeler a game between strangers leads to impersonal and non-emotional

metadata that is best suited for formal and public use, whereas a game with a closed group of people familiar with each other could be better suited for creating metadata for personal use and for sharing within the group, for example for photos from a particular event where all the players were present. In this latter case the game could either be played during an event or after it. The user interviews made during the first test support the concept of using the game as a scenario- or event-based game with a smaller group of players rather than as a persistent massively multiplayer game. Making the game closed for a group of players makes the game more socially attractive and encourages people to participate by peer pressure. This might well be the preferred way of playing using a mobile phone, which, after all, is a device primarily designed for communication and social interaction not for playing games [8].

4.2 Conclusions

A mobile multiplayer game that can be utilized for metadata creation for photos was tested in two different settings. The results suggest that mobile tagging games could be considered more as playful social applications than actual games and they seem to bring out similar behavior, e.g. joking, storytelling and playful displays of opinions, that is noticed around mobile snapshot photography. The games for PC studied in this research do not have significant social elements. In other words they concentrate purely on the gameplay between two players unknown to each other and do not provide any means of communication between players. This correlates with the intended end-use for the created metadata, i.e. metadata for a large-scale image search. Including social elements, e.g. means of communication or the possibility to choose your playing partner, could make the created metadata less neutral and more personal which would not suit the intended use in these cases. As an opposite example, a tagging game for PC that utilizes social connections and encourages social behavior could be built around Facebook and people's uploaded photos as Facebook has established itself as a prominent platform for social casual gaming.

Google's Image Labeler shows that games can effectively be used to overcome some of the motivational problems around the creation of metadata; however, creating a good game is not an easy design task and creating a good tagging game requires even more thought. The nature of created tags and how they can be utilized are affected by the game's design choices; the games need to be designed to produce the kind of metadata needed by the target application. For example, if the objective is to generate metadata for general search purposes, such as Google Images, the game needs to emphasize the creation of the most commonly used keywords. In Google Image Labeler this is implemented by the basic concept of using two players to find the same keyword without communicating to each other. Using a time limit for finding the words further

encourages the player to type in the things that first come into his/her mind when seeing the picture. Lastly, leaving any kind of social contact out of the game minimizes the kind behavior that could affect the general nature of the created metadata. The other major design factors are the platform and the games length. Persistent and long lasting tagging games are better suited for the PC or cross platform (PC and mobile) environment. The games for mobile phones only should be short, e.g. event or session based, and designed to be used in a social context.

Instead of using a full-blown game for metadata creation, some game like elements could be incorporated into other applications to make them more entertaining to use or to encourage users to compete in creation of metadata. A straightforward example would be using some kind of a ranking system for the users to show who has created the most tags or the best tags, if there were a way for the users to rate each others' tagging. In some cases rewarding users with virtual or real-world prizes for their achievements could help to motivate the creation of metadata. This could be as little as a trophy graphic or a title next to their name signifying, for example, a rank of a "Master Tagger". However, it should be taken into account that quantity does not stand for quality when talking about metadata; encouraging the user to create more metadata could degrade the quality to the level that it becomes a hindrance.

To summarize, making tagging more entertaining by including game elements in design is a concept that is worth considering in applications that require or benefit from the use of tags or keywords. However, there are important design choices that should not be overlooked; they have significant effect on the usefulness of the created metadata and on the appeal of the game or game-like elements.

4.3 References

- [1] The ESP Game. <http://www.espgame.org/>
- [2] Google Image Labeler. <http://images.google.com/imagelabeler/>
- [3] Marlow, C., Naaman, M., Boyd, D., and Davis, M. (2006). Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead. WWW2006, 2006. Retrieved from: <http://www.danah.org/papers/WWW2006.pdf>
- [4] MUPE. <http://www.mupe.net/>
- [5] Peekaboom. <http://www.peekaboom.org/>

- [6] Phetch, <http://www.peekaboom.org/phetch/>
- [7] Sarvas, R., Viikari, M., Pesonen, J., and Nevanlinna, H. (2004). MobShare: Controlled and Immediate Sharing of Mobile Images. In: The Proceedings of ACM Multimedia 2004, New York, NY, USA. ACM Press 2004. Retrieved from: http://www.cs.hut.fi/%7Ersarvas/publications/Sarvas_MobShare.pdf
- [8] Turpeinen, M., Sarvas, R., and Herrera, F. (2005). It is a phone not a console! Game Developers Conference 2005, San Francisco, CA, USA. Retrieved from: http://pong.hiit.fi/dcc/papers/Turpeinen_GDC2005.pdf
- [9] von Ahn, L., Ginosar, S., Kedia, M., Liu, R., and Blum, M. (2006). Improving accessibility of the web with a computer game. ACM Conference on Human Factors in Computing Systems (CHI), 2006. Retrieved from: <http://portal.acm.org/citation.cfm?id=1124772.1124785>
- [10] von Ahn, L., Liu, R., and Blum, M. (2006). Peekaboom: a game for locating objects in images. ACM Conference on Human Factors in Computing Systems (CHI), 2006. Retrieved from: <http://portal.acm.org/citation.cfm?id=1124772.1124782>

5. Utilising tags in an online encyclopedia

5.1 Introduction

Internet users have learnt to be active consumers and producers in social media services. They want to express themselves, participate in discussions and be connected to like-minded people. This can be seen both as a challenge and an opportunity for commercial online publishers. The publishers have to consider what kind of social media features users expect to find in the commercial services, and how users can be attracted to participate and stay involved in the service. Since there is a lot of free content and information available on the internet, innovative business models are also needed.

We studied this challenge by using a commercial online encyclopedia as a case example. The research questions were: Will social features increase the potential for a commercial service to stay and become more attractive to users, and how will new social media features and tagging in particular fit into a commercial service concept.

The analyses were made both from the user and service provider points of view. In the service provider side, it was studied, what the expectations are, and what kind of features they are willing to support. The service provider had already received signals from the field that users expect more social media features in the service. Possible features were evaluated with questionnaires, interviews and tests with users. Different features and opportunities for tagging were studied in particular.

The case study consisted of four steps:

1. Background study analysing the commercial online encyclopedia
2. The service provider view: Workshop with the service provider and interviews with the editors.
3. User view: Online questionnaire on the homepage of the encyclopedia, and user interviews.
4. Testing new features: user tests for tagging encyclopedia content in del.icio.us.

In this chapter we explain the process and the results of the case study. In making the conclusions, we utilise a more general framework that can be used for analyzing the focus and features of social media services. This framework was used to compare the features of Wikipedia and a commercial online encyclopedia. Also present and potential future features of the commercial encyclopedia were compared with the help of the framework.

5.2 Background study

As the starting point for the analysis of the online encyclopedia we used features listed in [6] for analysing tagging systems. We applied it in a wider sense and analyzed also opportunities for other types of user participation.

The service provider decides what users are able to do and the decisions are influenced by what service provider itself would like to gain with tagging and other social media features. The service provider may want additional benefits, like increasing the number of links within the service in addition to making the service more attractive and useful. The key aspect is to think about incentives driving participation. In the encyclopedia, potential incentives include future retrieval, contribution and sharing, attracting attention, playing and competing, self-presentation and expressing opinions.

There should also be clear benefits from tags, which could be made by users, by the service provider or by some specialist. Tags could provide an alternative way to organize content, and they could also be utilized for linking to additional sources: content within the service, other content on the internet, users' own content or some other content by the publisher.

Several opportunities were identified for increasing users' motivation to contributions and sharing, such as the opportunity to create timelines, map-views and themes with the help of user-created input. The service already contains support for user created notebooks and bookmarks, but these are visible only to the user him- or herself, and cannot be shared with other users. Users could also be encouraged to comment, rate, and discuss the encyclopedia articles. They could, for example, indicate topics that they would like to know more about. One opportunity would be to let users upload their own content, for example photos relating to geographical places or natural science. Gaining reputation as specialist in a topic could serve as a way of attracting.

These are all features that people have become used to expect in social media services. In the context of commercial online services, their inclusion is not self-evident. For example, the combination of community features within a pay-for-use service is not trivial. We studied what the user expectations were relating to new potential service features and which features they really would like to have. Also the expectations, ideas and attitudes of the service provider were studied.

5.3 Service provider view

The service provider view to the service was gathered in a workshop and in three interviews with editors.

5.3.1 Workshop

The analysis of the present features of the encyclopedia and the opportunities for future development presented in the previous chapter were used as the background information for the workshop with the service provider. The workshop participants had different backgrounds and roles, such as editors, technical developers and management. The key questions of the workshop were what kind of users and usage the encyclopedia currently has, and how user participation and user experience could be supported in the future.

The most active user groups of the commercial encyclopedia are educational institutes from elementary schools to universities, public administration and companies.

The strength of the commercial encyclopedia is in its edited and reliable content. This is also how the encyclopedia wants to differentiate itself from Wikipedia and other freely available web content. Therefore, the participants were concerned about the different ideas of letting users participate and “mess” with the reliable content. The editorial staff would like to check users’ content before making it publicly visible or at least the publisher provided content should be clearly separate from the user generated input. There should be the opportunity to limit the visibility of the user generated content only to a user, or to a user-defined group.

Teachers have given positive feedback on using the encyclopedia in teaching. Especially in elementary school the restricted search results are preferred over searching on the internet. In some cases “less is more”. The encyclopedia includes carefully selected links to additional information sources on the web. Anyway, it is a huge challenge to the editors to select and update the most relevant links in every subject area of the encyclopedia. One opportunity to overcome this would be to support tagging and bookmarking. This could be implemented either as a part of the service or by utilizing existing social bookmarking tools like del.icio.us. Tags could be utilized to improve browsing and linking documents, not only within the service but also elsewhere on the internet. The question remains, what would the users themselves benefit from tagging, and is it a feature that they would like to have.

The encyclopedia is not a place where people usually come to spend time. It rather is a place for searching answers to specific questions. Connecting there to other people is not what people have become used to. Also, because of the subscription fee, there is a

barrier for entry. The question is, would people find it interesting to network with other users within an encyclopedia service, and would that be something they are ready to pay for. The combination of a commercial service and social networking features is challenging. A natural starting point for creating communities would be relating to a specific subject. The question is, would the user communities be a part of the service or should the service be a part of some existing community.

5.3.2 Editor interviews

According to the service editors, the importance of editorial skills is emphasized in online encyclopedia publishing. Editors are becoming more and more like information officers, who have to find the right information and present it in a meaningful way. Careful consideration is needed to be able offer the information in right way since information retrieval on the internet differs substantially from what it is in traditional printed encyclopedias.

The editors found it important to interlink articles and facts within the service. The links are created manually, which is time-consuming, so there is room for improvement in tools for automating this task.

According to the editors the online encyclopedia should promote comparing facts. Also entertaining ways of finding and consuming information could be developed. Thematic aggregates of related articles are one possible way to this direction. Multimedia searches to selected quality sources were also regarded as potential future features.

Highlighting and covering topical themes would be useful. Editors proposed that users' suggestions could be used to select topics for more in-depth-coverage.

The editors did not believe that users could be of help in tagging and categorizing articles. The meaning of a term can be different for different people and in different contexts. User-generated tags could be confusing and they should therefore not be shown publicly.

5.4 User view

The current use cases and future needs of online encyclopedia users were studied with an online questionnaire and more detailed interviews. Additionally, we tested tagging encyclopedia articles with the help of an external service with three different user groups: active del.icio.us users, high school students and journalists.

5.4.1 Questionnaire

The online questionnaire was answered by 38 encyclopedia users. 21 of the answers came from women and 17 from men. The ages of the respondents varied from under 15 to over 60 years, but most of the respondents were between 20 and 50 years old.

Half of the respondents used the commercial encyclopedia daily or weekly, whereas 26 percent of them had just started using the service. The encyclopedia was considered to be an easy, quick and trustworthy source of information. It was mainly used as an additional source besides Wikipedia and Google.

The most wished-for features and improvements concerned the search functionality and linking the content both within the service and to external information sources. Respondents suggested that commercial encyclopedia content could be utilized in connection to web news, articles, databases, Wikipedia and dictionaries. The users would also like to have more multimedia content in the encyclopedia. However, all current users found it very important that the content in the service is reliable, and that the new features would not risk it in any way.

5.4.2 User interviews

Five of the questionnaire respondents were also interviewed. Two of them were women and three men, and their ages were between 20 and 60 years.

The usage situations for the commercial online encyclopedia varied a lot. Some interviewees use it for checking facts and spellings of words; some create exercises and quizzes for children. The ways of using were quite traditional, and the attitudes towards the online service were not much different from that to traditional printed encyclopedias. Surprisingly, the interviewees were not very interested in the suggested new features and functionalities.

Only little interest was expressed for letting users produce their own content in the encyclopedia, because it was seen as a threat to the trustworthiness of the service. Furthermore, the interviewees were not interested in producing content in a commercial service, instead if they decided to contribute, they would do it in Wikipedia. One interviewee suggested that experts should be given the opportunity to contribute by writing of articles using their own name in order to build and strengthen their reputation.

The interviewees were shown demonstrations of how the service could be enhanced in the future. The most desirable feature was customized search for additional information in pre-selected sources. Especially finding more multimedia content was regarded interesting. Pre-selecting and restricting the search sites increases the reliability and relevancy of information.

Information visualizations, such as timeline and maps, were considered interesting but not very important. Interviewees thought that these could be used in schools and for studying, but only few other use cases were mentioned. Interviewees' opinions varied relating to how entertaining the encyclopedia service should be. Some found quizzes and crosswords interesting, whereas others found them unnecessary. The interviewees assumed that playful features could help schoolchildren in comprehending the content.

Interviewees were also asked about tagging in an online encyclopedia, and they were quite sceptical about it. They believed that too many or inappropriate tags would disturb and harm the reliability, and wondered about how all the encyclopedia content could get tagged. Tags would not be useful if they existed only in some of the articles.

The interviewees did not find clear benefits in tagging. They saw some opportunities in organizing the content, but only for their own personal use. The social aspects of tagging were not found to be very interesting. According to the interviewees, the tags should at least be moderated before publishing. The possibility to limit the visibility of tags was considered very important.

5.4.3 User tests

One problem in tagging within a commercial service is that users do not want to bind themselves to one service and have their tags only there. Therefore it was tested, if tagging commercial content with the help of an external open service would work and if external tags could also be utilized within the commercial service.

del.icio.us was used as the external social bookmarking service because of its simplicity and popularity. Tagging encyclopedia articles there was tested with three different groups: five del.icio.us users, two high school classes with their teachers, and three journalists. The test consisted of an autonomous testing period of approximately a month and personal interviews after that. The del.icio.us users also answered a pre-questionnaire about their del.icio.us usage and tagging habits, and they participated in weekly blog discussions during the testing period.

The del.icio.us users were used to finding information on the Internet and did not find commercial encyclopedia particularly useful for them. We still wanted to see, what kind of tags they would create for encyclopedia articles. Our test users had different strategies for selecting tags: del.icio.us recommendations, current information retrieval need, and based on a common agreement about tags with colleagues or a standardized vocabulary. The users expressed the need for more support in selecting tags more systematically and easier and more useful tag bundles (tag bundles is the del.icio.us term for grouping tags).

The students and journalists were more natural target groups for the encyclopedia. The test usage of del.icio.us showed that especially in the school context, an external collection of links works well. It is better than an encyclopedia intern bookmark service, because information is searched from other sources as well. The whole class of students used del.icio.us with a common username and used tags to indicate the links made by each project teams.

The tags were not as useful as it was assumed. In some cases, tags gave additional information, but most of the tags only repeated the title of the article. Different languages (Finnish, English) and different ways of writing terms and concepts are also a problem in utilizing tags within the encyclopedia. The average number of tags per bookmark was relatively low: the journalists made on average 2.6 tags per bookmark, whereas students only 0.7 tags per bookmark, from which count, the student project team name tags were excluded.

The user test proved that the links between different services are essential whether the services are commercial or free. No service can offer all necessary information and therefore good links should be provided. The value of commercial encyclopedias is in thorough and well moderated articles. According to the users information should be easy to find within the service also without tags. Finding articles that other people have found worth tagging would be of interest. In this case, the tag itself would not be of interest, but the fact that somebody has taken the trouble to tag the article could indicate that there is something special and interesting in the article.

There are also challenges in collecting the encyclopedia bookmarks in an external bookmarking service. Sometimes a bookmark is not made, because the user doesn't want others to see, which information she or he is interested in. Commercial encyclopedia was compared to a private note book, whereas del.icio.us is mainly public. Also the need for a password for logging into a commercial service makes it more difficult to combine the commercial service with an external social bookmarking system: accessing the content requires typing in the password.

5.5 General framework for social media features

5.5.1 Introduction to the framework

In a previous study we proposed a general framework that gives media companies a systematic way of analyzing social media features in their own services and comparing them to those in other [1]. The framework divides the features into two main groups: Concept and system, and Content and user (See Figure 15).

In this study we used a slightly updated version of the framework to analyze the commercial subscription-based encyclopedia and Wikipedia. The updated framework is explained in the next chapters.

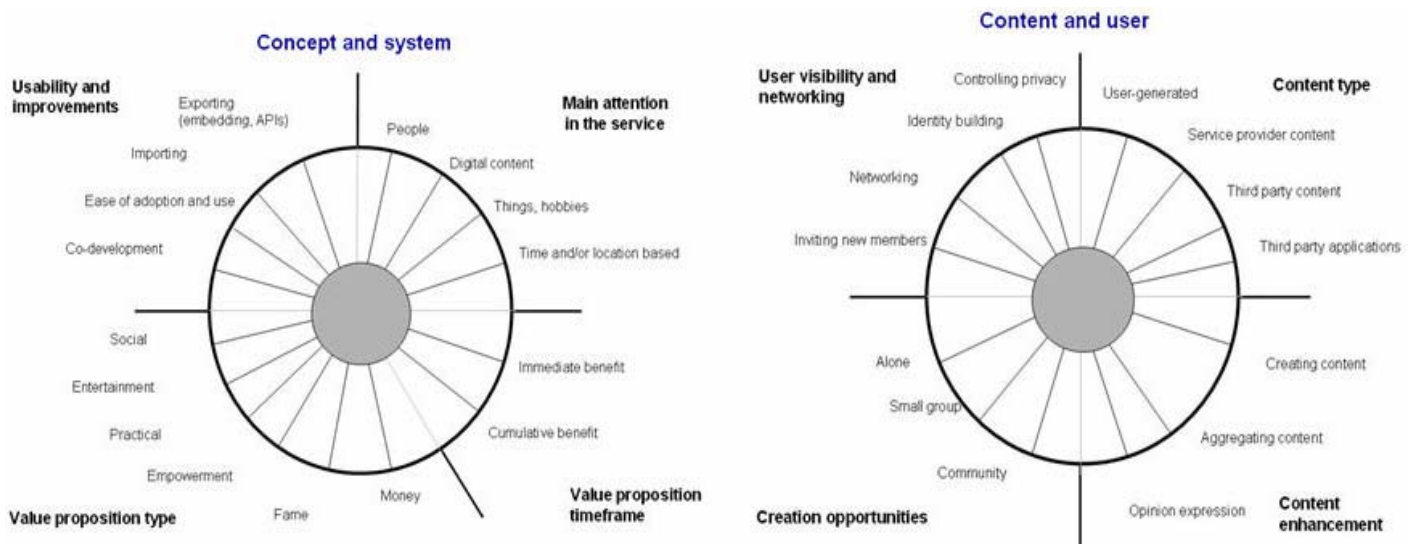


Figure 15. The framework for analysing social media features of an online media service.

5.5.1.1 Concept and system

Social media services have been developed for many different purposes and the first aspect is to identify, what the service mainly focuses at. We have identified four main focus areas: digital content, people, things & hobbies, and time and/or location related information. Social networking services like MySpace or Facebook focus on people; eBay is an example of a service that focuses on things. Different kind of calendar or event related services as well as travelling related services focus on place and time information. Many services cover more than one of these areas, especially the focus on people is important in many social media services.

The benefits for the users are analyzed in two dimensions: what kind of benefits the services offer to the users and what is the time perspective for these benefits to become evident for the user.

The benefits for users can be divided into the following sub-groups:

1. Opportunity to create and maintain social contacts and be part of the community.
2. Entertainment and exploration, enjoyable activities and/or content.
3. Practical benefits like getting opinions or advices or learning.
4. Influencing and participating, empowerment
5. Reputation and fame, where the scale varies from an appreciation within a small community to opportunities to become globally famous.
6. Gaining money directly or indirectly based on the activities in a service.

A service may offer one or more benefits to its users. The challenge of the service provider is to find the correct combination of benefits, also taking account the time perspective. It will be easier to attract new user if the benefits of the service are immediate. Cumulative benefits are important for long term commitment.

Opportunities to co-development of the service as well as the ease of adoption and use can be seen as part of the users' benefits. Many social media services are usually free to use, but some additional features may cost. Users can easily familiarize themselves with the service before deciding to pay for the extra features or functionalities. Easy importing of user's content or information could be supported in order to lower the barrier for starting to use the new service. Also when ending the use of a service, some exporting features like user's own content could be supported to relieve users' fear of "walled gardens".

5.5.1.2 Content and user

One of the key questions to decide when creating a service is what kind of content there will be in the service: is all content user-generated or is there also content from service providers or some third parties. Services may also offer opportunities for third parties to create applications for users.

Another aspect to consider is what kinds of features are offered to the users for creating and aggregating content. The interaction opportunities may vary from rating and recommending to making content aggregations like playlists and reading lists to really creating and modifying content.

Another dimension is how the co-operation with other users is supported: is creating and managing groups supported, and how the visibility of content can be controlled. Another aspect relating to visibility is how users are able to control the visibility of their user profile including personal information, social networks and content. The visible presentation of the user identity can be created utilising various sources. It can be explicitly created by a user him- or herself, or it can be created automatically based on the activities of the user, or other users can influence it. For example in eBay, buyers judge the trust worthiness of the sellers. It is important how the connections to other users are supported. Social networks may also be used in marketing the service.

5.5.2 Case: Wikipedia and the commercial online encyclopedia

Figure 16 shows the comparison of the features of Wikipedia and the commercial encyclopedia with the help of our framework. Also the present and potential future features of the encyclopedia are compared. The framework illustrates the areas where the development opportunities exist (Figure 17). The focus on utilizing the framework should not be in debating where the exact position on each axis is but to discuss, whether these different aspects should be supported in a service, and if so, what would be the best way to do it.

The biggest difference between Wikipedia and the commercial encyclopedia is that Wikipedia content is user generated and it creates communities. However, a big part of Wikipedia users are using it only for searching information rather than creating it. For some of the users, Wikipedia is an important channel to self-fulfillment and a way to participate for a common good. Important features in Wikipedia are that it is easy to use and the user can easily change roles from consumer to producer. Wikipedia also offers some opportunities for identity building. Empowerment is one of the biggest motivations of active Wikipedia producers [5].

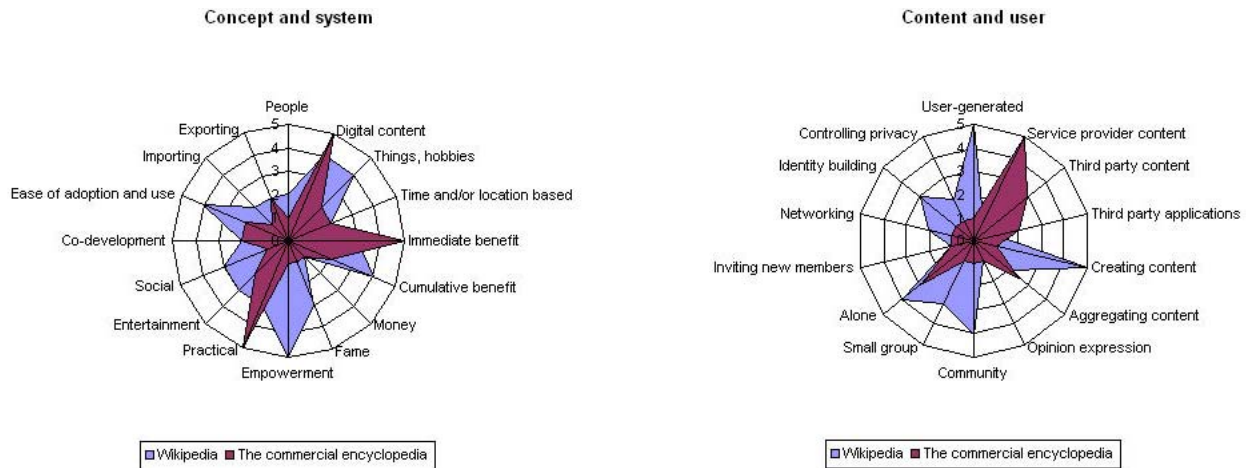


Figure 16. The differences between Wikipedia and the commercial encyclopedia illustrated with the help of our framework.

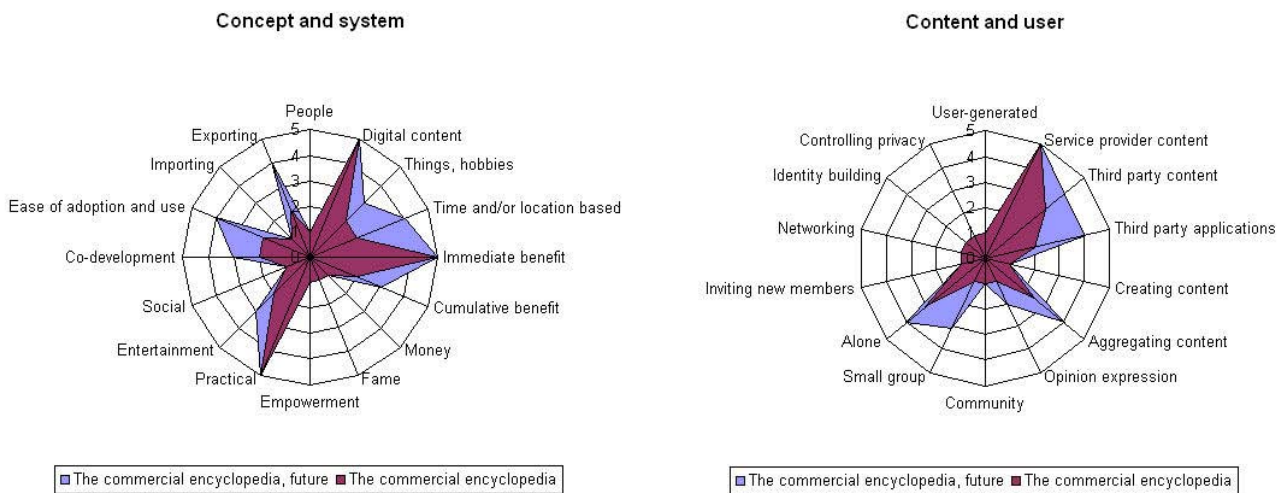


Figure 17. Present and potential future features of the commercial encyclopedia described with our framework.

The present commercial encyclopedia is used mainly for a practical purpose: to find information (Figure 17). This gives an immediate benefit to the user. Main focus is on digital content which could be enhanced to cover also time and location based information. This is one opportunity to offer more cumulative benefits to users. Although the existing encyclopedia already contains quizzes, more entertainment features could be added. Also more attention could be paid to improving the ease of use and the findability of the content in the encyclopedia. Building social and community features that would attract a lot of users are not easy to combine with a subscription-based encyclopedia and this was not what the users expect in first place either. It is more important to think about how to support users in using the information outside the encyclopedia, for example by offering different exporting features. At service provider level, one opportunity would be to offer content through APIs to the existing services and communities.

The essential question for the commercial encyclopedia is how to support combining its existing content with other content available on the internet or by the publisher itself. Supporting linking to third party content and finding the most relevant and high-quality links would help users to aggregate content from different sources. User-made tagging is one potential way to increase the number of links between related, relevant content. Third party applications could be utilized to support aggregating content in different ways and also to support entertainment features. There must be some incentive for the third party to create and offer such applications; if the incentive exists, it would make it possible to offer and test the popularity of the applications with small or no up-front investment. The encyclopedia is mostly used alone, the user can do his or her memos and bookmarks but it does not support any co-creation with other users. Better support for small groups as well as for opinion expression are potential features to be developed in the future.

5.6 Conclusions

Both users and the service provider were concerned about the idea of letting users participate and possibly mess the original trusted content. If there would be some user created content, the minimum requirement is that the user created parts be clearly marked and separate from the publisher provided content. Extra features for user activities should be optional and offered so that those who like to participate can do it, but the service can also be used in the traditional way. The new features should be hidden from those who do not want to mix original content with users' input.

The quality and reliability is highly appreciated especially by the users in school environment or public sector. Extra features could be developed specially based on their needs. For example tools that support collaboration could be offered for students doing team work.

In order to appeal to larger user groups, it would be essential to develop new innovative features for linking encyclopedia content to the web and other sources. Linking could also facilitate effective information search, as was seen when testing the idea of customized search engines. Also the content that users have tagged in external social bookmarking services could be used to suggest related content to the encyclopedia article.

The content of the commercial encyclopedia can be bookmarked for example by using del.icio.us but the problem is that the encyclopedia content can only be seen if the user is logged into the encyclopedia service.

Another opportunity is to develop features that support utilizing the content of the encyclopedia as a part of other existing services. Instead of creating communities within the encyclopedia, the encyclopedia service could be linked to be as a part of some existing communities. This means developing APIs so that content could be offered as a part of the other services or tools. If users could make tags or aggregations within the service, it would be important that they could export their tags and aggregations out of the service as people are afraid of losing the work or materials that they have created or collected.

Our framework gave a good starting point for exploring the development opportunities of the encyclopedia and also comparing its features with existing social media services. A commercial encyclopedia could complement its existing strengths with features that support easy findability, linking and aggregating content, as well as small group collaboration.

5.7 References

- [1] Bäck, A., and Vainikainen, S. 2007. Enhancing traditional media services utilising lessons learnt from successful social media applications – Case studies and framework. Elpub 2007 Conference, June 2007. Proceedings, IRIS-ISIS Publications, 2007. Pp. 371–380. ISBN: 978-3-85437-292-9.
http://elpub.scix.net/cgi-bin/works/Show?148_elpub2007
- [2] Bäck, A., Vainikainen, S., Näkki, P., Reti, T., Sarvas, R., Seppälä, L., Hietanen, H., and Turpeinen, M. 2006. Semantically supported media services with user participation. Report on the RISE-project. VTT, Espoo. 99 p. VTT Publications 612 ISBN 951-38-6857-5; 951-38-6858-3. <http://www.vtt.fi/inf/pdf/publications/2006/P612.pdf>
- [3] EIAA 2007. Shifting traditions: Internet rivalling TV in media consumption stakes. Press release 12.11.2007. <http://www.eiaa.net/news/eiaa-articles-details.asp?lang=1&id=154>
- [4] Golder, S., and Huberman, B. 2005. The Structure of Collaborative Tagging Systems. <http://arxiv.org/ftp/cs/papers/0508/0508082.pdf>
- [5] Kuznetsov, S. 2006. Motivations of Contributors to Wikipedia. ACM SIGCAS Computers and Society. Volume 36, Issue 2 (June 2006).

- [6] Marlow, C., Naaman, M., Boyd, D., and Davis, M. 2006. Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead. CollaborativeWeb TaggingWorkshop at WWW2006, Edinburgh, Scotland. <http://www.danah.org/papers/Hypertext2006.pdf>

- [7] Moore, T. D., and Serva M. A. 2007. Understanding Member Motivation for Contributing to Different Types of Virtual Communities: A Proposed Framework. SIGMIS-CRP'07, April 19–21, 2007, St. Louis, Missouri, USA.

- [8] Nov, O. 2007. What Motivates Wikipedians? Communications of the ACM. Vol. 50, No 11. (November 2007)

- [9] Xu, Z., Fu, Y., Mao, J., and Su, D. 2006. Towards the semantic web: Collaborative tag suggestions. Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland, May. http://www.ibiblio.org/www_tagging/2006/13.pdf

6. Intelligent Tagging – Case Tilkut

6.1 Introduction

Web links have traditionally been saved into browser bookmarks or favourites, copy-pasted into Word documents or sent via email to oneself as well as others who could possibly be interested in the link. These forms of information storing do not support collaborative information processing or sharing for wider and new audiences. The links are not very searchable either.

Nowadays there are several social bookmarking and clipping softwares available like del.icio.us, Connotea, Notitio.us, and Clipmarks. Bookmarks and clips are typically organized using *tags* in those services. An active tagger is quickly faced with large and messy tagclouds, which do not support finding content so well any more. The feedback from the del.icio.us user tests (see 5.4.3) also highlighted the need for better tag management. The Tilkut service was developed to support more intelligent tagging with some semantic support.

6.2 The aim of the Tilkut service

Tilkut is a prototype of an online service for collecting, organising and sharing user selected clips out of web pages. A clip consists of a URL, text and image copied from the original web page, as well as users' own remarks and tags. If the clip contains event information (date and place) they can be viewed on a map or a timeline. Clips can be made either public or private, and they may be posted to a blog as well. Public clips and events can be browsed by all users with the help of tags. The user interface of the main page can be seen in the Figure 18.

The goal of the Tilkut application development was to research, how tagging and tag management could be better supported with some new features, and to test suggestion-based semantic tagging by utilising existing semantic datasets. Still another goal was to develop methods for creating and utilising microformats. Microformats are light-weight metadata formats that use only XHTML classes and other attributes. [1] When event information (date and place) is saved according to the microformat specification, information from different sources can be combined into the same collage and shared with other people and used in other services as well.

The screenshot shows the Tilkut website interface. At the top left, the word "Tilkut" is written in a large, black, serif font. To the right of the title, it says "Logged in as 'pirjo'" and provides links for "sign out", "my account", and "home". Below the header is a "Clips" section. The first clip, titled "[1 of 4] Google's April Fools' Day Hoaxes Go International", includes a short paragraph in Finnish, a small image of the Google logo, and a list of tags: "Google", "aprillipila". The second clip, "[2 of 4] Tietoviikko: Nettijätit yhdessä Facebookin kaatoon", features a photo of Gary Hamel and a paragraph in Finnish, with tags including "MySpace", "Yahoo", "OpenSocial", "artikkeli", "Google", "Facebook", and "Tietoviikko". On the right side, a vertical menu allows users to filter clips by various criteria, with "Companies" and "Google" selected. The menu includes categories like "Advertising Age", "Apple", "Boston Dynamics", "CERN", "CIA", "CNET", "Discovery Magazine", "Economist", "Facebook", "Finnair", "Forrester", "Google", "HIIT", "HS", "hse", "IBM", "IDC", "Inside Higher Ed", "Jakob Nielsen's Alertbox", "JHS", "Kansalliskirjasto", "Karolinska Institutet Univer", "Kate Davis", and "Kiasma".

Figure 18. Tilkut home page with the most recent clips on selected topic to the left and tag categories and tags to the right.

The purpose of increased semantics is to enhance further handling and searching of clips and tags based on user's interests. This includes creating collections of clips and viewing them on a map or a timeline, as well as better browsing of clips based on folders, tag categories and tags.

One of the new features in Tilkut is the three-level tag classification including folders, categories and tags. The clips are organized into folders that the user may create by herself for different purposes. The user can select for each folder only those tag categories that she wants to use for organizing the clips and their tags. For example, a

folder called “News” might be defined to contain only the tag categories “topic”, “place” and “source”. Instead of writing all tags in one text field, the tags are added directly to the correct category, which makes it easier to handle the tags and find clips later on.

6.3 Implementation

Tilkut software consists of a web application and a Greasemonkey script for Firefox. [2] The script is used for selecting the web content that will be stored into Tilkut. When the text is selected with a mouse on a web page, a Tilkut button appears next to the selected text. When clicking the button, a new window opens with a form for saving the clip data and metadata. The title, selected text, URL and first pictures from the page are automatically picked from the page and used to prefill the form. The user may modify the information and add his or her own remarks and tags. If the script recognizes a date in the selected text, date and place information are automatically filled out as an event in a separate event form.

Tags can be added to tag categories that are topic, type, products, companies, places, project, importance, and misc. The user may select, which categories to use in each folder. When adding tags, users’ previous tags in the same folder and category are shown below each category. The user can click those tags to add them to the current clip as well. If the user starts to write a new tag, suggestions are given automatically from the General Finnish Ontology YSO and a place ontology Kulttuurisampo-paikka (in the place category). [3] The user may select to use the suggested terms by clicking the correct tag in the list. The user interface of tag categories can be seen in the Figure 19.

Folder Add/edit folders...

Topic tags (comma separated)

 Your earlier tags:
 työ, sähköposti, yso:work, yso:e-mail, yso:innovations, yso:ideas, IPR, yso:patenting

Type tags (comma separated)

 Your earlier tags: uutinen, yso:news, testi, yso:www, web page, yso:blogs

Products tags (comma separated)

 Your earlier tags: Owela, Innomanager, SMS

Companies tags (comma separated)

 Your earlier tags: Daily Mail, VTT, Valkeus

Places tags (comma separated)

Kulttuurisampo-paikka
 Helisevänjoki
 Helmelänmäki
 Helmijärvi
 Helsingin maalaiskunta
 Helsinki
 Helylä

Figure 19. Tag categories and suggestions from ontology.

The clips are stored into the Tilkut web application, from where the clips can be retrieved using tags. The event information is saved also as microformats (hCalendar, geo), which makes it possible to utilize them in other applications accepting microformats as input. The map view for the clips has been implemented using Google Maps API [4], and the timeline view with SIMILE service [5]. The timeline view can be seen in the Figure 20.

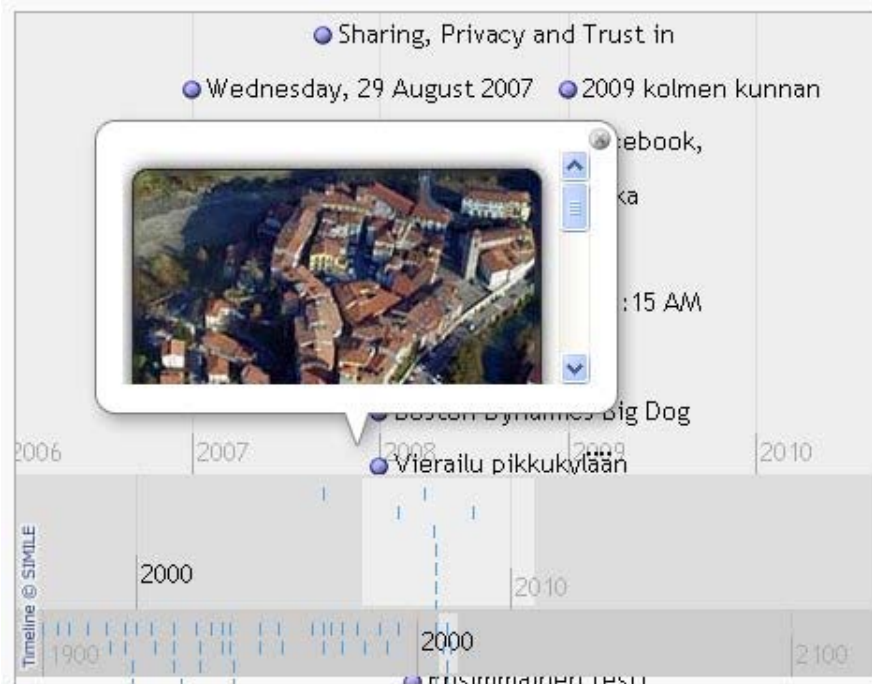


Figure 20. Timeline view of clips.

6.4 User tests

The Tilkut service was developed based on interviews with online encyclopedia and del.icio.us users and experiences from del.icio.us tests (See Chapter 5.4). In order to validate the concept as well as to get feedback and future development ideas, a small scale user study was done.

6.4.1 Users

The test users were recruited via Owela, VTT's Open Web Lab [6]. Users interested in sharing knowledge on the web and new ways of categorizing information were asked to participate in the research. No previous experience of tagging or social bookmarking was required. Initially eight users were selected, but one of them could not take part in the test and one dropped out after the initial interview because of lack of time.

Four out of the six test users were women and two men. The ages of the users varied from 25 to 60 years. All users were experienced web users, but their experiences of using bookmarking services varied from no experience or little testing to heavy using with several links per day.

6.4.2 Methods

The user test consisted of an initial interview and briefing (four face-to-face, one Skype call, and two Skype video calls), independent test period of ca. three weeks, and an end interview. At the end of the initial interviews four users were taught the basics of how to use Tilkut and three users were asked to test the service by themselves. This served a function of a light-weight usability study to test the intuitiveness of the user interface. The instructor answered test users' questions and explained the use of the features that they did not try out by themselves.

During the three weeks' test period a blog was used for communicating use experiences and problems. New questions were posted to the blog each week by the researcher, and users were asked to comment on them. Emails were used to remind the users about the test and answering the latest questions in the blog.

In the end interviews, the users were asked about their experiences with the Tilkut service, ways of using it, problems encountered during the test, and ideas for further development. During the interview, the users demonstrated how they had used the Tilkut service and what kind of clips and tags they had saved, as well as which features did not work as they had wished. Notes from the interview were made with pen and paper and later stored into Excel for analysis. The use statistics were retrieved from the database after the test period for more specific content analysis.

6.4.3 Results

6.4.3.1 Background of users

Three of the seven users had their own blog. Most of them have also tested many new social media services, although not all of them use them actively.

The users were used to saving links to web pages as bookmarks in their web browser. Almost all users had used, or at least tried out, social bookmarking like del.icio.us. Some users liked the simplicity of del.icio.us whereas for some it had seemed to be too minimalistic or too hard for everyday use. Sharing bookmarks as such was regarded quite positive.

According to the users, the quantity of bookmarks easily becomes a problem. The current services do not provide well-designed and easy ways for organizing bookmarks. Defining *good* tags can be hard even for people who do it professionally. One of the challenges is choosing a proper level of categorisation: a tag must be specific enough to be helpful for oneself but also general enough in order to be found by others.

6.4.3.2 Use statistics

During the test period 118 clips were added by the test users, which means 20 clips per person in average. The amount of clips per user varied between 10 and 30.

Tilkut was used both for work and hobby related themes, and both as a general note book and for some specific purposes. Three test users mainly used the default folder “Bookmarks” to organize their clips, whereas three users created their own folders for different themes, such as Books, Holiday trip, Blues music, and ToDo.

412 different tags were added into the eight predefined categories. Most of the tags were added into the Topic category (55 %), after which the most commonly used categories were Type (13 %), Companies (12 %) and Places (11 %). Products, Project, Importance and Misc were not very much used. The number of tags in different categories can be seen in the Table 8.

Table 8. The number of tags in different categories.

Topic	Type	Companies	Places	Products	Misc	Importance	Project	TOTAL
226	52	49	45	13	12	8	7	412

In total, 513 tags were used to describe 118 clips, which means 4.3 tags per clip on average. The real amount of tags per clip varied from 0 to 11 tags. 14 % of the tags were YSO terms. The number of different tags used by one user varied from 17 to 156.

6.4.3.3 Use experiences

The use experiences of Tilkut service were mainly positive. Most users found that the basic concept of clipping and tagging suited their needs and they could think of clear use scenarios for the service. Basic functions were easy and quick to use, although Tilkut is still at prototype state, and did not work completely correctly during the tests.

The background of users affected strongly on their way of using Tilkut. Some users used the service as a quick bookmarking tool and did not bother to add many or any tags at all. Whereas, some users wanted to have their clips and tags well organized and appreciated especially the possibility for more structured tagging in Tilkut than in current social bookmarking services.

Tilkut provides two different ways for organizing clips: folders and tags. Those users who created their own folders for different clipping purposes, found the folders very practical and handy for organizing their clips. For them, folders partly replaced the function of the tags, although tags were added as well. Quite interestingly, tag-like characteristics were wished for the folders, particularly adding one clip to multiple folders.

Since the number of clips in each folder was small, it was enough to search the clips only with the help of the folders, and tags were not needed for finding a specific clip. Because all users had different topics for their bookmarks, the benefits of tags in finding other users' bookmarks did not come out in the context of this test.

The grouping of tags into categories was not regarded very useful. Categories were too many and it was difficult to decide, into which category each tag should be put to. The burden of tagging grows, when the user has to make many decisions and remember his or her own way of using the categories last time. Only one field for tags could be shown at first, and more categories could be added later if needed.

However, some users considered it as a good thing that adding a clip requires a bit more work: tagging and categorizing serves as a check point to make sure the clip is worth saving and what it means to the user. One user supposed that the benefits of categories could probably be seen in the long run, since categories could make searching easier when the number of clips and tags grows. During the three weeks' testing period the usefulness of categories could not yet be evaluated.

One possibility to better utilize different categories would be to use different ways of tagging in different categories. Also at the moment the tag suggestions varied according to categories. It was suggested that for example places could be selected by pointing them on a map, whereas importance could be marked with stars rather than with written tags. The importance of clips also changes over time and should therefore be easily editable.

Tag suggestions from the YSO ontology divided users' opinions. Some users did not even want to see the suggestions, whereas some users would be ready to use only predefined ontology terms as their tags. The problem at the moment was that YSO suggestions did not work well in Finnish and therefore they could be utilized properly only when tagging in English.

At the moment the tag list for browsing the stored clips was not very informative and easy to use. More statistics, like popularity of tags and who has saved the same link, were suggested. Tag cloud could be better than the current alphabetical list. The service could also display the connections between tags and suggest the user some other search terms for finding similar content. Also editing tags should be made easier. Especially a mass edit mode was wished.

The users suggested several possible ways for using Tilkut. Teachers could create a list of additional study materials for students, who could participate by also adding links. Many possibilities were seen for workplace use and for example in collecting references and background information in research projects.

The group level between private and public clips was regarded very important. With the group functionality, Tilkut could serve both teams at workplace and e.g. travelling friends when planning a trip. In the group use, it could be helpful to define common tags for a shared folder already in advance. Another idea was to make a predefined tag vocabulary and use it as a to-do list for information retrieval. Here, tags with no clips would remind the user to search information also about the topics not yet covered.

6.4.4 Analysis of results

Two types of taggers were identified among the test users based on their priorities: ease of use, and precision of tagging. The first group of taggers does not want to do any extra work or spend time with categorising information. The second group values predefined vocabularies and hierarchies as well as creating rules for tagging within a team. In order to suite the different needs and purposes of these groups, the service should enable different ways of using and be modifiable. It is also worth noticing that one user may have different preferences according to the task and context.

Users appreciate tag suggestions and semi-automatic tagging although different users have very different preferences in how the suggestions should work. For example, the user should be able to select which ontologies to use for suggestions, and it should also be possible to switch off the suggestions completely. The tags at the original web page should be utilized and automatically added as tags in Tilkut as well. Tag categories could also be created automatically on a fly so that the user does not have to bother herself with categorizing tags. Collective intelligence as well as semantic web technologies could be utilized to find out tags' meaning and connections to others.

Social tagging and bookmarking at the work place will probably be a growing application area. Many users referred to the need of sharing web content easily at workplace and within a group of co-workers. In the workplace, the use of more structured tagging could be acceptable. If a team shared some clip folders, there could be support for using common tags or some relevant official vocabulary.

More structured tagging does not necessarily mean more complicated tagging. Hard restrictions can also make the service more pleasant to use. For example choosing a folder is a lot easier task than inventing good tags and organizing them into categories.

When creating a tagging service it could even be considered, if masses of tags are really needed or if a couple of folders could serve the same purpose more efficiently. However, it may be better to give users free hands to find their own ways of using tags instead of defining the use cases and purposes too strictly in advance.

The Tilkut service has potential especially in the workplace use, in the context of knowledge sharing within a team, for media reviews, and collecting weak signals. Tilkut may also be utilized as a personal notebook or as an extension for a blog. The service could also be added to many learning applications and be further developed to support group work.

6.5 Semantic Tilkut

This chapter describes how the tags and tagging related information in Tilkut could be expressed semantically by utilising existing ontologies. The purpose of adding semantics is to support better handling of tags and analysing clips and to enhance opportunities to search additional content.

The goal of the Tilkut service was to offer extra support for users to manage their tags. A three level, hierarchical categorisation was used: folders, tag categories and tags. Users can create their own folders. Folders describe the topic or the purpose why the user tags (e.g. analysing ideas, creating lecture, project). Tag categories help users to organise their tags and clips, and improves opportunities to information retrieval. The idea is that each tag has a unique URI, which will be used within the service among different users.

The ontology for describing tags in Tilkut is the combination of Tag ontology, SKOS and MOAT. SIOC is used for describing the content of clip as well as Tilkut service itself. FOAF is utilised for describing users. Also Dublin Core vocabularies are utilised.

Tag category and Category (corresponds to Folder or Purpose) classes were defined as subclasses of skos:ConceptScheme, as can be seen in Figure 21. MOAT ontology was used for describing tags, but also the properties of tags:Tag as well as skos:Concept were utilised. Since these ontologies are interlinked, this approach is possible.

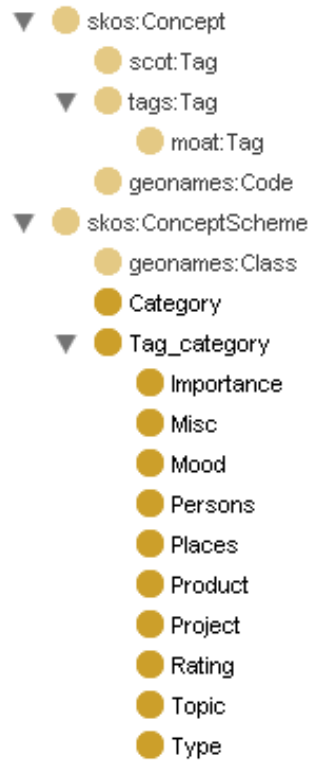


Figure 21. Category, Tag categories and Tag defined in the ontology.

The hierarchical structure of tagging in Tilkut complicates the implementation. Users might use the same tag in different tag categories. The information of which user has used which tag, in which category and folder and when needs to be stored. This information can be regarded as the context of tagging and tags: RestrictedTagging is utilised for describing it. The information is utilised when browsing the clips. The user can take different views on clips based on each user's folders and tag categories. Each Folder may have different tag categories associated since the user is able to select which tag categories to use in each of them. `dterm:has_Part` is utilised to describe the user selected tag categories in a certain folder.

Tilkut utilises YSO and KS_paikka semantic datasets for tag suggestions. Users are not forced to use suggested tags, they are able to add their tags using any words. This means that one tag may have a clear meaning expressed with help of semantic web resources in YSO or only the same label without exact knowledge about the meaning. The meaning of a tag is expressed with help of `moat:Meaning` class. The Figure 22 shows the relationships between the utilised classes.

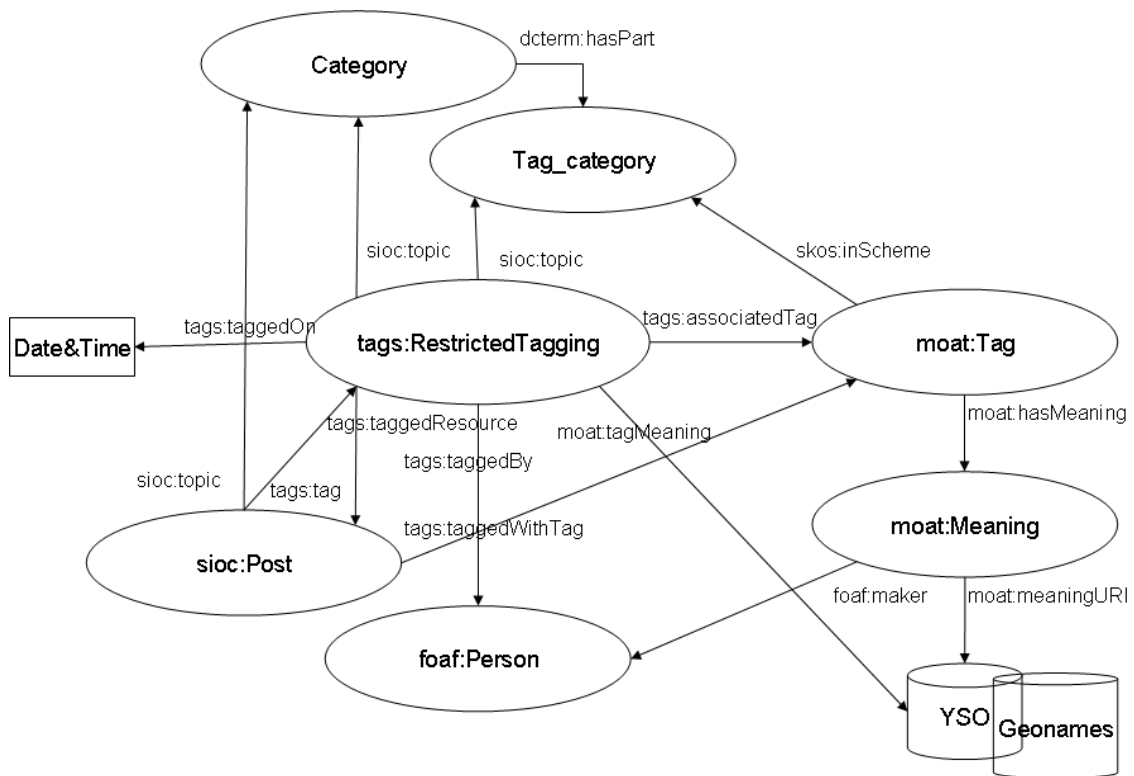


Figure 22. Interlinking Tag ontology, MOAT, SIOC and SKOS with the Category and Tag_category classes defined for the TILKUT application.

Some examples of the RDF code and SPARQL-queries can be seen in Appendices A and B.

From the user point of view, several tag categories may be confusing and not experienced as “easy tagging”. When the availability and scopes of semantic knowledge bases increases, they can be utilised more widely, also for tagging. More intelligent solutions that utilise this increased semantic knowledge can be developed. Users’ acceptance of using them will increase if they are easy to use and if the benefits of using them can be clearly seen. Also, with help of semantic knowledge bases applications may be developed further and this makes different tag categories less important. The meaning of tags and their categories could be inferred based on the semantic meaning. At least tag categories like topic, organisation and places could be inferred automatically. There are still some tag categories like importance and project which heavily depend on user’s purposes and viewpoints. Also, these are not necessarily so relevant for other users. We propose that automatic inferring and tag suggestions will be used as much as possible, but the final selection and evaluation is left for users to do.

6.6 References

- [1] Microformats. <http://microformats.org/>
- [2] Greasemonkey Firefox Add-on. <https://addons.mozilla.org/fi/firefox/addon/748>
- [3] Yleinen suomalainen ontologiapalvelu. <http://www.yso.fi/>
- [4] Google Maps API. <http://code.google.com/apis/maps/>
- [5] SIMILE Timeline. <http://simile.mit.edu/timeline/>
- [6] Owela, Open Web Lab. <http://owela.vtt.fi/>

7. Conclusion and discussion

Tags have been an important part of the Web 2.0 phenomenon. Earlier, tags and metadata were used only in work context and defined by professionals. The applications dealing with user created content needed some solution for content management, and tags turned to offer a solution to that.

This research project wanted to understand the role of tags and tagging as a part of user experience and to bring value to companies for offering media products and services. Answers to these questions were sought by making user studies both with systems that currently let users tag content as well as testing different opportunities to add tagging into an existing application (Jaiku, ZoneTag and an online encyclopedia) and by developing a new applications for tagging, Tilkut for the web and CamQ for playful mobile tagging.

Some key differences were identified between keywords and tags, and the project defined the tag as an automation of a keyword search. This means that a tag does not only describe a resource, but it also creates a dynamic link to other resources described with the same tag. Another aspect differentiating tags from keywords is the cumulative effect that is often present in applications that let people assign tags: many people tag the same resources and this makes the tags richer descriptions than keywords that are often only made by one person, or tell about what people are dealing with at different times.

Tags are created for different purposes, in different contexts, and in different kinds of applications. All these affect on what kinds of tags are being created and how eager users are to create them. Our results show that the social context is an important factor here. There are big differences in resulting tags depending on whether the tags are created only for personal use, for known friends or for public use. In some applications the visibility of tags is also restricted to these different levels. Even though tags are public, they may still be created only for personal or small group use, and are not necessarily very useful for other purposes, like adding metadata or as a user profile.

The playful mobile application for tagging showed potential, particularly when used among friends as a way to communicate and share experiences. To make tagging games attractive for large number of people is however challenging. Also, if the aim of offering the game is to get metadata for existing resources, the game should be such that users are likely to produce the kind of tags that the media house has in mind.

For most users, getting direct value out of tagging is the necessary motivator for making tags. In practice, this often links to tagging user's own content. However, the presence of other people and their tags also has the potential of increasing users' motivations, particularly if tags link to interpersonal communication.

Location tagging has most potential among smaller groups of people who communicate frequently. When used among people with a lot of common knowledge, location names may be nicknames for places, or context specific, like home, that are of little use for people outside the small group. There was also playful use or inside jokes with the help of location tags.

Location tagging overall is still very new and most users are unfamiliar with it. This sets high requirements on implementing such systems. Also, the concern about privacy came up in relation to location tagging.

In the encyclopedia case we looked at how an existing commercial service could benefit from tagging. The user and service provider interviews led us to conclude that there are challenges in combining existing services with new social media features like tagging. There are two main reasons for that. Both the service provider and users regarded the reliability of the content as key value, and were afraid that user participation might jeopardise this. Another reason is that the content is behind a password. Doing something like creating tags is not very interesting, if the tags remain within a closed system, and do not connect to and support user's other online activities; also, using an external service for tagging is somewhat problematic in this case, because opening the link requires a separate log in. For the encyclopedia service and its users, relevant links to resources elsewhere on the Internet would be more useful than service internal links that can be supported also without tags. External links could be best achieved with the help of applications that support tagging anywhere, and not only within a service.

The Tilkut application offers a more structured approach to tagging by offering different levels to tag management. Users may create folders for their different bookmark collecting purposes and select what kind of tag categories to use in the different folders. Eight predefined tag categories are offered to the users, and the user could leave out the irrelevant ones. The tag categories were defined based on published research papers where tags from systems like Flickr and del.icio.us had been analysed.

The user tests showed that people's preferences and habits vary a lot: some just want a quick and easy way of storing their links whereas others want and appreciate support for making more meaningful tags and use of predefined vocabularies. Tag suggestions and automatic features are welcomed, but users must be given opportunities to control when and how to use them.

Even though tags are used in many popular applications, it is still an unclear and unfamiliar concept to many users. This is also partly caused by the frequent use of the word with different meaning and purpose in different applications. In order to be able to benefit from using the tags to the full, the user must understand what each application offers via the use of tags.

Increased semantics will support better handling of tags and improved analysis of content, such as searching and aggregating content in different ways. Based on the semantic relations, search terms can be widened to related terms and the search need not be restricted only to the tag that the user has used. Also aggregations of content relating to the same topic can benefit from increased semantic knowledge. Also linking users with similar interests becomes more precise if the system had more knowledge about the meanings of tags.

The availability of semantic knowledge bases is increasing, which helps in creating semantic support also for tagging applications. Utilisation of existing vocabularies for tag suggestions can be implemented in a way that is easy for users to use with the help of new web interface technologies.

All in all, tags are a versatile and flexible way of letting people interact with content, system and even other users. When creating applications that utilise tags, it is important to keep in mind that people have very varying habits and needs, and successful features adapt to different use cases. Automation and suggestions should be utilised where possible, but the final decision must be left at the user. Tagging is still unfamiliar to large number of users, therefore ease of use and understanding why to tag are critical. Also, tagging cannot be the aim in itself, but a way to achieve something that the users find useful or fun.

Appendix A: Example RDFs of Semantic Tilkut

Example: A Tag and different meanings of the tag

One tag may have several meanings described by different users. Tag-category is expressed with `skos.inScheme` property.

```
<moat:Tag rdf:about="http://owela.vtt.fi/tilkut/tag1_Otaniemi">
  <tags:tagName
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Otaniemi</tags:tagName
>
  <tags:isTagOf rdf:resource="http://owela.vtt.fi/tilkut/clip_id8"/>
  <sioc:topic rdf:resource="http://sws.geonames.org/_643522"/>
  <skos:inScheme rdf:resource="http://owela.vtt.fi/tilkut/places_tags"/>
  <skos:prefLabel
rdf:datatype="http://www.w3.org/2001/XMLSchema#string"></skos:prefLabel>
  <moat:hasMeaning
rdf:resource="http://owela.vtt.fi/tilkut/tag1_Otaniemi_meaning_id2"/>
  <moat:hasMeaning rdf:resource="http://owela.vtt.fi/tilkut/tag1_Otaniemi_meaning"/>
</moat:Tag>
```

```
<moat:Meaning rdf:about="http://owela.vtt.fi/tilkut/tag1_Otaniemi_meaning_id2">
<foaf:maker rdf:resource="http://owela.vtt.fi/tilkut/Timo_foaf"/>
<moat:meaningURI rdf:resource="http://www.yso.fi/onto/yso/p10761"/>
</moat:Meaning>
```

```
<moat:Meaning rdf:about="http://owela.vtt.fi/tilkut/tag1_Otaniemi_meaning">
<moat:meaningURI rdf:resource="http://sws.geonames.org/_643522"/>
<foaf:maker rdf:resource="http://owela.vtt.fi/tilkut/pirjo_foaf"/>
</moat:Meaning>
```

Example: Tag category

```
<Places rdf:about="http://owela.vtt.fi/tilkut/places_tags">
  <skos:prefLabel xml:lang="fi">Places tags</skos:prefLabel>
  <siocservices:has_service rdf:resource="#geonames_web"/>
  <foaf:maker rdf:resource="http://owela.vtt.fi/tilkut/Timo_foaf"/>
  <sioc:has_space rdf:resource="http://owela.vtt.fi/tilkut/site"/>
</Places>
```

Example: Management of user categories (=folder) and tag categories related to the category

```
<Category rdf:about="http://owela.vtt.fi/tilkut/Bookmarks_sariv">
  <dcterms:hasPart rdf:resource="http://owela.vtt.fi/tilkut/project_tags"/>
  <dcterms:hasPart rdf:resource="http://owela.vtt.fi/tilkut/topic_tags"/>
  <sioc:has_space rdf:resource="http://owela.vtt.fi/tilkut/site"/>
  <skos:prefLabel xml:lang="fi">Bookmarks</skos:prefLabel>

  <foaf:maker rdf:resource="http://owela.vtt.fi/tilkut/author_sariv_foaf"/>
  <dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema#string">2008-
04-14T11:35:54+2:00</dcterms:created>
</Category>
```

Example: “Tagging context” – Information related to tagging from the user point of view

```
<tags:RestrictedTagging rdf:about="http://owela.vtt.fi/tilkut/tagging_tag1_Otaniemi">
  <tags:taggedOn rdf:datatype="http://www.w3.org/2001/XMLSchema#string">2008-
31-03T12:20:57+02:00</tags:taggedOn>
  <tags:taggedResource rdf:resource="http://owela.vtt.fi/tilkut/clip_id8"/>
  <tags:associatedTag rdf:resource="http://owela.vtt.fi/tilkut/tag1_Otaniemi"/>
  <sioc:topic rdf:resource="http://owela.vtt.fi/tilkut/places_tags"/>
  <sioc:topic rdf:resource="http://owela.vtt.fi/tilkut/Asuminen"/>
  <tags:taggedBy rdf:resource="http://owela.vtt.fi/tilkut/pirjo_foaf"/>
  <moat:tagMeaning rdf:resource="http://sws.geonames.org/_643522"/>
</tags:RestrictedTagging>
```

Example: Connection of Tag and Tagging to the content (post)

```
<sioc:Post rdf:about="http://owela.vtt.fi/tilkut/clip_id8">
  ...
  <tags:taggedWithTag rdf:resource="http://owela.vtt.fi/tilkut/tag1_Otaniemi"/>
  <tags:tag rdf:resource="http://owela.vtt.fi/tilkut/tagging_tag1_Otaniemi"/>
  ....
</sioc:Post>
```

Appendix B: Example SPARQL-queries of Semantic Tilkut

Example: The meaning of tags

Search tags that are connected to the meaning with help of existing semantic vocabularies. This knowledge makes it possible to fetch extra information and utilise it for enhancing knowledge about the content. Returns also several meanings for the same tag, as well as who has used tag in which meaning.

```
SELECT ?tag ?meaning ?uri ?maker ?link ?vocabulary
WHERE {
  ?tag moat:hasMeaning ?meaning .
  ?meaning rdf:type moat:Meaning .
  ?meaning moat:meaningURI ?uri .
  ?meaning foaf:maker ?maker .
  OPTIONAL {
    ?uri sioc:link ?link .
    ?uri dcam:memberOf ?vocabulary .
  }
}
```

Example: Meaning of the certain tag and in which categories the tag has been used

```
Prefix ex: <http://www.vtt.fi/Somed/Osmo.owl#>
SELECT ?tag_cat ?label ?meaning ?post
WHERE {
  ?Tagging tags:associatedTag tilkut:tag1_Otaniemi .
  ?Tagging sioc:topic ?tag_cat .
  ?Tagging tags:taggedResource ?post .
  ?tag_cat_class rdfs:subClassOf ex:Tag_category .
  ?tag_cat rdf:type ?tag_cat_class .
  ?tag_cat skos:prefLabel ?label .
  OPTIONAL {
    ?Tagging moat:tagMeaning ?meaning .
  }
}}
```

Example: Tags and tag categories relating to the posts of the certain user

Shows also the meaning of the tag if it has been created.

Prefix ex: <<http://www.vtt.fi/Somed/Osmo.owl#>>

```
SELECT ?post ?Tagging ?tag ?name ?tag_cat ?tag_cat_class ?label ?meaning
WHERE {
  ?post sioc:has_creator tilkut:pirjo.
  ?post tags:tag ?Tagging .
  ?Tagging tags:associatedTag ?tag .
  ?tag tags:tagName ?name .
  ?Tagging sioc:topic ?tag_cat .
  ?tag_cat_class rdfs:subClassOf ex:Tag_category .
  ?tag_cat rdf:type ?tag_cat_class .
  ?tag_cat skos:prefLabel ?label .
  OPTIONAL {
    ?Tagging moat:tagMeaning ?meaning .
  }
}
```

Author(s) Bäck, Asta, Melin, Magnus, Näkki, Pirjo, Vainikainen, Sari, Sarvas, Risto, Seppälä, Lassi & Vihavainen, Sami		
Title Tags and tagging Creating meanings, organizing, and socializing with metadata Report on the <täky> project		
Abstract <p>This report summarises the results of the <täky> project (User created metadata as meaning indicator and part of the user experience). The project wanted to bring understanding on the role of tags and tagging as a part of user experience and value to companies when developing future media products and services. The project also wanted to study how freely defined tags and more formal ontologies can be connected to each others. The project looked at tagging both in mobile and web applications. In mobile applications, the key question was what kind of a role does the user interface, application and the visibility of tags play. In web applications, the aim was to study user-friendly ways of creating and utilising tags.</p> <p>Tags were defined in the project as the automation of keyword search. This definition refers to that fact that a tag typically creates a dynamic link to other resources described with the same tag. Tags are more than keywords: they offer functionality. Another aspect differentiating tags from keywords is the cumulative effect of many people adding tags. In current applications, tags are not a single well-defined concept. The applications and their user-interfaces as well as the purpose of tagging affects what kinds of tags are being created. When adding tagging to an application, it is important to consider all these aspects carefully in order to be able to provide a valuable feature to users.</p> <p>A large number of user tests were made during the project, both utilising existing applications, such as Jaiku and del.icio.us, and applications that were developed in the project. Tilkut application was developed in the project for testing semantic support in tagging, and CamQ for playful mobile tagging. Results of user tests are presented throughout the report. Existing ontologies for describing tags and tagging are presented with examples. Opportunities and challenges in adding tags to a commercial service were studied with the help of interviews, questionnaires and by utilising a framework for evaluating and comparing social media features in existing applications and by comparing current and potential future features.</p>		
ISBN 978-951-38-7238-0 (soft back ed.)		
Series title and ISSN VTT Tiedotteita – Research Notes 1455-0865 (URL: http://www.vtt.fi/publications/index.jsp)		Project number 7857
Date October 2008	Language English	Pages 86 p. + app. 4 p.
Name of project TÄKY		Commissioned by Tekes, AinaCom, Nokia, Profium, SanomaWSOY
Keywords tags, tagging, social bookmarking, social media, mobile, presence, location, applications, web application, semantic web, semantics		Publisher VTT Technical Research Centre of Finland P.O. Box 1000, FI-02044 VTT, Finland Phone internat. +358 20 722 4520 Fax +358 20 722 4374

Tagging is a phenomenon of the Web 2.0 era. Tags are user generated keywords that are used and created in a variety of contexts, for a variety of reasons, with a variety of meanings, and in a variety of very different systems. Practically tags are an automation of keyword search. They also bring cumulative value when different people tag resources in open internet applications.

This report summarises the results of <täky> project (User created metadata as meaning indicator and part of the user experience). The project wanted to bring understanding on the role of tags and tagging as a part of user experience and value to companies when developing future media products and services. The project also studied how freely defined tags and more formal ontologies can be connected to each others. Several user studies were made both with mobile and web applications.

VTT
PL 1000
02044 VTT
Puh. 020 722 4520
<http://www.vtt.fi>

VTT
PB 1000
02044 VTT
Tel. 020 722 4520
<http://www.vtt.fi>

VTT
P.O. Box 1000
FI-02044 VTT, Finland
Phone internat. + 358 20 722 4520
<http://www.vtt.fi>
