ATHABASCA UNIVERSITY

INVESTIGATING THE EFFECTS OF NAVIGATION CUES IN A TAG CLOUD

BY

SASA DANILOVIC

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE IN INFORMATION SYSTEMS

SCHOOL OF COMPUTING AND INFORMATION SYSTEMS

ATHABASCA UNIVERSITY
FEBRUARY, 2016

Athabasca University
FACULTY OF GRADUATE STUDIES

The future of learning.

Approval of Thesis

The undersigned certify that they have read the thesis entitled

"Investigating the Effects of Navigation Cues in a Tag Cloud"

Submitted by

Sasa Danilovic

In partial fulfillment of the requirements for the degree of

Master of Science in Information Systems

The thesis examination committee certifies that the thesis
and the oral examination is approved

Supervisor:

Dr. Jon Dron
Athabasca University

Committee members:

Dr. Andrew Chiarella
Athabasca University

Dr. Vive Kumar
Athabasca University

Dr. Stefanie Panke
University of Northern Carolina

February 9, 2016

Acknowledgements

This thesis is a result of my newly discovered passion – tagging systems, and in even wider context, social software as a most probable direction for the infinite wisdom collected from the crowds. At the same time, social software presents itself as a big test of that same crowd, whether it is going to stem wisdom from the indestructible and everlasting obtuseness. This is where we, involved in moving the boundaries of science, have the moral duty to gently nudge these odds.

Although it is not common to write long acknowledgements in this type of publication, I will make an exception. It is my firm decision to include a set of principles in front of any larger work that I write in this area, an idea that I got from the great Isaac Asimov and his three Laws of Robotics. I believe these are the principles that any social (not limited to) software research and development should follow and *understand*. I would like to thank Dr. Jon Dron, who formulated these ten principles, which kept me on track when I was in doubt (Dron, 2008a):

1. **The principle of adaptability** requires that we must try to build small services that interoperate, that we must build to connect using open standards and, where possible, we should build as open source, so that others may adapt and evolve systems to suit local needs.
2. **The principle of evolvability** is that we must build deferred systems (Patel, 2003), systems whose structure is not fixed, systems that can change after the software designer has left the building.
3. **The principle of parcellation** requires that we must build systems in which there are distinct, ideally hierarchical ecological niches that are only weakly or occasionally connected

with each other. Where possible, such niches should emerge rather than be imposed by a designer.

4. **The principle of trust** makes it necessary for us to build the means to reliably identify reliability in people and resources, to protect ourselves from harm and to do this without resorting to top-down constraints.
5. **The principle of stigmergy** is that we should use signs to guide, but not to constrain, and to enable mechanisms to destroy those signs when they are no longer needed or are harmful.
6. **The principle of context** is that, when building social software systems, we must consider the entire virtual ecosystem in which they reside and remember that they are only a part of a much greater whole.
7. **The principle of constraint** requires us to be aware of the constraints that we build into our systems and to use them to enhance learning, much as an architect influences use of a building through the placement of walls, windows and doors.
8. **The principle of sociability** is that attention must be paid to the total system's capacity to enable social presence and communication and that, where possible, this should be embedded throughout.
9. **The principle of connectivity** is that nothing should exist in isolation, everything should influence everything else, much as the beat of a butterfly's wing might affect the weather in another part of the world.
10. **The principle of scale** is that we must be aware of the large and the small in our social systems, and ensure that, where possible, the large should arise out of the small in an endless iterative cycle of renewal.

I am thankful to all who contributed their time and effort to help me complete my research:

- Dr. Jon Dron, for the guidance, inspiring discussions and friendship. Above all for selfless support throughout my entire learning journey.
- Dr. Andrew Chiarella, for the ideas and various perspectives offered at the very beginning of my analysis, when I was confused with which direction to take.
- Dr. Vive Kumar, for constant challenges during my research proposal development.
- Dr. Stefanie Panke, for joining the examination committee, providing very useful feedback and inspiring research used throughout this thesis.

Abstract

Tag clouds as a social networking software allow users to describe a specific resource or another user using metadata, and simultaneously gain an insight into summarized descriptions provided by the online community. By employing visual cues in a tag cloud, users match the power of those external associations with the internal knowledge, and can follow a multitude of navigational paths to a desired resource or topic of interest. This study analyzed the factors contributing to forming of such paths, focusing on visual aspects and user acceptance. The analysis of hundred studies partially revealed the combinations of visual cues having the highest impact on users when selecting a tag, along with the additional factors relevant to successful implementation of tag clouds. These findings, embedded in the proposed development method aspire to provide guidance with design efforts, realized through a tag cloud user interface software simulation.

Keywords: tags, tagging, tag clouds, visualisation, navigation, cues, recommender, folksonomies, social networking software, social networks, social navigation

Table of Contents

Approval of Thesis	i
Acknowledgements	ii
Abstract	v
Table of Contents	vi
List of Tables	ix
List of Figures	x
Chapter I - INTRODUCTION	1
Background	1
Need for Study	3
Statement of Problem and Research Questions	6
Significance of Research	7
Definition of Key Terms	7
Chapter II - LITERATURE REVIEW	9
Social Navigation	9
Recommender Algorithms	11
The Missing Layouts	13
Visual Elements	15
Motivation and Tag Cloud Usability	17
Navigability of Tag Clouds	19
Preliminary study	22
Results	26
Definition of Key Terms	28
Chapter III - RESEARCH METHODOLOGY	29
A Word on Weights	30
Research Design (Phase I – Systematic Review)	31
Publication Bias	37

Research Design (Phase II – Grounded Theory Coding) 38
Research Design (Phase III –Model Design) 41
Special Implications and Limitations for the Findings Model 43
Influence Diagrams and Additive Utility Function Primer 45
Modeling Process 48
Research Design (Phase IV – Developing the Software Artefact) 49
Functional Testing of the Artefact 49
Tools 50
Data Sources 50
Research Limitations and Considerations 51
Definitions of Key Terms 52

Chapter IV - DATA ELICITATION FINDINGS 53
Systematic Review Findings 53
Grounded Theory Topical Discovery 56

Chapter V - THEORETICAL MODEL AND LITERATURE ANALYSIS 65
Theoretical Model Overview 65
User Acceptance 70
Layout-independent Analysis of Tag Visualizations 85
Clustered Layouts 93
Cloud layouts 97
Sequential and List Layouts 100
Recommender Algorithm 104
User Motivation 108
User Behaviour in Tagging Systems 117

Chapter VI - DESIGN PRINCIPLES AND DISCUSSION 125
Domain Analysis (Scrutinize) 126
Layout Choice and Tag Visualisation (Envision) 128
Motivation and Behaviour (Convey) 133

Administrative and User Control Planning (Control) 137
Recommender Algorithm (Recommend) 141
Step 6 – Quality Analysis (Evaluate) 143
Reducing to Realistic Boundaries 145
Answering the Research Questions 145

Chapter VII - DEVELOPING THE SOFTWARE ARTEFACT 149
Scrutinize (Domain Analysis) 149
Envision (Layout and Visual Elements Selection) 149
Conduct (User Motivation and Behaviour) 152
Control (Administrative and User Control) 156
Recommend (Recommender Algorithm) 161
Evaluate (Quality Assessment) 162
Conclusions and Further Research 163

REFERENCES 166

Appendix A - LITERATURE USED IN GROUNDED THEORY CODING 177

List of Tables

Table I-1. An example of tag types and user motivations 5
Table III-1. An example of the preliminary selection criteria 32
Table III-2. Currency (left) and size weight scales (right) 35
Table III-3 Methodology weights explained 37
Table III-4 Coding weights rationale 39
Table III-5 Truth tables (example) 46
Table III-6 Truth tables (example II) 47
Table III-7 Open coding categories 48
Table IV-1 Literature weighting process (table trimmed to fit) 56
Table IV-2 Code quantity per category 58
Table IV-3 List of all categories and topics 62
Table IV-4 List of selected independent variables 63
Table V-1. Impact strength of the independent variables. TS – total statements, RS – relevant statements, TW – total weight, RW – relevant weight, IF – impact factor, IS – item significance. 72

List of Figures

Figure I-1. Common tag cloud types. Sequential (a), cloud (circular) (b), lists (c), clustered (d). 4
Figure II-1. Tagging system state logical 22
Figure II-2. Current system logical with user profiling. 23
Figure II-3. Current system state with problems 25
Figure II-4. Survey results 27
Figure II-5. ICARUS Interface 27
Figure III-1. Steps in the proposed research methodology 29
Figure III-2. Weights ratio 34
Figure III-3 Basic utility function – influence on dependent variable 45
Figure III-4 Additive utility function with multiple independent variables 47
Figure III-5 Modeling process 49
Figure IV-1 Topical relevancy of selected literature prior and after correction 53
Figure IV-2 Literature currency 54
Figure IV-3 Selected publication frequency by year 55
Figure IV-4 Topical representation graphs 63
Figure V-1. Influential topics on user acceptance (light blue – supporting topics, dark-blue – categories). 65
Figure V-2. Aesthetics category with the associated topics (light blue – supporting topics, dark-blue – categories, gray – low impact, red arrow – negative impact). 66
Figure V-3. The assessment of different layouts (light blue – supporting topics, dark-blue – categories, gray – low impact, red arrow – negative impact, purple – insufficiently researched topic). 67
Figure V-4. Recommender algorithm and its associated topics (light blue – supporting topics, dark-blue – categories, red arrow – negative impact, purple – insufficiently researched topic). 68
Figure V-5 Motivational topics through functional perspective (light blue – supporting topics, dark-blue – categories, gray – low impact, purple – insufficiently researched topic). 69
Figure V-6. User behaviour in tagging systems, functional perspective. 69
Figure V-7. The entire findings model (light blue – supporting topics, dark-blue – categories, gray – low impact, red arrow – negative impact, purple – insufficiently researched topic). 71
Figure V-8. Cloud layout variant: white lines (left), tag length and readability (right) affected by tag size (tagul.com). 87
Figure V-9. Background colouring. 91
Figure V-10. Most frequent scanning pattern in a tag cloud 1-4 (left); the proposed quadrant distribution (right). 95
Figure V-11. Topical segregation by colour, varying tag size (left), same tag size (right). 96
Figure V-12. Separating topical clusters by colours (left), displaying all relationships (right). 97
Figure V-13. Wordle visual variants (www.wordle.net). 99
Figure V-14. Alphabetical list layout (left), frequency-based signalling (right). 101
Figure V-15. An example of suggesting a tag using colour. 101
Figure V-16. Use of extreme colours (left and right), framed tags (bottom). 101

Figure V-17. Scanning patterns example, alphabetical (left), random (right). 103
Figure V-18. Vertically aligned list. 103
Figure VI-1 Proposed design principle (SECURE) 126
Figure VI-2 Visual factors’ known relations. 128
Figure VI-3 Layout type selection rationale. 131
Figure VI-4 Clustered layout visual design implications 132
Figure VI-5 Cloud layout visual design implications 132
Figure VI-6 Sequential and lists layouts visual design implications 133
Figure VI-7 Mechanisms for increasing user motivation. 137
Figure VI-8 Administrative control transition to user control 139
Figure VI-9 Recommender algorithm perspective 142
Figure VI-10 Quality monitoring and control 144
Figure VII-1 Dual and single view interface 150
Figure VII-2 Visual cues 151
Figure VII-3 Domain expertize proposed solution 153
Figure VII-4 Harvesting expertize, user perspective 153
Figure VII-5 Rating the entire tag set 154
Figure VII-6 Timestamp 155
Figure VII-7 Tag filtering 156
Figure VII-8 Transparency (help function) 157
Figure VII-9 Administrator panel 158
Figure VII-10 Colour pattern change 159
Figure VII-11 User settings window 160
Figure VII-12 Tag importance on mouse-hover 162

Chapter I - INTRODUCTION

> Social navigation employs the behavior of a group of networked users of a system to influence the behavior of those who follow. Social navigation systems do not determine what people can do, but influence and persuade. (Dron, 2005)

Background

One of the main trademarks of transition from Web 1.0 to Web 2.0 was of social nature: emerging collaboration networks significantly influenced the way users interact with the software. Providing tools that allowed the users to contribute to online resources' description using tag words achieved the most important form of collaboration, known as *social tagging* (Gupta et al., 2010).

Social tagging serves as a principle that allows the user to annotate resources for classification, sharing, and search, and further extrapolating them to recommend similar resources or other users with similar preferences. As large collections of metadata, tags (*folksonomies*) have the primary purpose to aid in serendipitous browsing and content organisation (Lohmann et al., 2009; Melenhorst et al., 2008). Using folksonomies as a solution for navigation stems from low financial cost, high scalability, low user entry barriers, while simple to deploy and maintain (Filho et al. 2009; Gupta et al., 2010; Sanchez-Zamora & Llamas-Nistal, 2009). From users' perspective, folksonomies are self-guiding and stimulate accumulation of knowledge, especially if viewed through the community influence (Kang & Fu, 2010).

Tag clouds represent one of the most widespread form of social software navigational interface, allowing users to browse through social networks' resources using tags set by the online communities' members. Unlike the traditional taxonomies that utilise pre-set classifications, tag clouds have a more demanding task – presenting not only the resource a user is looking for, but also accounting for the popularity within communities, which is a complex overlap to summarize and visualize (Fu, Kannampallil, Kang, & He, 2010). This creates the platform for efficient navigational cues, high personalization, and attractive visualization (Gupta et al., 2010).

Tags can perform various tasks in social network software applications:

- *Indexing*, which provides better experience than simple bookmarking.
- *Search*, where the tag size or other cue reflect the resource popularity.
- *Ontology generation*, useful for cataloguing and general hierarchies.
- *Demarking the social areas of interest*, by acting as descriptors.
- *Improved browsing*, by navigating to a topic of interest.

Tag clouds however exhibit the absence of functionalities supporting many navigational objectives, such as *topical narrowing* or *finding a specific term*, usually focusing on a specific application domain, and varying in appearance. Some examples are (Gupta et al., 2010; Mezghani et al., 2012; Venetis, et.al., 2011):

- *Flickr* (tagging photos)
- *Delicious* (tagging entire URLs)
- *Blogger*, *WordPress* (tagging own posts)

- *Facebook* (tagging photos, assigning likes)
- *YouTube, Metacafe* (multimedia tagging: videos, podcasts, music, etc.)
- *Yahoo* (tagging based on helpfulness of the answer provided: like or dislike)
- *Digg, Slashdot* (tagging news)
- *Yelp, CitySearch* (reviewing businesses or products and tagging the reviews)

With the widespread of application domains and their increasing popularity, exploring the navigational cues' effects on user browsing opened the possibility of standardisation by analysing visual features and widening the application domain, resulting in a first-of-a-kind theoretical framework.

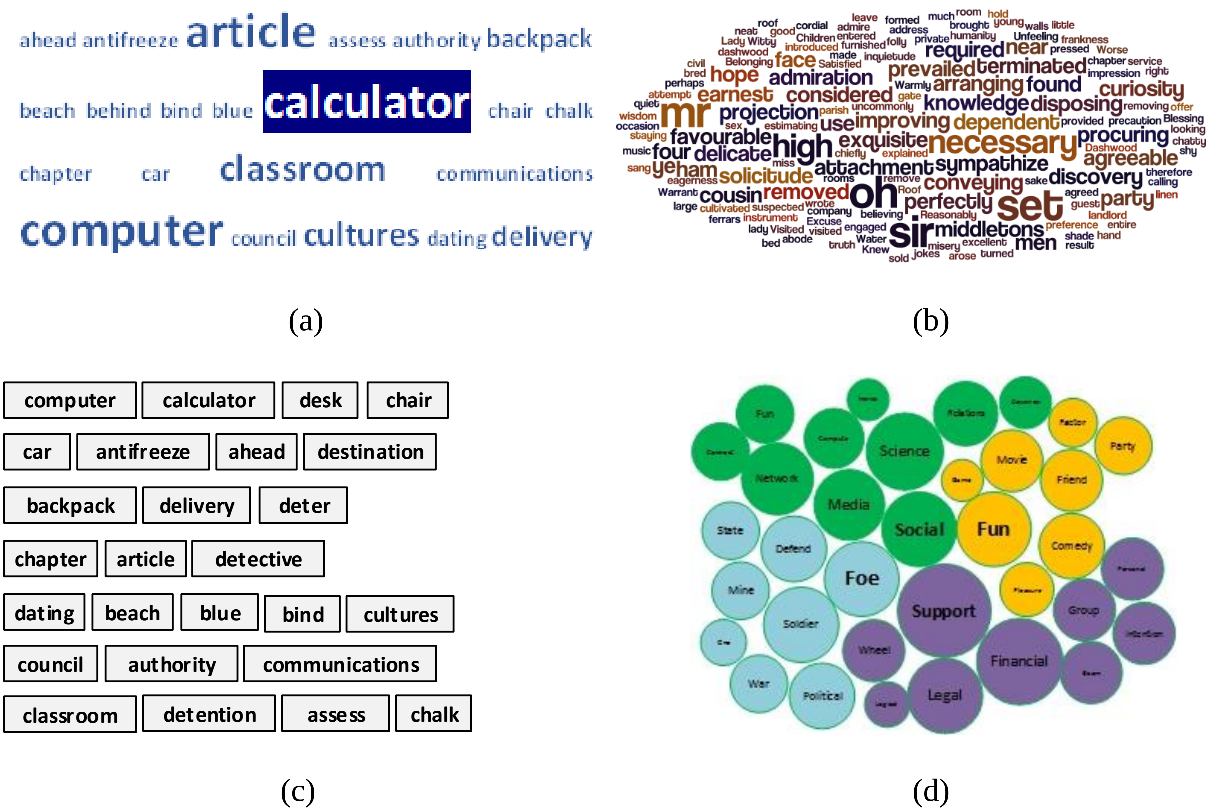
Need for Study

An increased users' adoption rate and the ability of the computation algorithms to visualize folksonomic datasets, inspired the recent technological development of tag clouds, thus creating a synergy. However, those systems exhibit relative navigational inaccuracy, and inadequate catering to user vocabulary or preferences (Gupta et al., 2010; Milicevic et al.,2010). A deeper investigation into problems revealed tag clouds' low success rate in serving high volume of users with varying preferences, coupled with tag popularity biasing, tag semantics and computational factors.

The current research is largely trying to rectify these problems through recommender systems research and implementation, by developing tag suggestion mechanisms that fit into the existing visual interface patterns, such as tag clouds or search engines (Han et al., 2010; Skoutas & Alrifai, 2011; Zanardi & Capra, 2008a). However, most researchers are investing little effort in analyzing tag cloud visual interfaces, apart from rare cases (Chen et al. 2009; Christie, 2011; Filho, 2010), mostly when deploying a novel system. In evaluating the possible solutions to these problems, it is hard to ignore that even the most current and widely accepted recommender systems can underperform if only focused on one aspect of user interaction: providing more accurate recommendations (Fu, et al., 2010). This had to be complemented by a second research direction and develop a suitable graphical user interface, as an aid in resolving these current problems with tagging systems.

Figure I-1.

Common tag cloud types. Sequential (a), cloud (circular) (b), lists (c), clustered (d).



The growing number of social sites and variety of purposes is further stressing this problem, by increasing the number of user motivations for using the tag clouds and tag types, hindered by inexplicit relationships (Table I-1).

Table I-1.
An example of tag types and user motivations

User Motivations	Tag Types
Future retrieval	Content-based
Contribution and sharing	Context-based
Attract attention	Attributive
Play and competition	Ownership
Self-reference	Subjective
Opinion expression	Organisational
Task organisation	Purpose
Social signalling	Factual
Money	Personal
Technological ease	Self-referential

Therefore, the *exploratory part* of this research aimed at discover the factors of visual interfaces that can improve the navigational accuracy by lessening the influence of the underlying recommender algorithm. The *explanatory* part describes the balance of these factors, extrapolating the relationships to form the most fitting theoretical and implementation foundation for a tag cloud interface, considering both recommender algorithms and user motivations for using a tag cloud.

Statement of Problem and Research Questions

The research goal was to design a tag cloud interface that would offer improved user browsing experience in the social network services' tag cloud environments. Before developing the software artefact, there was a clear need to design a theoretical presenting recognized benefits and flaws. To achieve this goal, these questions emerged as a step towards identifying the relevant tag cloud browsing elements:

1. *Which significant factors the relevant research suggest affect the likelihood of an individual choosing a particular navigation path in a tag cloud?*

Given that mere identification of the factors is not enough to inform a theoretical model, the following question further expands the research scope:

2. *What evidence does the relevant research provide on the interrelationship properties of these factors in respect to structuring the navigational paths?*

By defining the relationship strength and influence direction (positive versus negative), the factors involved will become a subject to certain changes when designing the software artefact, which lead to third research question:

3. *Which factors primarily define a tag cloud's successful implementation?*

Answering this question required a survey on several factors found to affect the implementation, such as user acceptance, application and knowledge domains, social context, and others. By diverting the attention to social and user context, answering the third question assumed a meaning distinct from technical implementation.

Significance of Research

The current research in the field is widely supporting the need for innovating underlying recommender algorithms in tag clouds (Cress, 2013; Fu, 2010; Skoutas & Alrifai, 2011). However, although these parts are interdependent, researchers often neglect examining the influence of visual factors on user browsing motivations (Allam et al., 2012; Oosterman & Cockburn, 2010). This justified a wide scope of the first research question: the field was an uncharted territory, which induced a possibility of discovering more factors or consider some of the existing irrelevant during the proposed research. Focusing on visualization also helped mitigate the possibility of findings being unidirectional or experience scope creeps, neither which were desirable. In directing the second question at not only at further exploration and quantifying the relationships, opened the possibility for more intelligent user navigation guidance in the concrete software implementations. Finally, since in this case the user acceptance served as a guideline in discovering the success of the theoretical model and the software artefact, it called for the investigation into relevant factors, resulting in user-centric implementation. By segregating the logical system modules to interactive and computational, followed by the analysis of these individual parts, appointing all of these factors as either depended or independent variables within the theoretical model (Fu, Kannampallil, & Kang, 2010; Skoutas & Alrifai, 2011).

Definition of Key Terms

Web 1.0 – World Wide Web logical organization structure in which the users cannot proactively contribute to the content, only view it.

Social Network Services – a set of users (profiles and social links) and the relevant services offered for various interactions.

Web 2.0 – a World Wide Web structure that allows user collaboration and proactive web content changes.

Folksonomy – collaborative annotation of resources, leading to user-driven classification.

Social tagging – assigning labels to resources belonging to social network service, comprising a set of users, a set of tags, and a set of resources.

Tag cloud – a visual representation of grouped text data, used for user interaction with a social software.

Cold start – the inability of the system to cater to the user because of the insufficient preferential information.

Knowledge domains – the areas of increased human interest or activity.

Chapter II - LITERATURE REVIEW

> With no basic model of visual processing on which we can support the idea of good data representation, ultimately the problem of visualization comes down to establishing a consistent notation. If the best representation is simply the one we know best because it is embedded in our culture then standardization is everything – there is no good representation, only widely shared conventions (Ware, 2012).

Social Navigation

Dron (2005) coined probably the best coherent definition of social navigation: ”Social navigation employs the behaviour of a group of networked users of a system to influence the behaviour of those who follow”. Since tagging and tag browsing represent only a subset of social navigation “class”, they inherit those same characteristics. In a narrow context of social navigation, there are three major influential factors to tagging systems, commonly agreeable in most of the relevant literature: communities of users, recommender (sub)systems, and tag visualisation (Cress et al., 2013; Milicevic et al., 2010; Skoutas & Alrifai, 2011). Users create communities, communities influence recommender systems, recommender systems influence user navigational decisions, and the symbiotic recursive relationship continues. However, these factors and effects are difficult to analyse in an isolation because of the self-organizing nature of social systems (Dron, 2005). Despite this problem, by directing the research efforts at examining those relations and performing factors segregation, the closer the science will be to discovering how to successfully serve user preferences across a crowds.

Still, the views on this topic are opposing, and creating a significant conceptual gap in research direction, whether the social navigation should favourize crowd-preferred cues or the individual (internal) associations. By testing their PETAC system, (Christie et al., 2011) gained positive feedback for user-controlled content filtering. A similar result applies to allowing the social activity insight, or “what content are other users of the system viewing”, while content personalisation recorded a negative feedback. On the other hand, Schoefegger and Granitzer (2012) found that personalized results yield better user acceptance. Held et al. (2012) also found the community influence is important on navigational cue selection, although, in their study, the strength of individual over popular associations was more significant. Furthermore, successful matching of the community-driven popularity will not warrant a successful task completion. (Dron, 2005) found that people respond to typical tag-cloud cues such as tag size, emphasis and list position in inconsistent ways, often behaving contrary to the intent of tag interface designers (for example, by deliberately not selecting the largest tag from the list). Tags often lack a grading scale or standard other than the abstract meaning, applied across communities and contexts (Dron, 2008). Cress et al. (2013) argue the community influence on users’ tagging habits occurs as an indirect navigational effect in social environment. They also found the user vocabulary styles directly depend on the community presence (or absence). Neither the user-centric nor crowd-centric designs provided convincible results so far, especially if measured against the user-acceptance. All this implies the need for systems that will cover wider application domains, but simultaneously allowing more flexibility in interaction, and in that way catering to wider *user groups*. Unfortunately, the current state of these systems is far from quality guidelines for technological aspect of Web 2.0, especially for the concept of universality of access (Derntl et al., 2011). Contrasting with the problem of information intricacies, and exploration of resources with dissimilar facets (Carpendale et al., 2012), hindering generalizations, especially in visual sense. For example, Flickr users benefit from an emphasised photographic content and prefer thumbnails over tags (Diaz et al., 2009), as contrary to e.g. Delicious users who benefit from semantic representations. Since some application domains carry significant thematic and motivational differences, it increases the need for theoretical modeling – while it is not necessary to achieve the universal appearance, the logic behind successful designing should be at least similar. The examples are present in most of the modern literature on systems design. Once proved, these user acceptance paradigms in an e.g. desktop applications, no software developer makes acute changes any longer.

Recommender Algorithms

The more accurate, complete and relevant tags are to individuals' needs, the more likely they are to be of value. According to several survey papers on tagging systems, main problems in providing accurate navigational cues are semantics, cold start, tag noise, correlating tag sets, and vocabularies (Chen et al., 2009; Christie et al., 2011; Gupta et al., 2010; Milicevic et al., 2010). Possibly one of the most important findings related to recommender algorithm is in identifying semantical flaws: an error in the initial construct will cause sub-optimization in following the navigational paths, leading to the lowered recommendation accuracy (Cress et al., 2013). Schoefegger and Granitzer (2012) compared a community-based tagging system with a personalized one using user modeling and inferred that richer user profiles gain higher personalized results. This leads to a cold-start problem: results' precision grows with user interaction, but until that occurs the system will be suboptimal and consequently underused. The study also showed a decrease in personalized results precision with richer user profiles, which leads to conclusion that tags alone are insufficient in upholding the profile efficiently. Skoutas and Alrifai (2011) proposed a model using tag co-occurrences frequency analysis, random walk on tag graphs, tag diversification and rank aggregation. The results demonstrated that tag diversification and weighting have no notable effect of the computational performance of the system when contrasted to the typical frequency-based tag ranking, but do provide for more accurate results. They had also found the navigational cost decreases with the tag cloud growth, and the selectivity increases, following the number of diverse (new) tags.

Regardless of previously mentioned problems, tagging has become deeply rooted in the users' habits when interacting with current Web 2.0 software, especially social networking services. The constant evaluations of the current various tagging techniques to systems design are not only desirable but also necessary as they provide both lateral and iterative improvements to previous ones, hopefully leading to some extent of standardization across the Web 2.0 platform. One course of action is the visual interface study, to provide simplistic and abstract separation – attracting the crowds through combined attractiveness and functionality, and catering to the individuals by providing better recommendations. Although recognizing recommender algorithms' performance is important, designing a model that will not highlight these problems any further should be imperative.

While some of the research (Chen et al., 2009; Christie et al., 2011) aimed at creating both novel interfaces and algorithms, they had neglected visualization as equally important to successful navigation, considering it a collateral. In evaluating these problems, it is hard to ignore that even the most current and widely accepted recommender systems can underperform if focused on providing more accurate results (Candan et al., 2008; Fu et al., 2010; Kang & Fu, 2010). Thus, many studies incorporate *a priori* tag cloud interfaces without considering alternatives, which creates an effect of incremental diversification of solutions. This approach is arguably contrary to scientific as it involves a significant amount of guesswork – inadequately designed user interface can hinder an assumedly good recommender algorithm, often stemming from the lack of features standardization. Zamora and Llamas-Nistal (2009), point out the lack of default visualization in tag clouds as a major drawback.

Another often reproduced practice is promoting a mechanistic view of the system - researchers often use the means of automation to test their newly developed ideas. This involves collecting a publicly available data dump and employing different algorithms to perform a comparative analysis of their solution (Au Yeung & Iwata, 2010; Liang, 2009; Trattner, 2011). By not involving real users, they interpret better performing system as an increase of computational performance (or accuracy of predictions) of the recommender algorithm. However, users may not realise this difference, especially if the system is (and often is) relying solely on this aspect (Hearst & Rosner, 2008). Although most of tag cloud research aims at user behaviour and motivation, it rarely finds its way into literature review when presenting such a solution, in other words, not considering user preferences and behavioural patterns.

The Missing Layouts

Many researchers point out the importance of cloud layout and its effects on user behaviour (Gwizdka & Cole, 2013; Oosterman & Cockburn, 2010; Skoutas & Alrifai, 2011), but there has been little comparative research focusing on this aspect. Lohmann et al. (2009) conducted a study on the effects of different tag cloud layouts on navigational cues and the results implied that a particular layout will perform best if it is task-bound, thus steering away from the application domain universality. However, there has been little research into hybridizing different layouts to assimilate the layout-dependent benefits, or to assess whether certain compromises are viable. Skoutas and Alrifai (2011) although not directly examining layouts in their research, point out the need for different layout consideration when implementing the underlying recommender algorithms. Oosterman and Cockburn, (2010) stress the need for comparative analysis of tag cloud layouts and visual elements, also supported by Deutsch et al. (2011). Though these studies are useful after already making a layout selection, the narrow research scope does not allow for any compromises or guidance on layout type selection.

Some studies (Rivadeneira, 2007; Sanchez-Zamora & Llamas-Nistal, 2009; Schrammel, 2009) assess various layouts, however this occurs only within one category (for example, variants of clustered or hierarchical tag cloud). Although this is helpful when considering the visual variants of a single layout, by providing an input for subsequent studies, still it is insufficient for any abstracted design considerations because of its narrow scope. Similar effort, made by Kaser and Lemire (2007), who slightly expanded the layout scope, still considered only a part of the palette. They underlined the limitations to the existing software types that position the tags automatically, and the need for a properly defined metric for clustering. In their later research, and using heat maps, Schrammel et al. (2009) found no significant deviations in user attention within hierarchical tag cloud variants, which also points out to the necessity for scope widening. However, they determined the attention increases from top to bottom, and from left to right, which is an excellent modeling consideration. The current research rarely assesses tag cloud layouts in a truly comparative way (other than recognizing the need for it), thus disclosing a clear knowledge gap. The lack of control tag clouds deprived of nearly all size or color variations, and neutral tag disposition further prevents objective inferencing. There were partial exceptions (Dron, 2005; Helic, Trattner, 2010a; Waldner et al., 2013), however, resulting in limited statistical evidence within corpora.

Pérez García-Plaza et al. (2012) qualitatively evaluated the possibilities of clustering tags, building on the ability of clusters to taxonomize the information based on its relevancy. Their information filtering ability showed improvements to conventional tag clouds in widening the application domain, also supported by Fujimura et al. (2008). During the visual design of their tagging system Di Caro et al. (2011) agree with the clustered advantages, and further point out the ability to preserve context in such form, in accordance with findings of Chen et al. (2009). However, clustered layouts suffer from spatial consumption, which carries noticeably low user acceptance (Chen et al., 2009; Christie et al., 2011; Danilovic, 2013). Cui et al. (2010) found that closer distribution of tags in a circular cloud will provide higher navigational cue value, while semantically coherent tags will retain better spatial stability. Advantages and disadvantages of different tag cloud layouts vary, which opens the possibility for layout features hybridization, a relatively unexplored area.

Visual Elements

Bateman et al. found that changes in visual properties of a tag carry an overall low impact on the quality of user interaction with the tag cloud. Certain visual elements have higher influence on navigational cues, especially size and location of a tag, supported by several related studies (Chen et al., 2009; Dron, 2005; Rivadeneira et al., 2007; 2009). Dron et al. (2005, 2008) support this view but also finds that other elements (colour, list order, saturation) can acting as competing, thus increasing the complexity of the navigation cues. The research results of Chen et al. (2009) did confirm some of those findings for the clustered tag cloud layout, which is encouraging as a research pointer. Halvey and Keane (2007) found the same visual elements have much higher influence on the users' navigational path choice.

According to some studies, color can have neutral to negative connotation affecting readability and contributing to visual clutter (Lee, 2010; Waldner et al., 2013), however, its positive influence on user satisfaction is significant (Viegas, 2009; Waldner et al., 2013). Dron (2005) found visual factors interdependent, with explicit navigational cues having the biggest effect on the user behavior. Combining tag weight and colour can be a dominating cue (Bateman, 2008), especially if the application domain performs within simplistic boundaries, such as social annotation software (A. F. Chiarella, 2012; A. Chiarella, 2011). Viegas et al. (2009) proposed a closer academic examination of already operational visualization alternatives. Further analysis into the exact uses of different visual properties would be suitable, especially if those scale in effect close to the semantical cues and having the similar potential for causing erroneous paths (Cress et al., 2013).

In the design of their tag cloud (SparkClouds), Lee et al. (2010) had several considerations, such as: space conservation and readability balance, tag size proportionally reflecting the tag importance, using gradients instead of plain colours for background and foreground, proper contrasts, etc. Since high colour differentiation is present in nearly all tag clouds, they have based their system on a belief that too much colour increases visual clutter. However, in an empirical comparison with other tag clouds, SparkClouds showed no significant deviation; also, the computational requirements to produce gradients led to system's instability and decreased performance, with a direct effect on user satisfaction. The research is useful in proving the tag colour is not one of the key navigational cues relevant to user acceptance. Waldner et al. (2013) based their study on nominal and ordinal placed tags, while retaining circular shape of the cloud. Although they claimed that tag colours can affect readability, they also voice the need for further study since the factual effects on the navigation cueing are unknown. The results showed that faded colours have low user acceptance, however coloured tags could be suitable for categorization. Although the colour and tag shape can carry negative cognitive impact, considering them during design time can contribute to user acceptance.

Dron (2005) performed a study encompassing many of the key visual elements and limited but well-rounded layouts, focused on revealing the relations with navigational behaviour, in an environment similar to tag clouds. The assessed elements included text positions, sizes, word-lengths, list orders, etc. The results pointed out to the high interdependence of the visual factors, with explicit navigational cues yielding the biggest effect on the user behaviour.

Motivation and Tag Cloud Usability

Hearst & Rosner (2008) suggest that tag visualization should use the opposite logic to typical data visualization, and the design must account for that. This finding is in line with (Danilovic, 2013) research results; the proposed model scored rather low aesthetically, hypothesizing the users will sacrifice some of the performance to increased visual appeal. Findings of several studies implementing novel systems support this finding (Seifert, 2008; Viegas et al., 2009), even when comparing tag cloud performance to a search engine (Christie et al., 2011; Kuo, 2007).

Apart from functional ones (e.g. content generation, information retrievability and refindability), the hedonic perspective received little attention (Allam et al., 2012). Another related problem that appears is examining of user motivations and behaviour by employing tag use analysis. One of the main negative effects of this approach is an introduction of personal bias, since it is impossible to realistically quantify user acceptance, leaving the whole process to a liberal interpretation. Milicevic et al. (2010) instead suggest using interviews to provide access to conscious intents behind the tag cloud use. Just because 1.1 billion people use Microsoft Office, does not imply that they are happy with the product, and any use analysis based solely on passive observation will produce biased results, at the least.

Despite their popularity and seeming intuitiveness, tag clouds may not provide the most effective user interface for accessing information. Oosterman and Cockburn (2010) claim that in current technological state tag clouds underperform for most navigational tasks and propose further research. Viégas & Wattenberg (2008) call for innovation in design to achieve high user acceptance of tag clouds. Kuo et al. (2007), acknowledging tag clouds' social significance, however find them an inadequate tool for discovering relational concepts, a concern supported by Waldner et al. (2013). Helic et al. (2010b) claim that tag clouds only perform well in theory, current implementations are far from being efficient, especially in the user interface ergonomics. Fu at al. (2010) provide one possible explanation, crediting poor information value presented by tag clouds to tehnological infancy. They also perceive tagging as a form of Web 2.0 conversation, stemming from presonal user knowledge. Carpendale et al. (2012) underline the need for a clear user task definition keeping in mind real users, and call for a better support of visual exploration by diversifying features. Panke and Gaiser (2009) note tag clouds as a self-learning environment in which the interface transparency is questionable for all users, i.e. their expertise levels. Although addressing the immersive worlds, Dron (2014) suggests further investigation into user reaction to interfaces and algorithms, by using approach that goes beyond stigmergic effects. This notion is interesting considering the monolithic dominance of purely behavioural (aiming to assess the current state) or purely technical studies employing mechanistic and computational logic to satisfy user needs. User motivation and behaviour received little consideration before design time.

Even though all tag clouds have their functional advantages and disadvantages, there were nearly no attempts to offer higher user control through at least limited visual customization. This rigid practice contrasts the essence of social aspect and creates yet another gap – deterministic approach to developing tag clouds. For example, if clustered layouts excel at preserving context of information, and alphabetical lists perform better with countries, directories, nomenclatures, etc. (Halvey & Keane, 2007), then the design choice commonly must become bound to a specific application domain, thus lessening the universality. This sets a speed bump to easy migration from one domain to another, and repeatedly introduces a new learning curve for the end user. Another problem arising is rich visualization or high user control, with the ability to overpower the user (Di Caro et al., 2011).

Navigability of Tag Clouds

Tag clouds' integration in social networking software was successful for two reasons: the ability to use the aggregated resources' descriptions to form an interactive content (Filho et al., 2010; Lohmann et al., 2009), and the power of attractive visualizations widely accepted among users (Viégas & Wattenberg, 2008). These descriptions form a user interpretation of the website, unlike keyword matching (Chiarella, A. F. & Lajoie, n.d.). Its widespread does not have roots in technological or systems design innovations, but for providing a basic functionality easily integrated in different contexts (Panke & Gaiser, 2009). Tag clouds can be an efficient navigational tool (Berlocher, Lee, & Kim, 2008; Melenhorst et al., 2008), and preferred to a search engine for serendipitous exploration of a dataset (Sinclair & Cardew-Hall, 2008), especially for topical novices. Apart from the ease of use, tag cloud presence can lessen the text-based content dominance (Deutsch et al., 2011), and prevent information overload (Allam et al., 2012). By acting as an activity summary, tag clouds can be a valuable asset to online communities, mirroring both groups' and individuals' interests in a fun manner, as opposed to serious and businesslike (Viégas & Wattenberg, 2008).

Hearst and Rosner (2008) consider data visualization logic inadequate for tag clouds and stress the design must account for that. Through the analysis of the Wordle cloud, Viegas et al. (2009) credit its popularity to the innovative graphic interface. And while appeal is important to attracting users and subsequent user acceptance, its navigability is questionable when compared to other search interfaces. Oosterman and Cockburn (2010) found tag clouds' colour coding to be inferior to tables in error rate and search speed, especially for finding maxima and minima in datasets. Tag clouds are an evolutionary step from text searches, however only in a supporting role for easier query formulation (Carpendale et al., 2012; Filho et al., 2009; Waldner et al., 2013). Since links between tag-sharing contents lack explicit visualization, they are difficult to identify and follow (Lohmann et al., 2009). Tag clouds excel in relaying generic concepts and the impressions of a domain, however lack the accuracy for finding a specific term or topic (Trattner et al., 2011). Besides, the lack of organized structure in growing social systems and user-generated data (Hotho et al., 2006) can negatively affect navigation paths structuring. Helic et al. (2010b) set the this unstructured data as the main obstacle for the recommender algorithms to be efficient in supporting most of the typical tag cloud navigational tasks in practice. This "imprecision" however can be valuable to the education environments where accidental learning takes place while exploring a path to the desired goal, similar to browsing in a department store (Gwizdka, 2010). The absence of central authority empowers tagged resources' shared meaning (Derntl et al., 2011), and interpreting concepts is driving tags' selection (Fu, Kannampallil, &

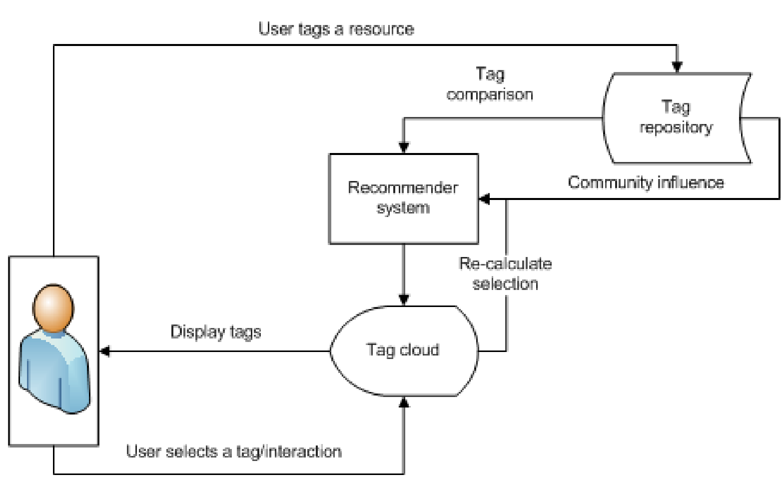
Kang, 2010). Since concept interpretation occurs at individual level, it lacks the same meaningfulness. Even though data visualization and recommender algorithms are inseparable, it is necessary to isolate and identify the visual limits of a tag cloud before examining the latter implications and the potential improvements. For example, competing visual cues are troublesome because different people interpret them differently (Dron, 2008b), and even the most sophisticated recommender algorithm cannot compensate for this. On the other end of the spectrum, if users were to browse a resource set without any cueing aids, it would require increasingly longer identifiers to locate a specific resource (Chi & Mytkowicz, 2007). Employing either emphases or information obscuring can effectively point out specific patterns in a dataset, however this depends on the designer's goals and navigational assumptions (Dörk et al., 2013). Since representation of search results is a fundamental aspect of information retrieval systems, and directly affects users' ability to assess the information relevancy (Gwizdka & Cole, 2013), it is important to examine *how* before *what* to signal.

The pioneering social sites in this area (Flicker, Delicious) have recently started replacing tag clouds with other content navigation systems or ordering, probably caused by low use. The root cause for this perhaps can be in stagnant content (Dron, 2008b), with little to offer to long-time users. Simultaneously, tagging sites such as Wordle, Tagaul, or Tagxedo have blossomed, allowing users to use own keywords to create interesting printable tag-cloud-themed artefacts, such as coffee mugs and t-shirt decals. This effect is not surprising since artistic visualizations, unlike scientific, hold no aspirations for universal truth (Dörk et al., 2013). Therefore, the efficiency of tag clouds as a sole navigational tool needs further investigation, especially in forming longer navigation paths and the means of signalling them to users.

Preliminary study

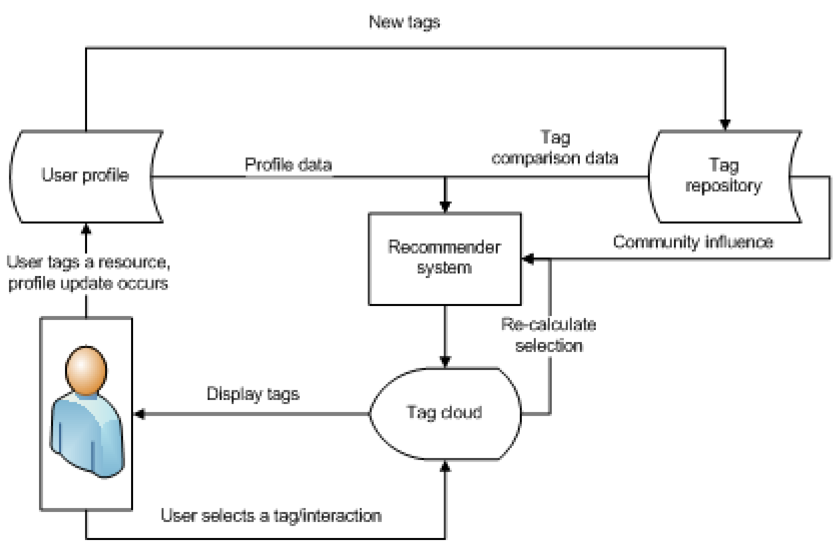
Tag clouds inherently strive to reduce the user navigational effort through automation, and provide a good (although not perfect) organic mechanism against spamming because of the strong community influence on the recommended results. They excel at topical narrowing or broad classification, and have a lower cognitive load than using a search engine, while providing a simple visual interface for the underlying database. The most obvious tag cloud disadvantages are information obscuring, higher interaction effort in reaching the desired answer, and inadequate ability to narrow the search to the exact object of interest (Gupta et al., 2010; Milicevic et al., 2010). The analysis of system's logical operation (Figure II-1) shows user interaction with the system and all related influences throughout the tag cloud, which serves as an interface and enforces community influence in tagging habits. Whenever a user tags a resource, the tag repository database keeps the record of the chosen tags, later consulted by the recommender algorithm.

Figure II-1.
Tagging system state logical



This system does not consider user profiling directly as not all systems do, while diagram 2 depicts such a system (Figure II-2). Depending on the recommender algorithm type, the displayed tags will have *different dispositions* in a tag cloud. User profiling however can occur even with systems not implicitly designed for such role, for example, when user views a specific resource.

Figure II-2.
Current system logical with user profiling.



With user navigation in a tag cloud this effect is obvious when different users choose different navigational cues and paths. The difference between these systems is in navigational accuracy for the user, and the cold start effect associated with systems having no user profiling. If the system has no user profiling, collecting the user data every time, reduces the accuracy of the navigational cues.

Even the best recommending algorithm can fail in providing the accurate navigational cues if tag placement in the cloud is inadequate. Some criteria for visualising tags in the cloud can be:

- Tag’s ability to describe the resource compared to the other tag used to describe the same resource.
- The number of other resources a tag can describe, compared to other tags.
- The ability to compensate for the other resource’s less describing tags, etc.

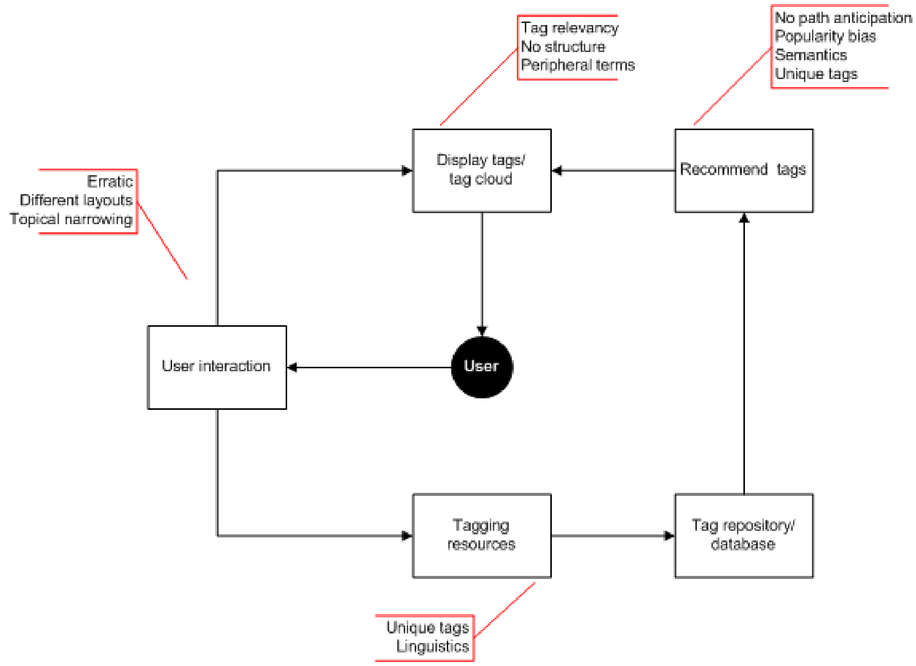
Since tag cloud visual interface is impossible to separate from the recommender algorithm, it is necessary to include any delimitations of these subsystems in the related problem analysis. From a deeper literature review, it is possible to identify tagging problems:

User interaction

- Users are unsure that topical narrowing will lead to the expected goals.
- Tag clouds rarely imply topical relevancy vector distance to decentralized tags.
- There is no structure, making reversible navigation impossible.
- Users tend to focus on the central keywords in a tag cloud, although often the marginally positioned ones could be more relevant.
- Tag clouds usually intersect the user preferences with all the resources and users, leading to generalized, all-purpose cloud.
- The tag clouds are often erratic in appearance, which can create user confusion with navigation.
- Different layouts of tag clouds across the different social networking services lead to steeper learning curve.

Figure II-3.

Current system state with problems



Recommender algorithm

- Users cannot anticipate the next offered set of tags before selecting it, making navigational paths obscure.
- Tags can point to several resources, which makes it hard to predict navigation path likelihoods against the user preferences.
- Tags are popularity biased against user preferred.
- Tag semantics are difficult to extract.
- Unique tags are difficult to identify and remove from the results.
- Spamming.

Following figure II-3, the problems are present at every stage of recommending, further intricating both the problem identification and the associated responding actions (Figure II-3).

Results

Addressing the common tag cloud visualization problems resulted in *ICARUS* tag cloud interface simulation, mainly inspired by the Kolline tagging and search system (Filho et al., 2010). The interface was designed in HTML and embedded in a web page with other tag cloud layouts specifically designed for comparative testing. Twenty seven participants recruited online have completed the survey, which allowed for user feedback during design as opposed to formal evaluation of a finished system. For testing purposes, all interfaces contained static links in an identical structure. Users performed two most typical tasks relevant to tag clouds and evaluated each feature importance, collecting the data for future development. The intent was not to discover quantitative response times and accuracy but to identify reactions when presented with a novel environment interface.

The results of the study were revealing (Figure III-4). Unlike Lohmann et al.(2009), ICARUS scored well in finding a specific term, however the absence of the recommender algorithm could have hindered the results, and other layouts lacked the ability to display large sets of tags at once. Both research indirectly support that clustered layout is more suitable for topical narrowing, but that task surprisingly received moderate importance rating. By following Bateman et al. (2008) design guidelines, the results supported their findings on aesthetics, since circular tag layout received the highest score, followed by ICARUS. A low score is probably a result of vertically placed tags, which may not had been easy to read. Chen and Santamaria (Chen et al., 2009) found smaller sized tags have less chance of attracting user attention, which is counterproductive with the topical narrowing and setting the tag association. Research results confirmed their findings, since the participants rated this feature in the circular cloud environment to be of minor importance.

Figure II-4.
Survey results

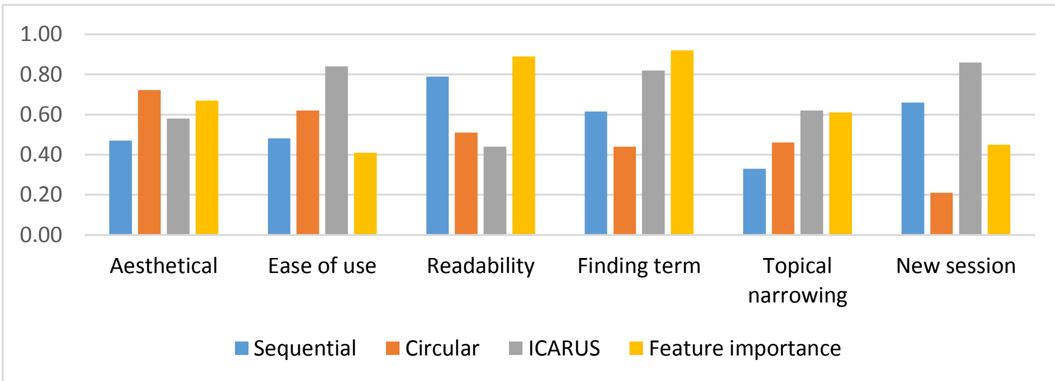
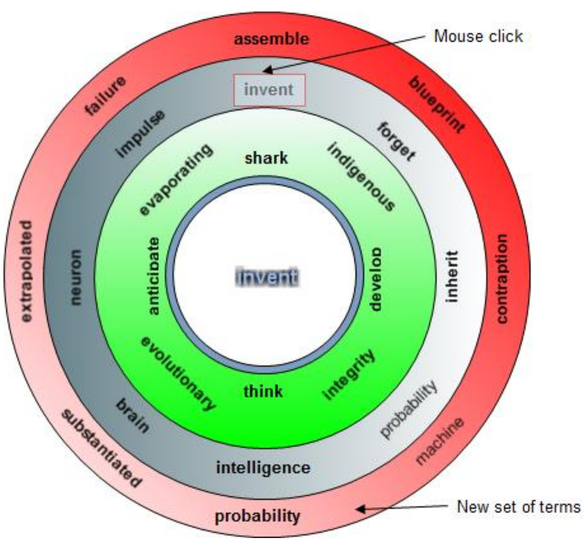


Figure II-5.
ICARUS Interface



The contrary finding stands for to "ease of use", in which ICARUS scored higher than the alternatives. Since aesthetics got a low score, another contributing could have been large space consumption, inherent to clustered interfaces (Chen et al., 2009; Christie et al., 2011; Weiwei Cui et al., 2010). The readability scored the lowest for the ICARUS, which is perhaps attributable to vertically positioned text and the mitigation between the interface and tag sizes.

Definition of Key Terms

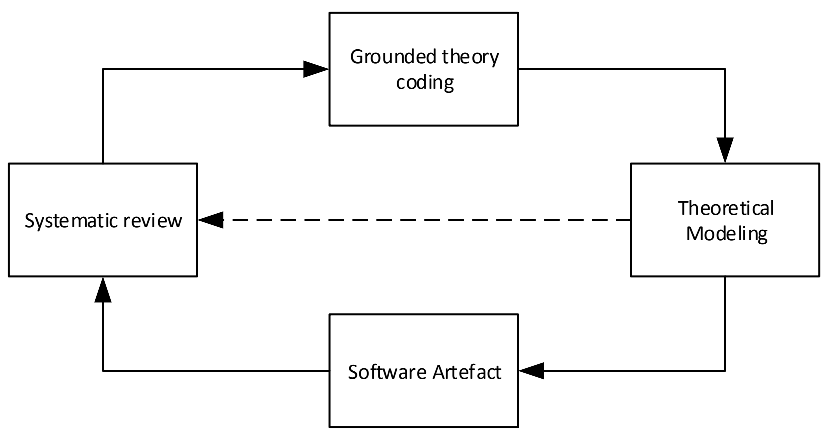
ICARUS – **I**nteractive tag **C**loud with **A**daptive **R**ecommender **U**ser **S**ystem.

Tag weighting – evaluating the relative tag correlation to a user preference.

Chapter III - RESEARCH METHODOLOGY

The absence of solid theoretical base from which to draw some principles justified the choice of qualitative methodology for this study, preventing more linear research flow or more statistically informative research methods. Therefore, filling the gaps in previous knowledge and theoretical modeling are in the focal point of the study. Grounded theory was the main data collection method, supplemented by systematic review literature selection criteria, comprised of four phases.

Figure III-1.
Steps in the proposed research methodology



The chapter describes the research framework employed in these four phases, starting with systematic review, followed by the grounded theory coding, to model framework and finally developing the software artefact (Figure II-1).

A Word on Weights

In the later stages of methodology methods and techniques selection, it became clear that both systematic review and grounded theory suffered from descriptive taxonomization: both employed liberal interpretation of either literature selection (*acceptable* or *not acceptable*), or grounded theory codes (*evidence exists or not*). The goal was to reduce personal bias in both cases, and without assigning some guiding criteria, that bias would have inevitably emerged. It is easy to find a wide range of literature both confirming and refuting a specific result, however not all carry the same weight nor quality. A typical systematic review would aim at classifying the primary research, extrapolating effects under observation, and then saturating the topic with secondary and tertiary ones (Keele, 2007; Kitchenham, 2004). That kind flow was difficult to follow because of missing quantitative data types, and a high number of variables involved. Therefore justifying the use of systematic review only for literature selection, followed by grounded theory data elicitation contributed to the overall stability.

A similar problem is present with statements: grounded theory does not distinguish the evidence strength in cases of liberal interpretation: "*We believe that*" versus the ones stemming from processed data "*The results showed*". Attempting to reduce this descriptive taxonomization and its negative effects resulted in a weight protocol design, uniformly applied to any statement considered. Although weights do not deprive statements of personal bias considering the weight system does not adhere to any scientific method or use statistically supported weights, it adds consistency to the process. That means that all publications and statements undergo the identical evaluation, which at least promotes higher objectivity.

Another reason for using weights stems from theoretical model design: categories expressed only in quantity of relevant statements (as dictated by grounded theory) are not effectively relating to their outcomes. It would have been impossible to note the amounts of influence without using a form of quantification, which grounded theory does not support. For example, presenting a discovery that *tag size* attracts most attention would have high certainty, but it would become difficult to compare its influence to e.g. *tag colour* without introducing a narrative bias.

Research Design (Phase I – Systematic Review)

Using systematic review to select the literature based on its relative importance and results validity, complemented grounded theory that provides no such means. A method initially considered was meta-analysis, however a preliminary study revealed it was nearly impossible to find a sufficient amount of similar publications, using the same research variables, methodologies, or even having the same goals. Since systematic review does not need similar samples and study effects, it imposed itself as the most suitable choice.

Besides systematic review, assigning weights to publications using fuzzy values instead of descriptive categories (Khan, et al., 2003; Meline, 2006; Treadwell, et al., n.d.), increased the quality of selected literature corpora. Another benefit is the ability to create refined acceptance filters, and manipulate the corpora towards the desired topical area (Chapter IV).

The first step of systematic literature reviews frames the filtering questions, thus giving consistency to literature selection. These questions, based on the scope of this study formed the following subtopics:

User acceptance (UA) – according to the preliminary study, studying this aspect had proven insufficient, equating acceptance to system's efficiency. More recent research (Farooq et al., 2007; Viegas et al., 2009) had identified factors such as aesthetics, ease of use, higher user control, familiarity, etc., are the factors which better stimulate user acceptance, assigning higher scores to studies that considered these factors.

Visual properties (VP) – this category is in the focus of this thesis research questions, therefore providing a satisfying selection criteria.

User motivation (MT) – since the navigational cues in social navigation are not static pointers but a folksonomic system, it is important to understand factors ranging from motivational to cognitive, and why users tag to why they select a certain tag. Understanding the user perspective reduces the influence of mechanistic views, and provides for better ergonomics.

Table III-1.
An example of the preliminary selection criteria

Publication/Criteria	UA	VP	MT	RA
Publication 1	Yes	No	No	Yes
Publication 2	No	Yes	No	Yes
Publication …(n)	No	No	Yes	No

Recommender algorithms (RA) – although the effects of the recommender systems were not in focus, it is an integral part of tag clouds and considering some relevant aspects was necessary, especially for modeling purposes.

Assigning fuzzy value ranges supplemented this liberal literature selection method (Feng, 2004; Shapiro, 2009; Zadeh, 1975) using three criteria:

- Currency weight (CW)

- Size weight (SW)
- Methodology weight (MW)

The weight values quantified these criteria based on the *perceived importance* on an interval scale, ranging from zero to one. The intervals provided an advantage to newer and richer publications, while not presenting a barrier to older or shorter but informative ones. By assigning the highest value to the methodology used in a publication (Table III-3), and adding weights for each category provided the final result:

Document Weight (DW) = (CW) + (SW) + (MW).

By setting the number of preselected publications to hundred as a quantity that can produce sufficient inferences, yet it did not carry an excessive grounded theory coding overhead.

<u>Example</u>*: if a document is older and short, obtaining the minimum allowable weightings in respective categories, but has a good methodology, the following can applies:*

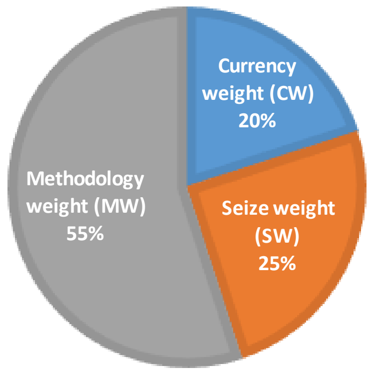
DW = 0.1 + 0.06 + 0.55; DW = 0.71

This result would rate the publication just below the median of passing score and the maximum allowed value. The opposite logic applies if the publication is newer, richer (longer), but the methodology employed inadequate (speculative, not universally applicable results, etc.), then:

DW = 0.2 + 0.25 + 0.09; DW = 0.54

Figure III-2.

Weights ratio



Here, the publication would not have passed the threshold and considered for grounded theory coding. This method does not differ from any employed by a typical systematic review, but instead of using descriptive criteria (Khan et al., 2003; Meline, 2006; Treadwell et al., n.d.), assigning numerical values. These values are useful in two ways: by applying weight to grounded theory coding, the document weights contribute to the overall evidence accuracy, and secondly, two-part weighting process reduces personal bias. Because the literature occurred first, and assigning coding weights as a second step, it further solidified this process. In the third step, multiplying the grounded theory weights by previously obtained document weights, reduced the effect of favouring a certain evidence or theory, using the following criteria:

Currency weight (CW) – the currency grading of the publication reviewed. The newer-dated documents gained higher scores, however this scale had the least impact on the overall result. Newer research efforts aspire to break the chains of traditional practices *in this field* and provide novel solutions instead of correcting the flaws of the existing (older) systems. However, preliminary literature review revealed a significant decline of research work, and some of the higher-quality publications are older, therefore the scale had to have only a minor influence on the overall score, but remain within the basic guidelines of the systematic review method.

Table III-2.
Currency (left) and size weight scales (right)

Currency Weight	Size Weight
After 2009 (0.2)	>15 pages (0.25)
	8-14 pages (0.22)
Prior to 2009 (0.1)	5-7 pages (0.15)
	4> pages (0.06)
0.2	0.25

Tuning the scale in such a way it can enriches the value of a publication but alone has no power to eliminate or carry a significant impact, only gave a minor advantage to newer research (Figure III-2). Within the overall model, the extracted contradicting statements did benefit/suffer from this weight, while agreeable ones had a positive cumulative effect. Year 2009 was selected as a reasonable last six year period for "*scientific freshness*".

Size weight (SW) – where building the theory is the primary goal, publication size matters. Larger publications provide more detail and better descriptions of the underlying principles, and often provide longer elaborations on choices made. This scale is non-linear, since it is challenging to show any evidence in up to four pages, especially in peer-reviewed journal format, some long publications had no value other than simple surveying of the existing systems and approaches, but the latter was less frequent. From Table III-2, the increases among weight categories do not grow linearly, ensuring that short publications get low score, with weights reduced in steps of 3%, 10%, and 19% of the overall value. This ensured that publication size carried no significant penalty unless it was very short, but there is nearly a symbolic weight difference in top two ranges, 8-14 pages and over 15 pages (Figure III-2).

Methodology weight (MW) – weighs the methodology used in the publication based on two main criteria: relevance and quality. Although this research field offers many novel and interesting solutions, the publications employing methodologies for evaluating their relative effectiveness were given preference.

The weights grow linearly, targeting objectivity, except for the lowest score, subsequently adjusting it to prevent poor methodologies from reaching the threshold level (Table III-3). While assigning weights assured the publication quality, it was necessary to develop a method address the topical quantity in advance: there was a clear need for dominance of publications presenting the findings from areas of tag visualization and related user acceptance. This lead to creating second filter, addressing the topical representation applied already weighed publications. The filter employed four criteria, each with two possible states: *present* (1) or *absent* (0):

New system (NS) – a publication presenting a novel system, which was important in assessment of different perspectives when solving the existing problems.

Surveying/Assessing (SA) – reviews and assessments of the existing solutions, important to avoid the proven erroneous approaches. *Recommender algorithms (RA)* - a publication focuses on recommending algorithms.

User analysis or Cognitive (AC) – the analysis of user acceptance, user behaviour, motivations, visualizations, and navigational cues.

Table III-3

Methodology weights explained

Weight	Weight criteria description
0.550	High relevance to user acceptance and tag cloud user interface (UI), or, high relevance to tag cloud UI, the experiment conducted on a larger sample, or, a deep analysis of user acceptance with user motivations investigation, or, good research methodology, novel system, and user tested, large sample.
0.412	High relevance to UI, no user acceptance, or, good research methodology, tag cloud-related cognitive issues addressed, or, good research methodology, tag cloud-related motivation issues addressed, or, good research methodology, tag cloud-related user behaviour issues addressed, or, a sound research methodology, novel system, and user tested, small sample.
0.275	Surveys, or, good research methodology, novel approach, and automated/simulated testing, or, user cognitive/behavioural analysis from a dataset.
0.09	Methodology unclear, or, speculative inferences, or, presenting a mere concept with no detailed description.

These four criteria created a base for deciding topic relevance of document corpora, allowing for less effort during later grounded theory coding process.

Publication Bias

Despite the efforts to increase objectivity, there was notable publication bias affecting systematic review (Borenstein, et al., 2011; Keele, 2007; Kitchenham, 2004; Müller et al., 2013; Walker et al., 2008). Systematic review dictates the consideration for unpublished works, especially because they often contain negative results that are equally important – “the file drawer" effect. However, already complex methodology justified this bias (entitling grounded theory coding, modeling, and software artefact), and a pursuit in overcoming it would exceed the time assigned for this research.

Research Design (Phase II – Grounded Theory Coding)

The final goal was discovering the relations between relevant factors, and the grounded theory coding was the method for eliciting the variables from the published research. This entailed three stages, typical for grounded theory: *open coding*, *axial coding*, and *selective coding* (Axelsson & Goldkuhl, n.d.; Gibbs, 2011; Soulliere, Britt, & Maines, 2001). The research did not deviate from this practice, however, assigning fuzzy weights to statements in a similar fashion to the systematic review process provided a better results validation.

Using the combined results of all three stages to create a theoretical model, allowed for modeling, theory discovering, and real-time analysis (Gibbs, 2011). A high amount of latent variables justified this research type, highlighted through user-centric perceptions of the tag clouds' usability. Even within the quantifiable variables pool, mutual correlations and the relationship directions and strength usually are difficult to extrapolate (Allam et al., 2012). Adding any complex mathematical modeling would present a challenging task, especially with the absence of a theoretical base. Another reason that made grounded theory suitable for the task is a higher scientific value of the theoretical model than the concrete implementation of a software artefact.

Stage 1.

The coding process itself uses weights for every segment, allowing to represent variables' influence in quantifiable form, employing ordinal 0-100 scale.

Table III-4

Coding weights rationale

Weight	Rationale
90-100	A statement is supported by the actual data – a direct finding.
80-90	Previous finding confirms a statement, but a result of a direct finding. May involve confirmation bias. Does not affect the score significantly.
65-80	A statement is supported by the actual data, the research aimed at confirming or refuting the results from a previous study, but the conditions differ.
50-65	A statement is supported by data, but the finding is not indisputable, e.g. other factors may have affected the result.
35-50	Performing the research on a specific domain, e.g. Delicious, not random sample. This approach ignores user motivations and behaviour, therefore not necessarily universally applicable.
20-35	A statement is supported by an algorithmic or automated testing, simulation.
0-20	Persuasive interpretation, free interpretation, debatable interpretation.

Every assigned code was measured for own value without considering publication's weight, since this process was performed in the previous step. If there was an empirical evidence of some effect or event, it received a maximum mark (100), but interpreting results speculatively, e.g. what *may have* hindered the expected result of an effect, then it received a low score within the 20-30 range. This process did not dismiss opinions, but the impact on the model was significantly lower in contrast to the measurable findings (Table III-4). Using statement weights within one document had the power to undermine the monolithic dominance of certain publications in this area of science, which provided a base for vast number of newer works, and often following insufficiently confirmed interpretations. The decision on exact weight within the range depended on a degree of relevancy to this study. Once the coding was complete, the code weights were multiplied with the source publication weights. This final score was the impact this statement would have on a model's node:

*Total Weight (TW) = Document Weight (DW) * Coded Weight (CW)*

<u>Example</u>*: a statement received a coded weight score of 20, and the document weight is 83, implying a reliable source. The final weight for this statement rounds to 17% impact strength to a relevant model node (dependent variable), reducing the impact which normally would be a lot higher considering the already high publication rating. If the statement is of higher quality, weighting at e.g. 85%, and since a publication threshold sets at 55%, the final statement weight would be smaller than 47%. This way a quality statement always has a higher impact on a model than e.g. a liberal interpretation from a higher-quality publication.*

Stage 2.

The *axial coding* stage focused on identifying and classifying all the user navigational tasks, visual elements, and the motivations behind using tag clouds environment, resulting in definition of both *dependent* and *independent* variables. Statements categorization produced an entire taxonomy containing variables. Although giving advantage to the factors defined as important to user acceptance, the whole process did not dismiss the rest of the factors. This assured that during topical saturation of user acceptance is a guiding premise and not missing supporting evidence.

Stage 3.

Topical saturation was a part of *selective coding*, to further support the subcategories that had a low number of statements and relating them to core categories of *user acceptance* and *visual interfaces*. The topics either marked in the model as variables with inconclusive significance or removed did not achieve significant saturation.

Research Design (Phase III –Model Design)

In this phase, the relations assessment of independent and latent variables, and their effect on dependent variables resulted in forming the relations. The modeling used influence diagrams (decision networks), a method that was the logical choice for two reasons: it allows observation of different variable values, and its ability to use weights assigned in the previous phase. Another alternative considered was SEM - Structured Equation Modeling (Lei & Wu, 2007; Ullman & Bentler, 2003), dismissed because of its steep learning curve, and the overall result would not be achievable in the available time. Although SEM method excels in evaluating the relationships of different factors, it still needed clearly defined variables, which was not possible.

The relationships in findings model were calculated using additive utility functions (AUF), a suitable technique when uncertainty is present. Another technique considered were judgement matrices (Crawford, 1987; Tomashevskii, 2014). However weighing the statements using pre-set criteria dismisses the need for geometrical averages, because the low weights deviation would apply against the same scale. Two values define the independent variables in the model:

Average item or variable significance (IS), a measure of topical saturation around a specific topic, where a higher number of statements results in its greater importance within the model.

Average impact factor (IF), a measure of the independent variable's influence on a dependant variable, considers the ratio of publications and the total number of statements. Without calculating the IF value, any later influence calculations would suffer from publication and/or personal bias, as certain topics and variables had a high number of statements from a fewer number of publications, which could have led to possible inaccurate conclusions.

By using the following formula it is possible to assess the average variable significance:

$$IS = \frac{t}{100} - \frac{c * t}{l^2 * p}$$

where *t* is the total number of coded statements, *c* is the number of statements relating to the variable under observation, *l* is the number of publications from which statements were gathered for the variable under observation, and *p* is the total number of publications used in coding process. This formula increased resulting value precision because of its ability to detect variance in the number of publications.

The *IF* value uses first-stage weighing probability function (Dumais & Foy, 1996; "Statistical Significance," n.d.), and complements *IS* value through ALU function (IS * IF):

$$IF = \frac{c}{r * w_r}$$

where *r* represents the number of relevant statements, and W_r the sum of weights for the relevant statements:

$$W_r = \sum_{i=1}^{r} W_i$$

This statement weighing allows for a corrective action where a smaller set of studies under observation determined that a certain variable has an e.g. low impact on the overall model. This way it supports further experimental efforts, or as a novel idea featuring certain added functionality. Without this weighing calculation, if the impact was positive, the variable would not discarded from the model as insufficiently supported to be effective. If negative, it would present a risk of risking the model stability.

Special Implications and Limitations for the Findings Model

During the design of the model, the derived weights and ratios try to describe the overall influence strength of navigational factors and their effect on other variables. This study aimed at discovering most of the known factors and their relationship direction accurately, which is its true value – organizing factors and their relationships description. The model serves two purposes: supporting future studies of navigational models in an isolation from the rest of the system by pointing out the associated implications. Secondly, providing the outline of factors, which can improve tag cloud features by introducing novel solutions, or steering away from known poor practices. These benefits are the core of the model's composition, allowing for recursive research: if used to design a new system, then it is possible to collect user feedback and subsequently improve on its functionality by focusing on relevant parts.

The quantitative results of this research embedded in the model depict only a single perspective and do not imply universal accuracy, but only an interpretation based on a limited number of analyzed publications. The model provides a theoretical framework that can accommodate any future findings, especially with success or failure of certain navigational factors. The impact effect and ratio values allow for change, and any following research in this area can contribute to its accuracy. The model's purpose is not to define and extrapolate all factors or their respective interrelationships to the highest detail an effort would take years of research, involving significant empirical testing. Apart from this, two main reasons dictate higher abstraction levels: high variable interdependency which leads to increased complexity, and significant opposing views in this research area.

For example, considering tag size the simplest independent variable and a subject to manipulation, the results would vary across different layouts. For circular or rectangular tag cloud it can draw user attention if placed centrally, for lists it can be a keyword (alphabetical-new letter, significant), and for clustered to mark different topical clusters. Therefore, a possible conclusion that tag size not only depends on its physical size but also location and its role can convert it into a dependent variable, influenced by more variables, such as colour, shape, orientation, etc. The resulting model would have little to none of independent variables, rendering it unusable for experimentation where theoretical findings directly lead to practical implementations. Sometimes researchers argue against certain approaches, supported by their own results, therefore a high number of opposing findings would prevent inferencing for most of the relations. For these reasons, this model contains only variables and relations that exhibit a solid amount of certainty, even if at higher abstraction levels. Its intended purpose is to serve as an aid at design time and a summary main factors involved.

Influence Diagrams and Additive Utility Function Primer

The independent variables extracted in the previous stage created the foundation for calculating the interrelation strength using the additive utility function, by pre-weighing them (average significance per element) based on their relative impact to the dependent variables,.

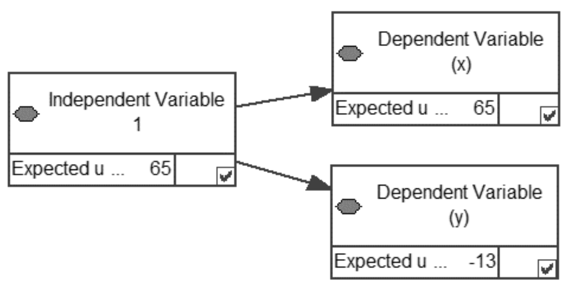
Let X be a set of model outcomes under observation, and R a set of real numbers; then the linear utility function $\int : X \rightarrow R$, with a utility function valid if and only if :

$$x\ , y\ \in X, and\ x \geq y\ \Leftrightarrow\ \int(x) \geq\ \int(y)$$

In the concrete example, a single independent variable has different influence on possible outcomes, although it has the same weight. A different outcome is obvious by noting the different values of the dependent variables (65 for x and (-13) for y).

Figure III-3

Basic utility function – influence on dependent variable



This means that a basic utility function will compare two or more possible dependent variable outcomes based on the preference of a single independent variable on them and expressed as a preference or no preference. In truth tables, this would appear as (Table III-5):

Table III-5

Truth tables (example)

Independent variable weight	Dependent variable x ratio	Dependent variable y ratio
65	1	-0.2
Total	65	-13

Since the function in this simplistic form considers this preference only under one variable when making comparisons, therefore the *additive utility function* expands it:

$(x_1, \dots . x_n), (y_1 \dots . y_n) \in X, (x_1, \dots . x_n) \geq (y_1 \dots . y_n) \Leftrightarrow \int (x_1, \dots . x_n) \geq \int (y_1 \dots . y_n)$, where:

$$X = X_1 * X_2 * \dots . X_n$$

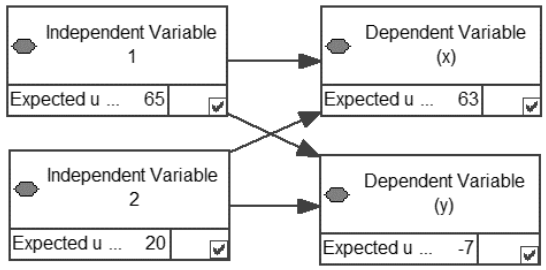
Considering the complex cases where the truth tables would grow in size and become difficult to analyze, it is possible to sum this as:

$$\int (c) = \sum_{i=1}^{n} c_i * w_i$$

This means that function performed over an augmenting set of criteria, as a product of criteria and assigned weights. Following this formula, an independent variable has a certain weight (w_i) and can affect various dependent variables at different influence strength (c_i) (Figure III-4 and Table III-6). The function allows an augmented outcome state of a dependent variable (or two or more possible states of a single variable), which is essential in modeling. By adding the weights of the independent variables and assigning the influence strength on the dependent variable, it is possible not only to view the possible outcomes, but to introduce new findings by focusing on specific topics.

Figure III-4

Additive utility function with multiple independent variables



Representing Figure III-4 by the truth tables, we get (Table III-6):

Table III-6

Truth tables (example II)

	Weight	Dependent variable x ratio	Dependent variable y ratio
Independent. Var. 1	65	1	-0.2
Independent. Var. 2	20	-0.1	0.3
Total		63	-7

This approach allows for coherent model development, where it is possible to note a specific problem based on augmented findings in the literature, described by the effect strength within the model structure. The discarded factors serve as an evidence of the considered topics.

Modeling Process

The first step in modeling process is to identify the first-level dependent variables (DV), derived from open coding stage, and directly addressing the research questions by applying a preliminary classification (Table III-7). Defining the open coding categories was considered most important, because they were expected to assume the state of dependent variables.

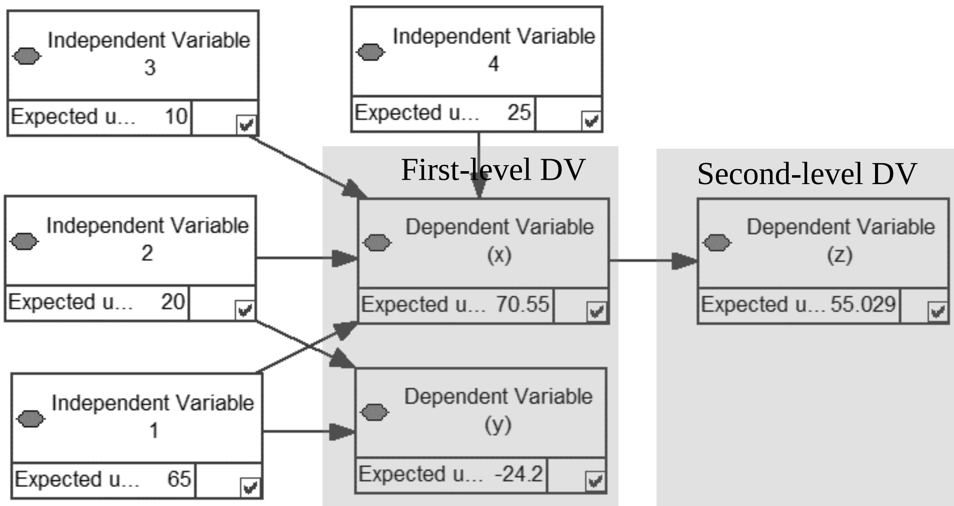
Table III-7
Open coding categories

Variable/code name	Type
User acceptance	Category
Visual layout properties	Category
Visual properties	Category
Cognitive properties	Category
Recommender algorithms	Category
Motivation and behaviour	Category

Defining and analyzing the relationships among the variables created an early model, representing its possible outcomes. As axial coding progressed, the second level of dependent variables were expected to emerge under the direct influence of first-level, and the independent ones. In theory, the level depth of dependant variables have no limit, but lower number reduces the uncertainty of the entire model because of lesser complexity, determined by the number of relationships. After defining all variables, the nature of the interrelationships were assessed, i.e. the directions of mutual influence (Figure IV-5).

Figure III-5

Modeling process



Research Design (Phase IV – Developing the Software Artefact)

Since this research revolves around tag cloud user interface, the artefact functioned as an interface reacting to user interaction, while simulating the content, allowing for model testing under the ideal conditions. The effort involved to model complex dynamic algorithms for visual representation justifies such decision, especially because the goal of the software artefact is to prove model's ability to guide the design decisions and visualize the findings.

Functional Testing of the Artefact

The software artefact testing focused on its functionality, i.e. how accurately represents the theoretical model, which was sufficient considering already high research complexity. The testing paradigm used during the development process was *feature inspections* (Jackson, 1997; Nielsen, 1993, 1994), an approach that met all the functional requirements. Conducting feature inspections at various development stages ensured the artefact features accurately interpreted theoretical model's requirements of inputs, and outputs - the early stages of testing followed Nielsen's heuristics guidelines (Nielsen, 1993).

Tools

The choice of coding software was MAXQDA, a software designed specifically for grounded theory coding, mostly because of the flat learning curve and a fitting price model. The modeling was performed in GeNIe, a freeware Bayesian networks software, also supporting additive utility functions.

Although considering many development environments alternatives, such as PHP, CSS, HTML, JavaScript, etc., the choice for this implementation was Microsoft Visual Studio Blend. Its main advantage was the ability to create rich user interface first, while inserting all the programming codes into Microsoft Visual Studio in C# programming language. This practice allows for easy addition of any recommender or visual classes and modules, in a programming environment that is familiar worldwide. Another benefit is its portable, platform-independent nature through Silverlight plugin, which is further customizable within a web page (size, position, variable size, user interaction styles, etc.) Finally, the decaying trend of JavaScript's popularity because of the security issues aided in shaping this choice – modelling tag clouds in an environment not in danger of becoming obsolete, avoiding portability issues in the future.

Data Sources

During the systematic review stage, the research focused on identifying and documenting the variables, methodologies, and conclusions from the literature, while the material selection had *priority* based on these criteria:

- Peer-reviewed publications
- Relevant books on social systems and user interfaces.
- Implemented solutions publicly accessible.

Research Limitations and Considerations

Preliminary literature analysis showed that some publications have biased approaches, often when presenting propriety systems; still, the findings of these works are important for this study. The assigned weights aim to objectify these types of publications by assigning lower values, however, it was not an accurate approach because of personal bias.

Since the navigational cues, user goals and visual elements vary in quantity, one of the challenges was to set the logical limit to the number of possible combinations and forms of these factors, which resulted in high model abstraction.

Although atypical interfaces catering to a narrow application domain (e.g. university library) can provide useful insight, the research goal was to seek the solution as close to one of “social nature”. This implied widening the application domain, at the expense of potentially reduced functionality.

The software artefact’s absence of the recommender algorithm did not allow for user testing.

By accepting the limitations, the research scope reduced and focused, which increased the overall feasibility and allowed the research to gain tangible value. The intention to address these limitations as a part of PhD program research dictated most of the scope adjustments.

Definitions of Key Terms

Fuzzy weights – quantitative expression/taxonomization of linguistic criteria.

Unique tags – tags that have low semantical and/or cognitive connection to other tags.

Multi-word tags – tags of higher complexity, especially in semantical and computational sense (e.g. white, house versus white house).

Chapter IV - DATA ELICITATION FINDINGS

This chapter summarizes the literature selection processes and the results gained through systematic review, followed by a discussion on grounded theory early findings. The next chapter contains a detailed analysis of every topically saturated variable.

Systematic Review Findings

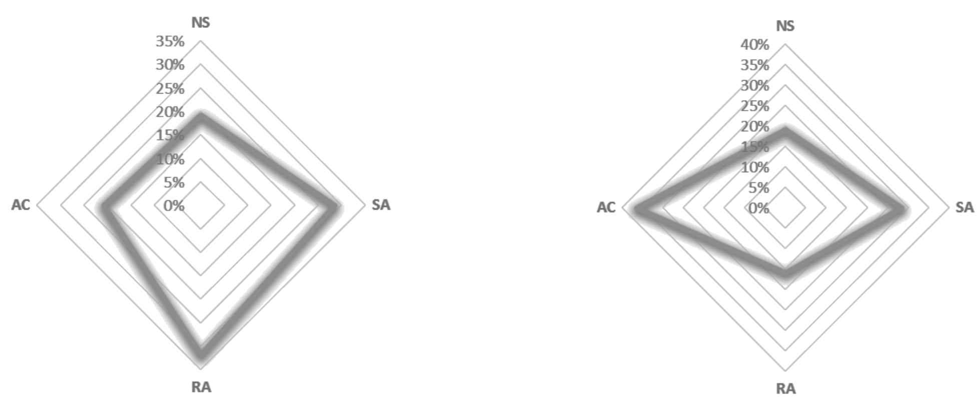
Based on the preliminary study, the set of criteria for the literature selection was:

- User acceptance
- User motivation
- Recommender algorithms
- Visual properties

Based on these topic, 187 relevant scientific papers and books were collected, catalogued and assessed, satisfying at least one the criteria.

Figure IV-1

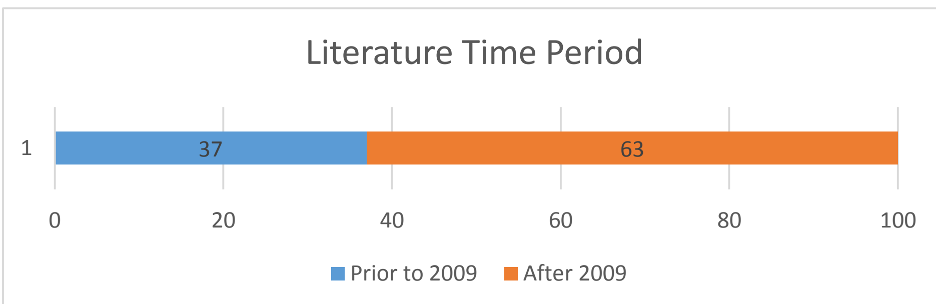
Topical relevancy of selected literature prior and after correction



The documents could have had from one to four states, since the previous step ensured none deviated from the tag cloud topic. The preliminary analysis showed that recommender algorithm was a dominating topic, followed by surveys and assessments, which was not desirable (Figure IV-1, left). Removing several publications remedied this problem, targeting the ones addressing recommender algorithms. DW filter assessed newly added literature (Chapter IV) and iteratively analysed it using the second filter until reaching the satisfying ratio.

Figure IV-2

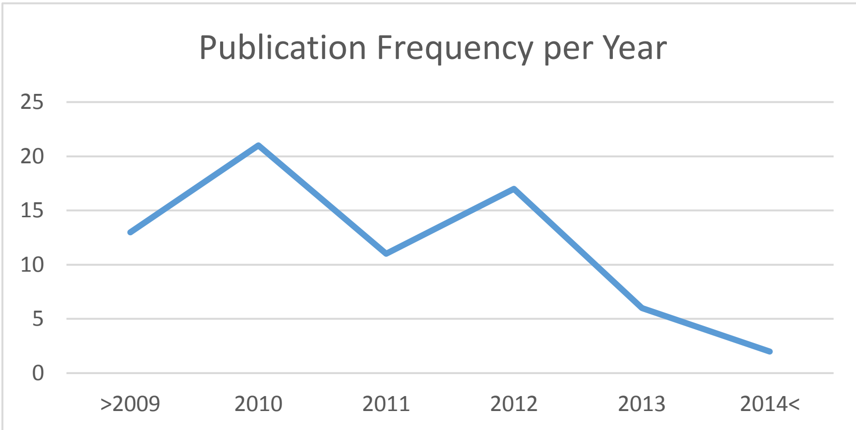
Literature currency



The newly selected document corpora leaned significantly towards the user and cognitive assessment (37%), followed by surveys (29%) and containing the lowest amount (16%) of recommender algorithm topics (Figure IV-1, right). Literature currency stabilized at 63% for post-2009 publications, proving the selected method not excessively restrictive (Figure IV-2). Even with currency filter adjusted to favour recent publications, there was a noticeable decline in the quantity of literature (Figure IV-3). This can is attributable to various reasons, one being fading interest of research community in tag clouds given the successful saturation of content-richer recommender systems (e.g. Amazon), over decreasing user interest, to the probable inability to develop the tag cloud framework resulting in an intelligent and intuitive navigational tool.

Figure IV-3

Selected publication frequency by year



The most probable cause is the implementational segregation, which stimulated developing unique theories, resulting in research taking multiple vectors leading to dysfunctional and non-unified theoretical framework. In other words, continuous innovation as opposed to theory building. Isaac Newton's famous quote: "If I have seen further it is by standing on the shoulders of giants" cannot accurately describe this research area.

Once topically narrowed, the applied threshold levels assessed the quality of the literature selection content. Since the literature already had meaningful weights, it was easy to adjust the threshold level by considering the intended 100 publications. With that criteria, the value settled at 0.55 (55%), which expelled the publications with lower quality for the following phase of grounded theory coding (Table IV-1). Not surprisingly, the selection process confirmed that most of the short publications had a weak methodology or methodology description.

Table IV-1

Literature weighting process (table trimmed to fit)

IDENTIFIER	TITLE	DATE	PAGES	CURRENCY WEIGHT	SIZE WEIGHT	METHODOLOGY RELEVANCE	TOTAL WEIGHT
Mishra2012	A theoretical a	Jan-12	4	0.2	0.06	0.1375	0.40
Mezghani20	A user profile	Jan-12	8	0.2	0.22	0.2750	0.70
Filho2009	A visualization	Jan-09	4	0.1	0.06	0.1375	0.30
Kay2006	Adaptive Hype	Jan-06	488	0.1	0.25	0.4125	0.76
Halvey2007	An assessmen	May-07	2	0.1	0.06	0.4125	0.57
Oosterman	An empirical c	Nov-10	8	0.2	0.22	0.2750	0.70
Klaisubun20	Behavior Patte	Nov-07	4	0.1	0.06	0.4125	0.57

For example, the publication under identifier “Halvey2007” gained a low score for currency and size, but the methodology and the results were sound, at the expense of introduction and brief literature review. The influence of methodology factor to the outcome proves the proper operation of grading scales and ratios. In 4 cases, adding up to 5% to a document score allowed threshold crossing, as those publications were relevant even though they scored low, which introduces low intentionally produced personal bias.

Grounded Theory Topical Discovery

The final document selection was imported into the coding software, assigning the initial topical codes to support the research questions:

Navigational properties – all the navigability factors (e.g. transparency, tag relationships), tag types, typical user tasks in a tag cloud environment, etc.

The second research question was broken down into two major code categories:

Visual properties – tag size, tag colour, contrast, tag length, lists, positions, etc.

Layout properties – cloud or clustered layouts, alphabetical, gaze distribution, etc.

The categories that related to the third research question:

User acceptance – user empowerment, information disclosure, different perspectives of system designers and the users of the system, visual preferences, enjoyment, etc.

Motivation and behaviour – social capital, domain experts versus domain novices, online vocabulary differences, the reasons for content generation, communities, etc.

Cognitive – eye fixations, tag manipulation, use plurality, tag reuse, tag orientation, tag re-findability, etc.

Two more categories further strengthened the model, as found relevant in the preliminary literature review:

Recommending - guidelines, approaches, novel designs, tagging frequency, tagging periods, user roles division, etc.

Performance – simplicity, tag reusing, autocomplete, platforms (e.g. mobile), efficiency, effectiveness, etc.

During statements elicitation process 15 documents were removed, found to have a low coding value. For example, a system involving novel interface only tested with automation without providing an insight into user perspective, or a presentation of a recommender algorithm with no evaluation had little to offer. Since the coding process was iterative, sometimes it was possible to sense the approaches that had failed in the past, and those publications have had little to offer. From the previous example, a high-performing system does not imply user acceptance. The remaining 15 documents, added during the selective coding, targeted areas where topical saturation was inadequate, and stabilizing document corpora at previously determined quantity of hundred. Appendix A contains the complete list of the analyzed literature (Appendix A).

Table IV-2

Code quantity per category

Variable/code name	Type	Quantity
User acceptance	Main category	265
Layout properties	Main category	280
Navigational properties	Main category	84
Visual properties	Main category	111
Cognitive properties	Main category	144
Recommender algorithms	Main category	203
Motivation and behaviour	Main category	365

Throughout the axial coding, a number subcategories emerged as topics naturally became more specialized. Certain subcategories integrated with more relevant ones, or reassigned to other main categories where topics diverged from the initial. An example of this is “Grid and sorting” subcategory, which was at first under “Layout properties” main category. That cognitive aspect diverged from visual, justifying their reassignment to “Scanning patterns” under “Cognitive” main category.

Probably the most interesting discovery emerged during the selective coding, during further saturation of insufficiently supported topics, with 15% of the overall literature corpora available for the task. After revising over 400 abstracts and 150 publications, there was little improvement to already unsupported topics: most publications focused on well saturated ones, resulting in improvements ranging from 0% to maximum of 15% per subcategory. This rather unexpected outcome, however mitigated by already planned modeling mechanism: most topics could enter the literature model, pointing to the areas for further research. Another reason for low saturation stems from earlier studies with solid conclusions, resulting in high certainty of particular operation or behaviour, e.g. tag size.

From the beginning of coding process, there were several different perspectives on usability and purpose of tag clouds, but three created significant pivotal points for many studies:

Functional perspective – in which the system designers believe systems' functional performance increase will positively reflect on the navigational quality. These efforts typically focus on improving the system performance or increasing the recommendation precision. Such systems mostly rely on users' analytical abilities and knowledge to distinguish the path to the desired goal.

Cognitive perspective – an approach in which perception of visual elements quality has the most influence on users' navigational choices. This research type focuses on structuring the navigational paths using visual cues, and guiding the users of the system by estimating favourable choices. These systems aim to guide the user by suggesting a navigational path.

Motivational and behavioural perspective – focused on research why users tag, tags use, and attempt to explain complex relationships within the tagging communities. Although the results of these studies rarely result in novel designs, they provide significant critical and innovative support to the existing ones. If presenting novel designs, they are mostly complementing the functional perspective. Despite much research effort on user behaviour and motivation, it is a topic rarely considered during design.

The following factors contributed to forming the additional dependent variables:

Aesthetics – the preliminary literature review clearly revealed that users are sensitive to this aspect of tag clouds, and essentially there would be no user acceptance without it. In contrast, a vast majority of existing tag implementations do not sufficiently focus efforts on the beauty of tag clouds. In a pleasantly looking interface, it is much easier to intentionally present the navigational cues that users need, therefore this variable partially answers the second and third research question.

Motivation and behaviour was further separated in two dependant variables:

Motivation – this variable reinforces both user acceptance and recommender algorithms, as it gives insight into the reasons for tag cloud use and tagging, and motivations for implementing certain features. Apart from user acceptance, this variable is the key to modeling the system in such a way to increase user attraction and lessen repellence.

Behaviour – this variable complements motivation in many aspects. If we can come close to understanding how users behave during interaction with this social software, it is possible to partially reduce the negative effects at design time. The design should be oriented in such a way to perform a positive influence. By studying this variable it was possible to find out the scope of tag clouds, what tasks tag clouds are suitable for, and their overall limits.

The early stages of literature review revealed the importance of tag cloud layouts, but also their variable performance, both visual and technical. For example, it is nearly impossible to convey a salient semantical relation in a chaotic cloud layout. It is difficult to reduce the interface size in clustered layout, to make it a non-dominant feature across the entire user interface. Therefore, examining the most dominant layouts in an isolation was necessary, which resulted in three more dependent variables:

Cloud – represents circular, rectangular and variable shape tag cloud layouts, since those share the same features. This layout is the most commonly used in social networking systems.

Sequential and Lists – layouts that resemble the traditional tables, most frequently used for alphabetical tag ordering and where the value of conveyed information is high.

Clustered – layouts that focus on representing the semantical or ontological structure of terms within a tag cloud, by grouping tags in clusters around several central topics.

Since model creation commenced immediately after axial coding it became obvious that not all topics can assume role of a variable (Table IV-4). However, those retained an explanatory function and either served as a supporting argument when explaining relations within a model. Although ideally topics would have transitioned to dependent or independent variable, sometimes there was a minimum number of publications supporting the topic, even if there was a high quantity of statements per topic. These topics were labelled as prominent areas for research. Some topics experienced similar condition, however because of the overall topical saturation: an adequately researched topic resulted in solid conclusions, however lacked other studies' confirmation.

Table IV-3

List of all categories and topics

Category	Topics	Number of Statements	Number of Publications
Motivation and Behaviour	Domain models	28	8
	User modeling	10	5
	Personalization	20	7
	Communities of interest	21	13
	Vocabulary	35	10
	User types	35	10
	Tag influence (collective knowledge)	64	21
	Tagging imitation	15	12
	Domain expertise	31	14
	Social and cultural capital	15	6
Recommender	Tag filtering	11	4
	Semantical efficiency	20	7
	Cold start	8	7
	Administrative control	5	4
	Tag currency	7	6
	Multiple views	31	21
	Tag ambiguity and spamming	9	7
Cognitive	Tag qualities	44	11
	Learning	26	9
	Associative	45	10
	Cognitive effort	22	16
	Scanning patterns	7	5
Navigational Properties	Tag types	10	3
	Reversibility	2	2
	Tag reuse	22	12
	Topical narrowing	16	10
User Acceptance	Transparency	42	12
	Social insight	21	8
	Aesthetics	40	15
	Functionality	99	33
	User control	29	18
	Readability	11	8
Layout Properties	Semantics	20	10
	Folksonomies	24	10
	Lists	13	9
	Alphabetical	16	12
	Interface functionality	47	17
	Clustered	34	14
	Grid and sorting	46	15
	White spaces	13	7
	Layout size	15	10
	Grid and sorting	22	11
Visual Properties	Font weight	6	3
	Tag contrast	9	5
	Tag shape	4	1
	Tag orientation	5	2
	Tag colour	15	11
	Tag length	14	5
	Tag size	43	23

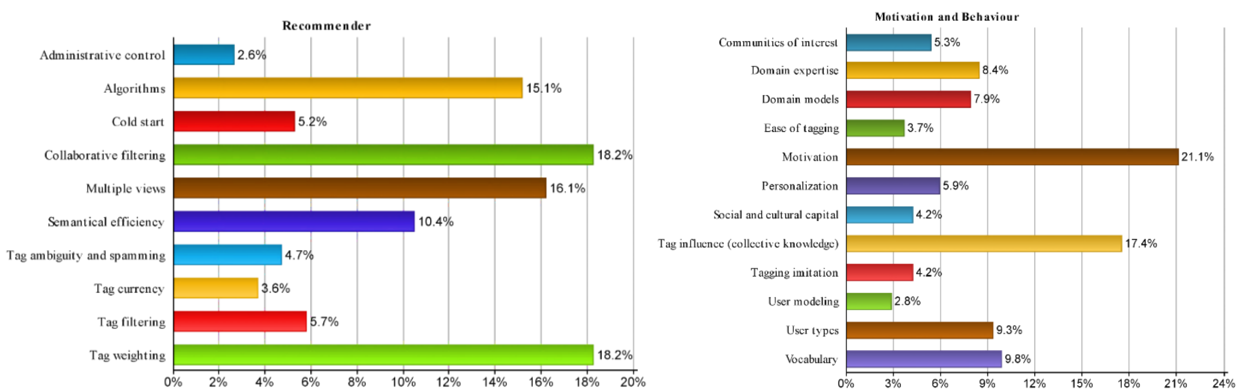
Table IV-4

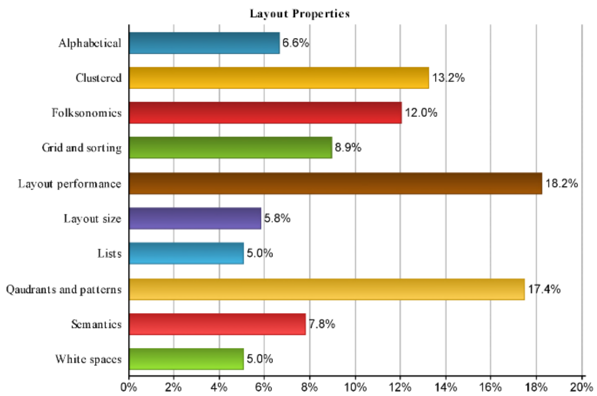
List of selected independent variables

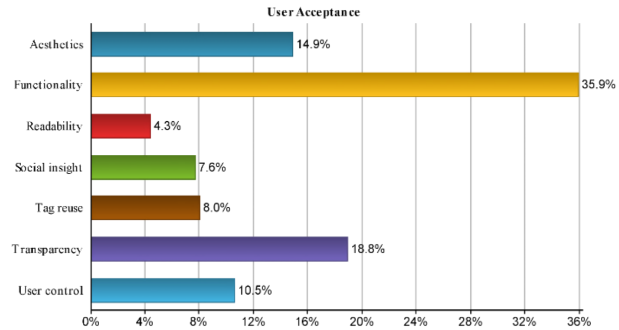
Variable name	
Tag size	Cognitive effort
Tag length	Semantical efficiency
Tag colour	Folksonomies
Tag contrast	Tag filtering
Font weight	Cold start
Readability	Tag currency
Inter-tag spaces	Tag influence
Layout size	Tag reuse
Grid and sorting	Tag qualities
Scanning patterns	Vocabulary
Semantical visualization	Domain models
Readability	Communities of interest
User control	Personalization
Social insight	Domain expertise
Transparency	Multiple facets

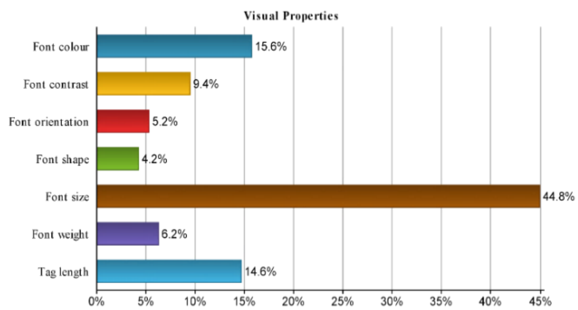
Figure IV-4

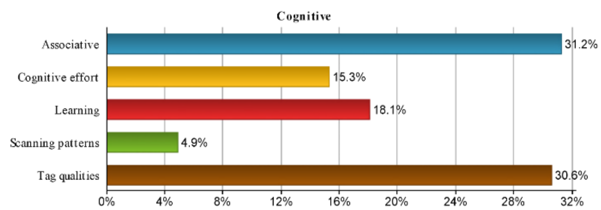
Topical representation graphs

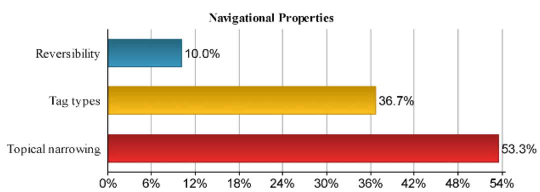



Layout Properties
Alphabetical 6.6%
Clustered 13.2%
Folksonomics 12.0%
Grid and sorting 8.9%
Layout performance 18.2%
Layout size 5.8%
Lists 5.0%
Qaudrants and patterns 17.4%
Semantics 7.8%
White spaces 5.0%
0% 2% 4% 6% 8% 10% 12% 14% 16% 18% 20%


User Acceptance
Aesthetics 14.9%
Functionality 35.9%
Readability 4.3%
Social insight 7.6%
Tag reuse 8.0%
Transparency 18.8%
User control 10.5%
0% 4% 8% 12% 16% 20% 24% 28% 32% 36%


Visual Properties
Font colour 15.6%
Font contrast 9.4%
Font orientation 5.2%
Font shape 4.2%
Font size 44.8%
Font weight 6.2%
Tag length 14.6%
0% 5% 10% 15% 20% 25% 30% 35% 40% 45%


Cognitive
Associative 31.2%
Cognitive effort 15.3%
Learning 18.1%
Scanning patterns 4.9%
Tag qualities 30.6%
0% 4% 8% 12% 16% 20% 24% 28% 32%


Navigational Properties
Reversibility 10.0%
Tag types 36.7%
Topical narrowing 53.3%
0% 6% 12% 18% 24% 30% 36% 42% 48% 54%

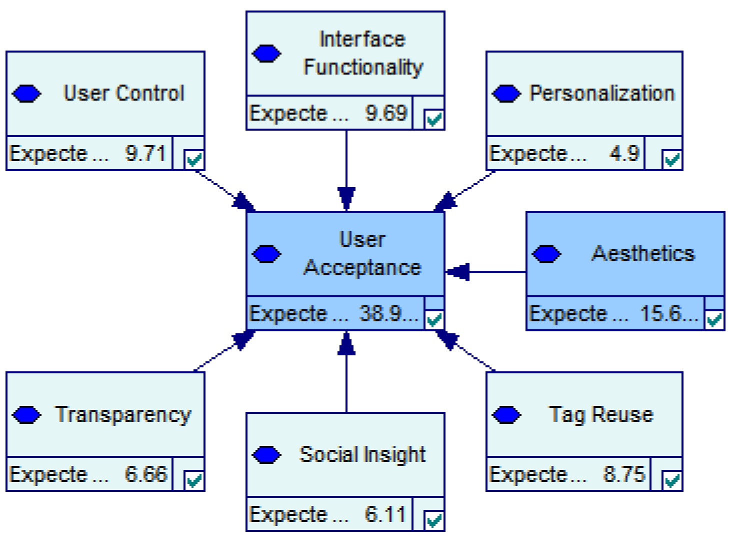
Chapter V - THEORETICAL MODEL AND LITERATURE ANALYSIS

This chapter describes the structure and evolution of the findings model, all the obtained factors and their interrelations. The theoretical model visualizes the relevant implications for tag cloud design.

Theoretical Model Overview

Since the pivotal points of this thesis are user acceptance and tag cloud visualization, modeling followed the same logic by setting them as the main topical categories influenced by several factors drawn from the analyzed literature (Figure V-1).

Figure V-1.
Influential topics on user acceptance (light blue – supporting topics, dark-blue – categories).

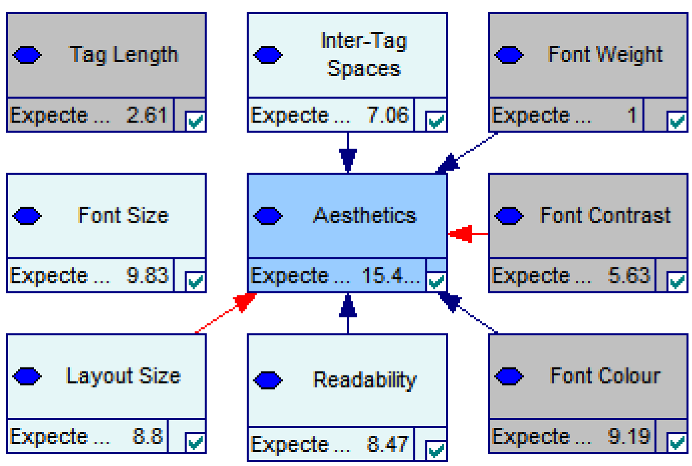


The numbers in the lower right corner represent topical significance, extrapolated from the quantity of publications supporting the topic and the number of relevant statements ratio against the overall document and statements corpora. A higher value represents well supported topic. However, this value does not reflect confidence levels accurately, since occasionally a fewer quality publications provided many statements, resulting in publication bias. By employing impact force values mitigated this problem, gained from the ratio of relevant statements and their weights against total number of statements and weights. As predicted, in minor number of cases there was an exception to this rule, when a topic had a high significance but low confidence level because of the conflicting research results.

Both tag cloud and conventional user interface design practice showed *aesthetics* to be crucial to user acceptance (Figure V-2). Note that tag length and tag size have no connections as expected, since their impact on aesthetics was inconclusive.

Figure V-2.

Aesthetics category with the associated topics (light blue – supporting topics, dark-blue – categories, gray – low impact, red arrow – negative impact).

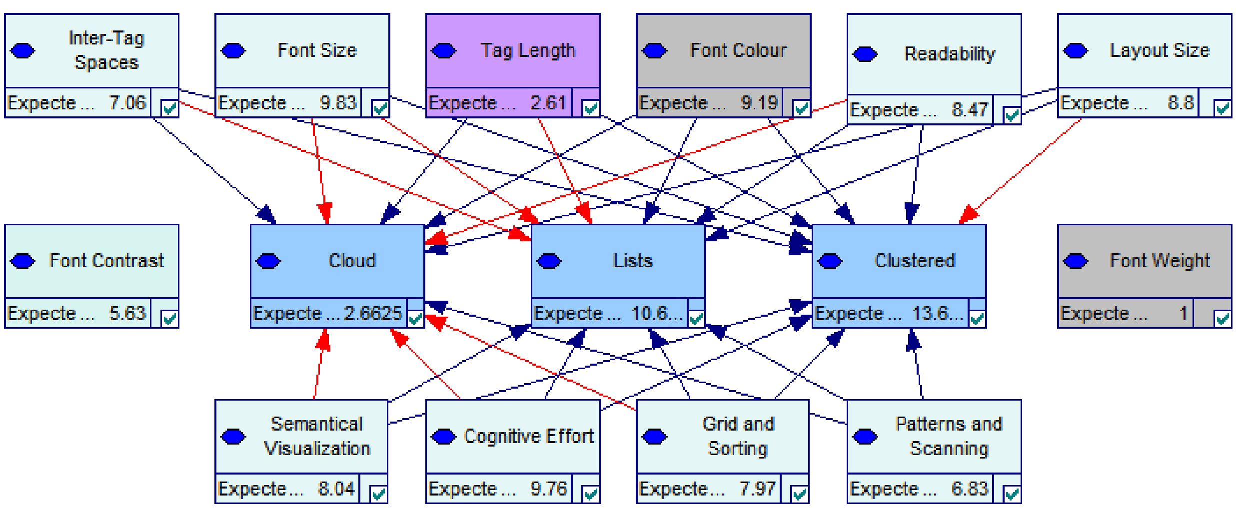


The gray colour marks independent variables with little influence that can be employed liberally as a navigational cue, and red arrows mark a negative impact on a dependent variable;

purple colour indicates insufficiently researched topics and uncertain effect at design time. The preliminary literature review revealed the need for layout consideration when designing user interfaces for tag clouds. As described in the previous chapter, three main layout types emerged as dominant by being the most frequently used, assessed from two perspectives: visual and operational/technical (Figure V-3). The technical perspective assesses the suitable role for each layout type, as not all equally adhere to different application purposes or domains. Although there are other variants of tag clouds, e.g. containing photos or domain-specific ones, they steer away from the traditional tag cloud and considered special cases, therefore out of scope of this research.

Figure V-3.

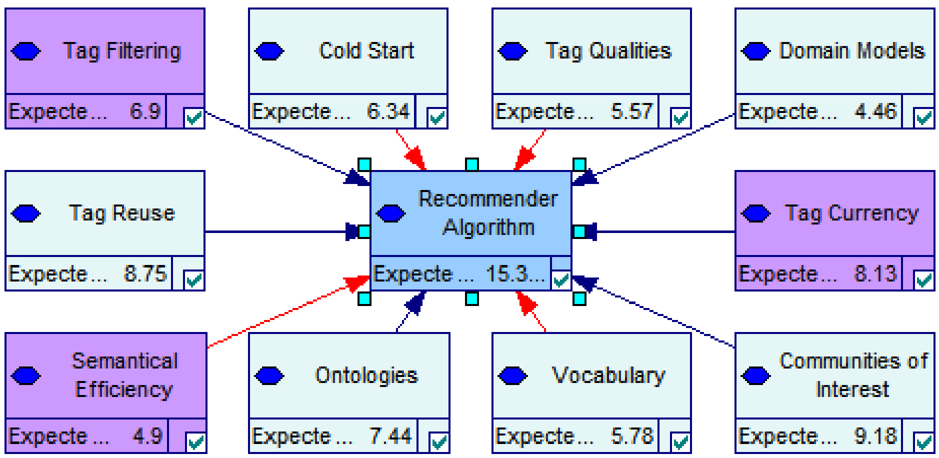
The assessment of different layouts (light blue – supporting topics, dark-blue – categories, gray – low impact, red arrow – negative impact, purple – insufficiently researched topic).



Unsurprisingly, the topics revolving around *recommender algorithm* presented the highest level of inconclusiveness, with about 30% required further research (Figure V-4). It was an interesting finding that major topics were often revisited, yet investing little research effort into exploring smaller or indirect factors valuable to system designers. For example, measuring the efficiency of the semantical relationships against user acceptance; instead, the focus was mostly on theoretical or computational benefits.

Figure V-4.

Recommender algorithm and its associated topics (light blue – supporting topics, dark-blue – categories, red arrow – negative impact, purple – insufficiently researched topic).

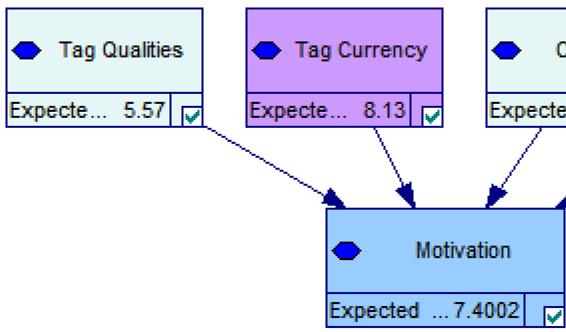


Those few studies that performed such research showed no significant improvements that were worth considering. This rather explicit approach was necessary to avoid employing a myriad of features of questionable value, but instead find an agreed upon set aligned in such a way to use their potential to the fullest.

Motivation was analyzed from systems design aspect, i.e. how to increase user motivation through either implementing a supporting feature or avoiding poor practises (Figure V-5). It was surprising to discover that few studies addressed the increase in motivation by setting up and overseeing a feature, other than measuring whether users liked or disliked it: even if users exhibited such a preference, that implies no increase in motivation over time compared with systems deprived of it. This resulted in fewer factors that can be manipulated to increase motivation.

Figure V-5

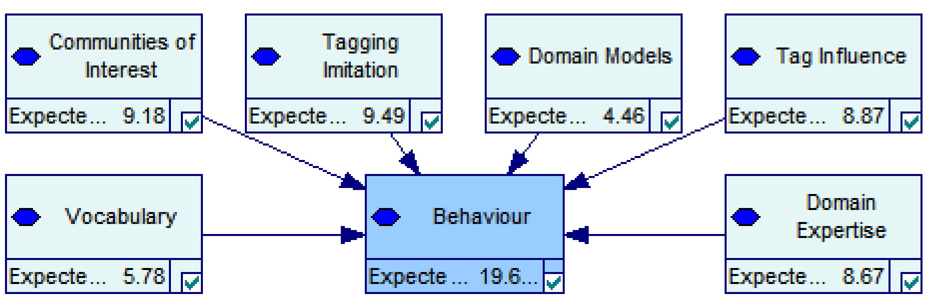
Motivational topics through functional perspective (light blue – supporting topics, dark-blue – categories, gray – low impact, purple – insufficiently researched topic).



The final category/dependent variable was user behaviour, an important topic as it directly supports user acceptance (Figure V-6).

Figure V-6.

User behaviour in tagging systems, functional perspective.



Although it achieved the highest saturation compared with other categories, this is the least cited topic in the reviewed literature focused on systems design aspect, providing little consideration for user behaviour in tag clouds at design time. To partially correct this, apart from functional implications, a section of this chapter discusses on selected topics on relevant user behaviour.

Figure V-7 depicts the entire model, with all of its relations, and Table IV-1 contains the supporting values of independent variables' significance and impact strength on dependent ones (Figure V-7 and Table V-1).

User Acceptance

Although there is nearly an unlimited number of factors and their combinations that can affect user acceptance, this study focused on those that are critical, setting a starting point for successful design practices of tag clouds. The biggest challenge was overcoming the established paradigm in the field which assumes that more accurate recommendation results, or results better tailored to the user of the system will lead to better user acceptance. It became obvious such paradigms stem from older studies because of topical novelty at the time, when computational accuracy in this sense was still developing, hence justifying those efforts. Another reason behind this computational perspective is novel concept of social networking, allowing scientists specialized in data mining to change their research direction to a tempting and almost inexhaustible pool of very dynamic data. There is a clear evidence in newer studies that user preferences and interfaces should be a priority over accurate recommending, as accuracy does not imply user acceptance (Skoutas & Alrifai, 2011; Viegas et al., 2009). This hypothesis follows real-life users' habits: few examine beyond the appealing exterior, for example, a typical driver spends 99% of the time directly behind a steering wheel, yet the exterior and shape of a car is one of the main selling points advertised.

Figure V-7.

The entire findings model (light blue – supporting topics, dark-blue – categories, gray – low impact, red arrow – negative impact, purple – insufficiently researched topic).

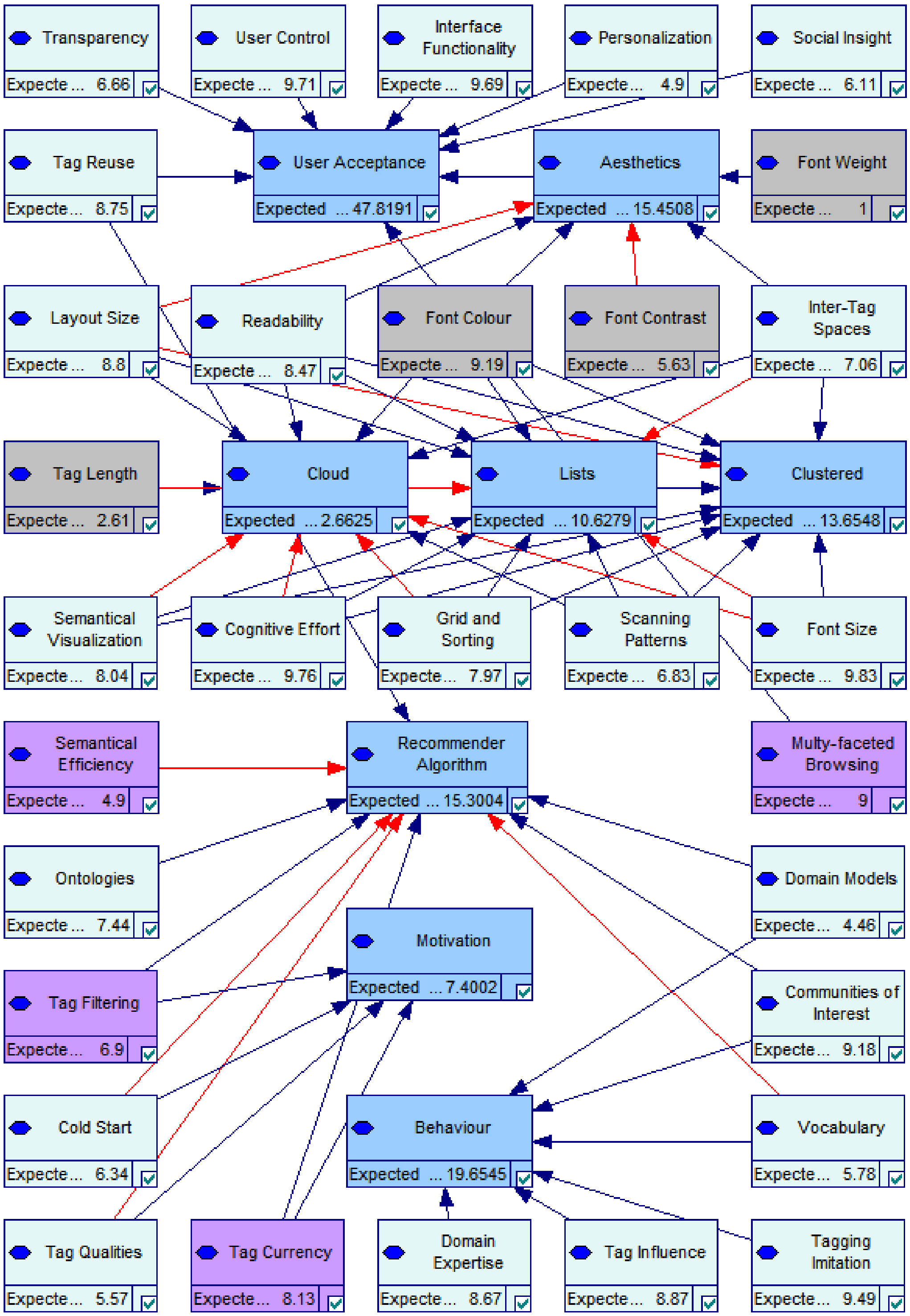


Table V-1.

Impact strength of the independent variables. TS – total statements, RS – relevant statements, TW – total weight, RW – relevant weight, IF – impact factor, IS – item significance.

Variable name	Node relation	TS	RS	TW	RW	IF	IS	Total
Aesthetics	User acceptance	41	37	3605	3303	0.9	15.6	14.04
Cognitive effort	Clustered	22	10	1206	557	0.21	9.76	2.05
Cognitive effort	Cloud	22	5	1206	-235	-0.04	9.76	-0.43
Cognitive effort	Lists	22	7	1206	614	0.16	9.76	1.58
Cold start	Motivation	10	8	1076	-590	-0.44	6.34	-2.78
Cold start	Recommender	10	8	1076	486	0.36	6.34	2.29
Communities of interest	Behaviour	19	17	1978	1345	0.61	9.18	5.59
Communities of interest	Recommender	19	8	1978	633	0.13	9.18	1.24
Domain expertise	Behaviour	30	20	2382	1563	0.44	8.67	3.79
Domain expertise	Recommender	30	11	2382	819	0.13	8.67	1.09
Domain models	Behaviour	29	20	1453	1453	0.69	4.46	3.07
Domain models	Recommender	29	11	1453	1453	0.38	4.46	1.69
Ontologies	Clustered	26	8	1436	531	0.11	7.44	0.85
Ontologies	Cloud	26	5	951	-357	-0.07	7.44	-0.54
Ontologies	Lists	26	6	951	141	0.03	7.44	0.25
Ontologies	Recommender	26	12	951	835	0.41	7.44	3.01
Tag colour	All layouts	15	13	810	698	0.75	9.19	6.86
Tag contrast	Aesthetics	9	7	392	-303	-0.60	5.63	-3.38
Tag size	Cloud	43	14	1740	-213	-0.04	9.83	-0.39
Tag size	Lists	43	11	1740	190	0.03	9.83	0.27
Tag size	Clustered	43	15	1740	649	0.13	9.83	1.28
Font weight	Aesthetics	6	5	227	180	0.66	1.00	0.66
Grid and sorting	Cloud	23	11	1545	-306	-0.09	7.97	-0.76
Grid and sorting	Lists	23	13	1545	649	0.24	7.97	1.89
Grid and sorting	Clustered	23	10	1545	514	0.14	7.97	1.15
Interface functionality	User acceptance	43	38	3675	3282	0.79	9.69	7.65
Inter-tag spaces	Cloud	13	7	850	-278	-0.18	7.06	-1.24
Inter-tag spaces	Lists	13	2	850	-110	-0.02	7.06	-0.14
Inter-tag spaces	Clustered	13	5	850	266	0.12	7.06	0.85
Inter-tag spaces	Aesthetics	13	5	850	-340	-0.15	7.06	-1.09
Layout size	Clustered	15	7	915	493	0.25	8.80	2.21
Layout size	Cloud	15	4	915	146	0.04	8.80	0.37
Layout size	Lists	15	4	915	-387	-0.11	8.80	-0.99
Layout size	Aesthetics	15	5	915	-457	-0.17	8.80	-1.46
Multiple facets	User acceptance	31	26	2187	1916	0.73	9.00	6.61
Multiple facets	Recommender	31	7	2187	571	0.06	9.00	0.53
Personalization	User acceptance	21	14	873	659	0.50	4.90	2.47
Readability	Aesthetics	12	9	870	660	0.57	8.47	4.82
Readability	Cloud	12	9	1404	-634	-0.34	8.47	-2.87
Readability	Lists	12	5	1404	-416	-0.12	8.47	-1.05
Readability	Clustered	12	4	1404	354	0.08	8.47	0.71
Scanning patterns	Cloud	26	9	1808	-481	-0.09	6.83	-0.63
Scanning patterns	Lists	26	11	1808	722	0.17	6.83	1.15
Scanning patterns	Clustered	26	10	1808	605	0.13	6.83	0.88
Semantical efficiency	Recommender	40	15	3045	1089	0.13	4.90	0.66
Semantical visualization	Clustered	20	9	2493	900	0.16	8.04	1.31
Semantical visualization	Cloud	20	7	2493	-607	-0.09	8.04	-0.69
Semantical visualization	Lists	20	5	2493	449	0.05	8.04	0.36
Social insight	User Acceptance	21	10	889	750	0.40	6.11	2.45
Tagging imitation	Recommender	15	9	1222	665	0.39	9.49	3.71
Tag currency	Motivation	10	6	647	321	0.30	8.13	2.42
Tag currency	Recommender	10	4	647	326	0.20	8.13	1.64
Tag filtering	Motivation	11	8	782	547	0.51	6.90	3.51
Tag filtering	Recommender	11	4	782	235	0.11	6.90	0.75
Tag influence	Behaviour	44	33	3364	2575	0.57	8.87	5.09
Tag length	Cloud	14	12	950	550	0.50	2.61	1.29
Tag length	Lists	14	12	950	-245	-0.22	2.61	-0.58
Tag length	Clustered	14	12	950	790	0.71	2.61	1.86
Tag length	Aesthetics	14	5	950	240	0.09	2.61	0.24
Tag qualities	Motivation	44	25	3125	1665	0.57	5.57	1.69
Tag reuse	Recommender	22	7	1715	519	0.10	8.75	0.84
Tag reuse	User acceptance	22	15	1715	1196	0.48	8.75	4.16
Transparency	User acceptance	52	31	4126	2750	0.40	6.66	2.65
User control	User acceptance	27	27	1923	1354	0.70	9.71	6.84
Vocabulary	Behaviour and recommender	35	24	2003	1778	0.61	5.78	3.52

Users also become attracted to the aesthetics of non-functional or counterproductive systems, such as fast foods, weapons, poor media content, etc. Microsoft has not significantly upgraded Windows NT technology for over 20 years, but visually redesigning each new operating system version. For tag clouds, it does not imply the system should be only aesthetically appealing and not efficient, but achieving an acceptable balance by prioritizing user perspective. Even the best functioning system will fail its intended purpose if it cannot attract users, however an aesthetically pleasing and erroneous one will repel users after a certain time. The analysis of studies under observation revealed eight major factors that can be manipulated and contribute to user acceptance of tag clouds with higher certainty (Figure V-2):

- Aesthetics
- Interface functionality
- Social insight
- Personalization
- User control
- Transparency
- Tag reuse
- Multi-faceted browsing

Aesthetics obtained item significance (IS) of (15.6), attributable to a cumulative ISs of the individual elements that affect it, some of which had negative influence (Figure V-2). This was an expected result – a single visual element with an extreme value can hinder visual appeal and repel the user more than poor recommendation quality. It also got a high impact force (IF) of

(0.9), a surprising finding considering that most studies from preliminary literature review focused on the recommendation accuracy (Liang et al., 2009; Milicevic et al., 2010; Zanardi & Capra, 2008b). The explanation for this result is in literature selection method, which lessened the impact of the older studies, and setting the bias towards user acceptance. In practical application this means any tag cloud design should employ a top-down approach, where interface appeal and functionality should have higher priority over recommender algorithm design (Lohmann et al., 2009; Viegas et al., 2009). This approach is likely to induce restrictions or mitigate certain computational dogmas, and introduce innovation. Aesthetics contribute to tag clouds' effectiveness and understandability (Lohmann et al., 2009; Waldner et al., 2013), probably because of increased user attention and the motivation to learn, and the ease of use (Allam et al., 2012). Although aesthetics is natively vague, user interface design should aim to be neither excessively dull nor too extreme with implementation of visual features. The aesthetical aspect is analyzed in detail later in this chapter.

Interface functionality had the highest IS (9.69), and a second highest score in the group for IF (0.79), since many studies covered the topic and supported it by empirical research. When compared to traditional interfaces, tag clouds have better user reception for simpler tasks, mainly for their intuitive interface and flatter learning curve (Ravendran, MacColl, & Docherty, 2012). Users who rate aesthetics high, apply nearly proportionally rating to the interface functionality (Lee et al., 2010; Viegas et al., 2009), which must be considered. A substantial evidence suggests a simplistic form of tag clouds as inappropriate for any complex data visualization (Dörk et al., 2013; Emerson, Churcher, & Deaker, 2013; Gwizdka & Cole, 2013), and many attempts failed in achieving user satisfaction, even when recommender algorithms performed well in simulations

(Diaz et al., 2009; Gwizdka & Cole, 2013; Lohmann et al., 2009; Ravendran et al., 2012). There is much disagreement on best-performing layout, being difficult to measure because of its dependency on recommendation quality. However, most researchers do agree that layouts' aptness is mostly application domain-dependant (Heckner, Neubauer, & Wolff, 2008; Milicevic et al., 2010; Oosterman & Cockburn, 2010; Tintarev & Masthoff, 2012). List and sequential layouts are probably the most suitable for information finding, retrieval, and categorizations, but often aesthetically displeasing and with low user attraction (Carpendale et al., 2012; Halvey & Keane, 2007; Oosterman & Cockburn, 2010). Clustered layouts best perform in displaying semantical and contextual relations (Deutsch et al., 2011; Oelke & Gurevych, 2014; Weiwei Cui et al., 2010), but their quality grows with spatial consumption, which yields low user acceptance by users (Chen et al., 2009; Steffen Lohmann et al., 2009; Oelke & Gurevych, 2014; Weiwei Cui et al., 2010). Cloud layouts have a high visual appeal and user acceptance, but are unsuitable for serious navigation other than serendipitous exploration (Lohmann et al., 2009; Seifert et al., 2008; Waldner et al., 2013). Probably the most interesting and surprising layout-related finding is that users appear not to favour towards 3D tag cloud interfaces, further supporting simplicity as one of its main trademarks (Diaz et al., 2009). This is promising for further research, since there is only one study supporting this hypothesis directly, but with solid empirical evidence, pointing to user unwillingness to conform to complexity during navigation. Interface functionality is often measured by users' ability to complete pre-set navigational tasks efficiently. This also relies on recommending accuracy and interface should entitle additional qualities, for example, user choices, simplicity of use, transparency, navigational delay reduction, etc. (Carpendale et al., 2012; Lee et al., 2010; Lohmann et al., 2009). One more important aspect to consider is the number of tags that occupy the allocated space for a tag cloud, since with increase in tag quantity the cognitive effort increases regardless of their relatedness (Schrammel, Deutsch, et al., 2009). What one user considers a meaningful relation is not applicable to another, potentially resulting in selecting any random tag to interrupt the navigation stall (Bateman, Muller, & Freyne, 2009; Cress et al., 2013; Held et al., 2012). This increases the navigational delay and the potential erroneous navigational paths. Users approach tag cloud to perform serendipitous exploration, and any extensive scanning can be detrimental to user experience. This is the result of users matching the external associations with internal ones, which innately causes a cognitive load (Fu, Kannampallil, & Kang, 2010). Some studies even propose about 30 tags on screen as the best balanced information delivery with quantity, since neither sparse nor to dense visualizations yielded significant success (Lohmann et al., 2009; Seifert et al., 2008). Paradoxically, the more overlaying functionalities a tag cloud interface offers (timestamps, ratings, etc.), the less efficient it becomes in navigational sense, probably because of the low cognitive value of words compared with richer navigational cues (e.g. pictures), evident in some studies (De Vita et al., 2011; Diaz et al., 2009). When empowered with rich features, an increasing tendency to create a visual clutter can appear confusing to users (Lee et al., 2010). Therefore, apart from the inherent simplicity tag clouds should keep, it is necessary to devote more attention to balancing the visual factors and ensuring that any addition of functionalities is visually discreet. If increasing the navigability with richer visualisations, then reconsidering tag cloud deployment to other interface alternatives probably is an appropriate course of action.

Social insight, with moderate results for both IS (6.11) and IF (0.4), is the effect of limited amount of studies addressing its direct effect on user acceptance. There is a significant difference of opinions whether employing social insight can increase user acceptance, however results showed no ill effects (Christie et al., 2011; Dörk et al., 2013; Klaisubun, Kajondecha, & Ishikawa, 2007). Most users prefer to see popular resources tagged by others, and questioning the appropriateness of this visualization is justifiable only if setting the application domain towards tag clouds used for personal information organization, such as banking, document organization, and information retrieval (Ravendran et al., 2012). Social insight is the foundation of tag clouds because it conveys a summarized information on online domain’s intended purpose and recent activities (Christie, Lueg, & Baghaei, 2010; Christie et al., 2011; Herlocker, Konstan, & Riedl, 2000; Klaisubun et al., 2007). Since popularity in not uniformly applicable across different user interests, some recommender algorithms introduce tags from user’s collection to provide tag sets that strive to balance it with popularity, attempting to better tailor to the individual users. Although this approach may seem fitting, there are cases when users simply want to have an insight into community’s shared resources or even follow other influential taggers of interest to them (Christie et al., 2011; Dörk et al., 2013). No matter how good machine-generated recommendations are, this social aspect must not be neglected, especially because providing social insight for users carries the lowest implementation cost, and no doubt should be an integral part of any social tag cloud, even as an choice. Since there is no consensus on this topic, the design perhaps should focus on providing the mechanisms for user control over this process, whether through multifaceted tag cloud (Skoutas & Alrifai, 2011), or user-controlled recommendation adjustment (Emerson et al., 2013). The first alternative has a potential of introducing visual clutter and confusion, especially if not made transparent, and the latter can present an obstacle for novice users by introducing a certain degree of complexity, thus taking away natively simple form of a tag cloud if not implemented discreetly. Another approach is to include a simple on/off button, allowing users to decide whether they want to follow popular activities or have personalized recommendations. Regardless of approach, the analyzed literature provided enough evidence that supports user diversity, and any deterministic recommending is probably inadequate unless the application domain is narrow, for example, shopping sites or blogs.

Personalization, with IS (4.9) and IF (0.5) extends on the topic of social insight, interpretable in two ways: the users' ability to browse own tags and resources for e.g. self-organization, or the ability of recommender algorithm to propose e.g. popular tags based on the user-preferred resources. The medium values have resulted from the lack of sufficient empirical evidence that would support high user acceptance of this feature, and similar to social insight, the findings in reviewed studies are contradicting. User preferences can be extrapolated from viewing a certain resource or by the tags from a personal profile user assigned to a resource. Although some studies found personalization to have medium to high user acceptance (Bateman et al., 2009; Christie et al., 2011; Panke & Gaiser, 2009; Sen et al., 2006; Tintarev & Masthoff, 2012), a question of highly personalized aspect of tag clouds naturally imposes (Sen et al., 2006). It can be detrimental to recommendation efficiency, especially if designed to rather aggressively narrow the results to the point of expelling any social insight (Tintarev & Masthoff, 2012). Users sometimes rely solely on public opinions in making navigational or resource choices (Held et al., 2012), finding personalized tag clouds redundant (Gedikli, Jannach, & Ge, 2014). Contrary to this finding, users whom perform a search using a single tag can be presented with higher quality resource recommendation if the algorithm considers tags from their personal profile (Bateman et al., 2009). Therefore, the efficiency of simpler tag cloud visualizations can increase with personalized results. Both personalized and non-personalized results to have their significance in navigational efficiency, depending on the task performed (Christie et al., 2011). However, some personalized implementations have positive user feedback on their efficiency (Christie et al., 2011; Vig, Sen, & Riedl, 2008). Wordle's success is a result of high personalization by describing an individual's activities (Viégas & Wattenberg, 2008), and PETAC are suitable as record keepers of personal information and organization (Christie et al., 2010). Most users have inclination towards demarking professional and private information, often using alternate online identities and information hiding (Panke & Gaiser, 2009). Therefore, separating personal tag clouds from social ones has the potential of improving the privacy aspect. Although personalized explanations are detrimental to effectiveness, but improve user satisfaction (Tintarev & Masthoff, 2012), the ideals for both effectiveness and efficiency may have been pushing tag clouds into decay over the last decade. Regardless of the continuing argument, deprived of personalization, tag clouds became a simple outline of site's recent or popular activities. This approach is of questionable value, being the same for all users. Considering the limited evidence points out the increase in user satisfaction, some degree of personalization should be included if possible, and perhaps balanced through broader user control of the results type or on-demand overviews. However, being a potential "opposing force" with social insight, the application domain analysis is necessary to set the correct early boundaries.

User control had a high item significance of (9.71) and impact force of (0.7), and it probably is the key factor to tag clouds evolution, however challenging to implement successfully, mainly because of the unavoidable increase in user interaction complexity. Since newer studies propose wider user control (Dörk et al., 2013; Emerson et al., 2013), in practice this can introduce a steeper learning curve if designed with overwhelming choices. Since tag clouds are not efficient for structured information exploration, introducing user control and added complexity can repel users who typically use tag clouds for tasks of serendipitous exploration (Hong et al. 2008). Recommender algorithms and related visualizations however have proven inefficient in catering to different interests if completely automated. Undoubtedly, many aspects of user control can potentially increase user acceptance and satisfaction (Dörk et al., 2013; Gwizdka & Cole, 2013; Oelke & Gurevych, 2014; Oosterman & Cockburn, 2010), based on several successful implementations. It is possible to mitigate some of the potential problems by introducing user controls with the highest user acceptance, hence balancing complexity with the simplicity of use. One example is Wordle (Viegas et al., 2009; Viégas & Wattenberg, 2008), perhaps the most popular tag cloud, where users can alter tag colour, size, orientation, number of tags, etc. Even though Wordle is highly attractive, and because of its rather chaotic layout reduces its effectiveness to being amusing or reflect topical impressions, still its popularity calls for higher user control, at least in certain visual aspects. Other user control types have also proven to have high user acceptance, such as sorting (Oosterman & Cockburn, 2010), filtering (Christie et al., 2010), explanations on demand (displaying tag rating or its relationships to resources or other tags), colour-coding (A. Chiarella, 2011; Dörk et al., 2013; Viegas et al., 2009), user feedback on system's efficiency (Dron, 2005), or multiple views (Carpendale et al., 2012). Some users aspire to hide a tag cloud (Lin & Chen, 2012), which can be favorable for clustered layouts in mitigating their space consumption to on-demand feature. The list of possible user controls depends on the design choices made for a tag cloud, which should stem from the application domain's purpose. For example, allowing users to disable social recommendations on a shopping site would be counterproductive, as the persuasiveness of the system would significantly decrease. Allowing a finer adjustment (without the extreme values) might however provide a better fit for user's preferences. Finally, setting up user controls is possible using discreet alternatives, such as a typical small gear icon that does not overwhelm the user but it allows for more detailed settings, or small icons that can provide different tag cloud views. That way more advanced options are available at the later stages of their familiarity with the system, opting to use them or not. It is also important to assess the amount of control assigned to users, and identify the ones that will aid in navigation and not create the unnecessary clutter, especially because simplicity is what a tag cloud must retain. If achieving a certain set of navigational goals requires too much control, this is the exact point in design flow to question the aptness of tag clouds for the task.

Transparency, with a medium IS (6.66) and IF (0.4), provided a surprising result considering the number of studies underlining its importance (Gedikli et al., 2014; Herlocker et al., 2000; Klaisubun et al., 2007; Tintarev & Masthoff, 2012). The explanation for average impact force on the model is in fewer investigations related to user acceptance, as opposed to suggesting the necessity of this feature. Most current tag clouds lack sufficient transparency, not so much because of a design flaw but due to their automated nature: providing the recommendations with no user involvement other than browsing. However, if a tag cloud aspires to offer more than automation, transparency plays a crucial role in two possible aspects. One is providing the information on mechanisms behind the navigational cues (Dron, 2005; Gedikli et al., 2014; Herlocker et al., 2000), thus flattening the learning curve through an increase in system's ease-of-use and explaining the recommendation rationale. The other implies broader user control over the system's features, where the effects of altering those features must be clear

(Dörk et al., 2013; Sen et al., 2006). Poorly designed or the absence of explanations can hinder user acceptance (Dron, Mitchell, Siviter, & Boyne, 2000; Herlocker et al., 2000; Viegas et al., 2009), and increase it even at the expense of efficiency (Gedikli et al., 2014; Herlocker et al., 2000). If the rationale behind navigational cues is not communicated clearly, then the system leans towards deterministic one, trying to control as opposed to providing control (Dron, 2005). Users prefer well-explained interface, especially domain novices, if those explanations do not add to visual clutter permanently, such as tag ratings or tagged tags (Sen et al., 2006). The transparency effect on user satisfaction appears high due to easier interpretation of presented results of recommendation (Gedikli, Ge, & Jannach, 2011) and increased trust in the system (Gedikli et al., 2014; Tintarev & Masthoff, 2012). The explanations should be presented in a form that a user is commonly familiar with (Gedikli et al., 2014), e.g. hover pop-ups, gear icon for settings (if any), help icon, brief training wizards, etc. To mitigate the problems associated with visual clutter, the design can entitle either discreet explanations, or the ones that appear on-demand, such as mouse click-and-hold. Another approach is to create a short video explaining the system's operation or principles, similar to CoREAD (A. Chiarella, 2011), cancelling the need for more space-demanding explanations. Finally, transparency should be simple since higher levels of explanations complexity can deter less experienced users, even if highly efficient (Herlocker et al., 2000).

Tag reuse with IS of (8.75) and IF of (0.48) is a subject to further research since there is a significant results discrepancy across studies, while giving little attention to its relation to user acceptance as opposed to its efficiency. The design can assume several forms, from autocomplete, drag and drop, to displaying tags already assigned to the resource or the ability to delete a tag (Ames & Naaman, 2007; Hong et al., 2008; Sen et al., 2006). It can source from either popular, personal or similar resource tag sets (Farooq et al., 2007; Gupta et al., 2010; Panke & Gaiser, 2009), based on the application domain. All these features have the potential for increasing user acceptance by reducing the cognitive load that increases workflow fluency, but can create problems for the design, commonly manifested through overtagging and visual clutter. When presented with a choice, most users will assign a certain portion of suggested tags, but not without changes (Panke & Gaiser, 2009). This can be useful to increase involvement and satisfaction with tagging experience, since users adapt the external associations to their own (Held et al., 2012). Providing tag suggestions not only improves motivation to use tagging systems, but also accelerate vocabulary convergence (Sen et al., 2006; Vig et al., 2010). Since tag suggestions can have significant benefits, one possibility would be to favour personal tags previously used, similar to CiteULike (Farooq et al., 2007; Mezghani et al., 2012). Blending them however with a limited quantity of popular ones, and potentially balancing motivation, tagging cold start and vocabulary acceptance can further improve this feature.

In relationship to recommender algorithms, tag reuse obtained a low IF of (0.1), with fewer studies addressing the existing problems. Although tag reuse can be useful, it is likely to introduce problems for recommender algorithms in a form of overtagging and linguistic differences. Overtagging is a popularity-based tagging where users take advantage of suggested tags to deliberately increase the popularity of the described resource, resulting in introduction of spam into the system, since those tags do not relate to the resource (Heckner et al., 2008). However, in all tagging systems there is no real mechanism to prevent users from *inputting* spam, which is also achievable by pasting the needed set of tags, an even easier approach than to manually select the suggested ones. Therefore directing the efforts in reducing spam should not aim towards restricting user actions, but deployed as a recommender algorithm filtering utility. Linguistic differences can lead to misinterpretation of legitimate tag (e.g. Vienna versus Wien) as spam by the recommender algorithm (Milicevic et al., 2010). Therefore, suggesting a tag format recognizable by the system has the potential of lowering this barrier. Finally, most suggestion mechanisms have to occupy an additional space, which can be detrimental to user acceptance, therefore requiring a delicate balance when fulfilling this feature. Limiting the quantity of suggested tags can possibly mitigate the problem of visual clutter.

Multifaceted browsing are changes in tag cloud's visualization and functionalities to adapt to a wider range of navigational tasks. Its impact values for user acceptance are high (0.73), with item significance of (9.00), however with fewer implemented systems. Even when employing a simple bi-faceted visualization, such as personalized and public views, it shows the tendency to cater to a wider range user types, and increase the motivation in using a tag cloud (Christie et al., 2011). Tag cloud layouts do not exhibit the same performance across various tasks, and users seem inclined to have a visualization which can be temporarily changed (Carpendale et al., 2012). This dynamic layout visualization can be transparent to users, whom can change it without knowing the underlying algorithm, by selecting the intended task, e.g. "popular resource" versus "exploring the relationships" views (Emerson et al., 2013). Multifaceted tag cloud can also strengthen the learnability (Carpendale et al., 2012), especially because not all users perceive efficiency in the same manner. An individual user may view a less efficient layout better performing for the task, especially if motivated by habitual use from another social website. It is possible to design a dynamic visualization of a tag cloud that can adapt its layout depending on the recommender algorithm used, such as similarity, co-occurrence, importance, semantical, etc. (Weiwei Cui et al., 2010). However, multifaceted navigation is difficult to set up within a limited interface space, especially if the results or result pairs are large (Skoutas & Alrifai, 2011). One approach would be to design a split-style interface, and allowing users to enter or exit the dual-view on-demand, however affecting rather limited tag cloud size. Another alternative, commonly present in desktop applications, is to employ tabular views. It would result in a more discreet interface with a flatter learning curve, but at the expense of not visualizing any comparative views if necessary.

From the recommender algorithm perspective, correlating multiple facets can produce more accurate results, such as high popularity with high relevance (Venetis et al., 2011), but also a topic or resource with timestamp or resource frequency with ratings. Users may also want to browse personal libraries, see what other users are doing (Christie et al., 2011; Schoefegger & Granitzer, 2012), or even compare both aspects. Christie et al. propose tri-faceted functionality of a tag cloud: the display of personally generated information, data type customization, and visualizing not only what is being posted but also what is being read (Christie et al., 2010). Single-faceted tag cloud will most likely suffer from a cold start (Carpendale et al., 2012), which is avoidable by introducing bi-faceted interface with popular and personalized tags. Because of all the proposed added features, perhaps the biggest finding of this thesis relates exactly to multifaceted browsing. Tag clouds can take two paths, one by visualizing simplistic datasets, such as ranked tags based on popularity or co-occurrence, or assume a more complex form carried out through multiple facets.

Layout-independent Analysis of Tag Visualizations

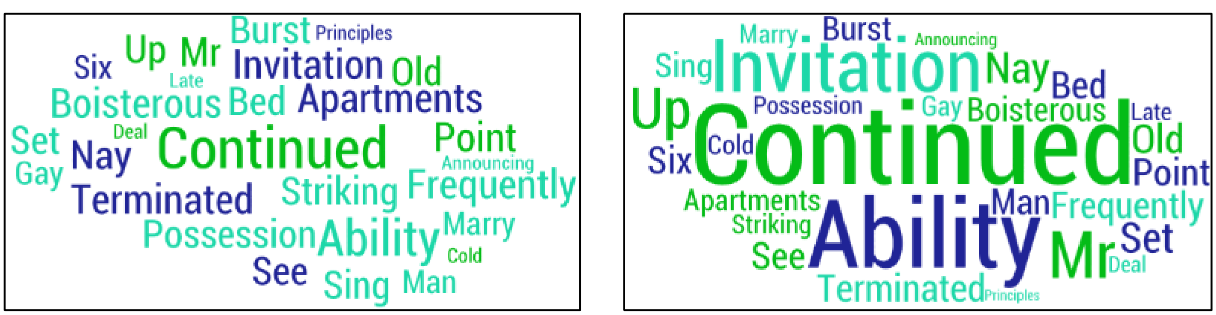
Tag size of a tag was the most visited topic within the context of graphical interface, with a consensus on its dominance in attracting user attention when scanning a tag cloud. However, opinions vary on its use, whether stimulating it to attract user attention to the main results produced by the recommender algorithm and it that way guide the user (Bateman et al., 2008; Chen et al., 2009), or reduced (or avoided) as detrimental by distracting from other tags, making them redundant (Candan et al., 2008; Deutsch et al., 2011). The navigational cueing conveyed using tag size, may not be precise as plain lists when looking for a specific information (Derntl et al., 2011; Oosterman & Cockburn, 2010). Waldner et al. (2013), find tag size more precise, however slower because comparison time with other tags is longer. The interaction speed significantly grows with increase in tag size (Halvey & Keane, 2007; Lohmann et al., 2009). There are two reasons for these contradicting results. Firstly, navigation cues have an guiding role by design, and some are stronger than the others. If presenting users with a task of finding a specific term, the internal associations have to assess that guidance against all the information in tag cloud. This requires more cognitive effort and decreases the navigational speed in comparison to serendipitous browsing. For example, alphabetical ordering performs much better for this task (Derntl et al., 2011). Secondly, the tag size does not perform the same in all layouts, as presented later in this chapter.

Both *tag size and tag length* showed no notable influence on aesthetics and user acceptance, except in list layouts. While it was impossible to extrapolate any connection between tag size and aesthetics, the cognitive effect remains unquestioned. Its role as a social signaller is two and a half to three times higher than other relevant visual properties (Dron, 2005; Schrammel, Deutsch, et al., 2009). It even dominates users’ internal knowledge during navigation, as opposed to smaller tags (Held et al., 2012). Users’ recall ability is also higher for larger tags (Rivadeneira et al., 2007; Schrammel, Leitner, et al., 2009), and overall a primary choice for selection (Bateman et al., 2008; Chen et al., 2009; Lee et al., 2010).

Tag size is favorable to tag cloud attractiveness by reducing the monotony of otherwise plain block of text (Viegas et al., 2009). Therefore, by increasing the size of tags on important topics novice users can interact with the domain content more easily, while subtly alerting experienced users can of recent changes. Being the most powerful visual navigational cue, using excessively large sizes has proven to be detrimental to user experience and implies caution in design. In its extremes, it can affect other visual properties in a negative way (Emerson et al., 2013; Viegas et al., 2009), especially layout size and inter-tag spaces. (Figure V-8). The problem is more prominent on mobile devices, where already constricting space can produce significant visual clutter (Kaser & Lemire, 2007; Lohmann et al., 2009). Its exact cueing value, compared to combination of other visual factors (e.g. colour and weight) is unknown, but those can affect its significance as a primary navigation cue (Bateman et al., 2008; Dron, 2005).

Figure V-8.

Cloud layout variant: white lines (left), tag length and readability (right) affected by tag size (tagul.com).



Probably the biggest challenge in signalling using tag size appears with longer words that inherently attract attention, where expectedly weaker cues can become dominant (Emerson et al., 2013; Lee et al., 2010; Viegas et al., 2009). One solution is to truncate longer or multiword tags, however with the potential of creating confusion with users as it affects the cognitive context (Chen et al., 2009). While it is possible to impose both upper and lower limits to tag length, it would cause a selection bias for the recommender algorithm, excluding important words because of their excessive or short length. Alternatively, moderate tag size reduction of longer words in certain layouts might mitigate most of the unwanted or accidental cueing effect (Figure V-8).

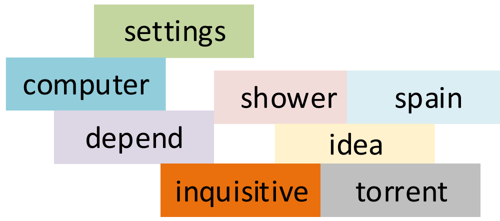
Based on the preliminary study, the increase in *layout size* reflects negatively on user acceptance, mostly prominent in clustered (Chen et al., 2009; Danilovic, 2013; Lohmann et al., 2009). Surprisingly, the result revealed that only a few studies directly examined this topic considering user preferences, overall resulting in a small negative impact (IS 8.8 and IF (-0.17)). However, a circumstantial evidence suggests this impact is greater, mainly from implementations that tried to reduce the layout size by accepting the results of previous studies (Chen et al., 2009; Diaz et al., 2009; Emerson et al., 2013; Weiwei Cui et al., 2010). When analyzed from a technical perspective, increase in tag cloud size decreases the navigational cost (Lohmann et al., 2009; Skoutas & Alrifai, 2011). This effect is more prominent in interfaces that contain rich content, such as media items, folksonomies, or even semantical structures. Conversely, there is no decrease in navigability if reducing the size of a sequential layout (Helic et al., 2010b), which points to better fitness of tag clouds for simpler visualizations and tasks. In mitigating problems with space-demanding layouts, some studies propose user control over interface size upon visualizing the initial set of results within pre-set minimum space (Carpendale et al., 2012), or based on the predetermined number of tags (Helic et al., 2010b). This may be inappropriate for clustered or list layouts because imposing the limit on the number of relationship nodes can reduce comprehension. Probably the best universal solution is to design the interface boundaries before recommender algorithm and populate the available space with the maximum number of tags (Helic & Strohmaier, 2011; Kaser & Lemire, 2007), and concentrating efforts on increasing the effectiveness of the navigational cues. Considering there is no definition (or even a hint) what constitutes an average satisfactory layout size regarding user acceptance, the application domain's purpose should guide any choices. Larger size can be justified in domains that primarily rely on tag clouds for search and exploration.

Inter-tag spaces in tag clouds negatively affect aesthetics and user acceptance, unless used to intentionally separate semantical or folksonomic groupings (Chen et al., 2009; Deutsch et al., 2011). A high item significance (7.06) and low negative impact force (-0.15), results from smaller number of studies focusing on relevant visual appeal and relation to user acceptance. However, those few that did, stressed out the importance of this factor when designing a tag cloud (Deutsch et al., 2011; Emerson et al., 2013; Kaser & Lemire, 2007). There is a strong tendency to consider inter-tag spaces as a collateral resulting from increasing the tag size of certain tags (being a primary navigational cue). Many algorithms are commonly used to dynamically gather tags as close as possible, without considering layout-dependent alternative approaches. For example, clustered layouts use inter-tag spaces to their advantage (Figure V-11), while they are detrimental to sequential layouts' visual appeal (Figure V-8, right). In the latter case, considering a combination of cues other than the tag size is rare, regardless of their potential appropriateness. A pre-set equal line spacing (Kaser & Lemire, 2007), may be applicable to sequential and list layouts, but not to other. In most layouts a large tag size increases inter-tag spaces and reduces the associative properties of tag groups (Chen et al., 2009; Weiwei Cui et al., 2010). Alternatively, smaller tag sizes increase the cognitive effort and reduce readability, mostly caused by overlapping tags (Chen et al., 2009; Lohmann et al., 2009). There is little evidence that inter-tag spaces directly affect aesthetics and user acceptance, and it can be assumed that only the extreme values can be potentially detrimental. Since it is impossible to generalize this visual effect, its role is discussed within layouts context later in this chapter.

Tag colour had one of the highest significance values within the findings model, with overall high impact on the aesthetics (IS of 9.19 and IF of 0.75). Even when using extreme values, colour does not to have any negative effect on navigability. Users rarely describe interface as “too colourful” or having “too little colour” (Viegas et al., 2009). Liberally manipulating this variable can therefore to achieve the desired visualization goal, for example, topical segregation. As a navigation cue, tag colours have variable value, mainly dominated by tag size, and a potentially leading role in layouts that suffer from increased tag size as a main signaller, e.g. sequential. If combined with the increased font weight, colour can be even more powerful navigation cue than the large tag size alone (Bateman et al., 2008). Since tag colour has a high impact on aesthetics and therefore on user acceptance, this visual property can increase interface attractiveness (Viégas & Wattenberg, 2008). One practice in this category that should be avoided is applying colours to tag background (Figure V-9), since users are critical of it (Lee et al., 2010; Waldner et al., 2013). One successful implementation of using color as a main signaller is evident in CoREAD, a social annotation software, presenting users with colour-coded terms (A. Chiarella, 2011). Since all the reviewed implementations applied background colouring/highlighting across all tags, it would be interesting to study the effects of selective colour-coding. For example, mouse-hover visualizing the association to other tags through background highlighting, with a lower likelihood of being visually intrusive.

Figure V-9.

Background colouring.



Tag contrast (intensity, saturation) had a low significance (5.63), and a notable negative impact (-0.6), because of determinate research conclusions this visual property has a negative impact on aesthetics and user acceptance (Bateman et al., 2008; Carpendale et al., 2012; Waldner et al., 2013). It is a design choice with little room for design compromise, resulting from poor performance as a primary or secondary navigational cue. Users perceive faded tags as illegible, for example, where they intentionally aspire to represent less important tags or their frequency (Lee et al., 2010; Waldner et al., 2013). Furthermore, Bateman et al. (2008) found no relation between variable tag intensity and its individual success as navigational cue, advising no more than 10% of variation in levels as it can affect readability. Since such subtle differences are not visible enough for cueing, extending this range results in poor readability (Waldner et al., 2013). Therefore, when designing for a tag cloud, varying tag contrast should be avoided because of the high risk to aesthetics and low benefits as a navigational cue. Possibly the only exception is to purposely fade a large pool of tags that not related to a mouse-hovered one, to avert users from selecting those tags. Additionally, tag intensity could suggest the level of someone’s reputation within a social system, e.g. teachers versus students (Dron, 2014), or acting as a tertiary navigational cue, such as timestamps (Gambette & Véronis, 2010a).

Font weight gained low significance (1.0) and moderate impact on the visual interface (0.66) since only three studies addressed it, but with strong supporting evidence. While users rarely distinguish it from a regular font in an aesthetic sense, there are indications this is a powerful navigation cue (Bateman et al., 2008; Dron, 2005; Rivadeneira et al., 2007). This especially applies to layouts that cannot vary tag size significantly, and in possible combination with colour. However, further empirical research must support its relative navigational value.

Readability (IS 8.47, IF 0.57), has the highest impact on aesthetics, since reduced readability is the most directly criticized effect by users. This factor should act as the main balancer of other visual properties, since any visual compromises or chosen set of navigational cues cannot compensate for the lack of readability. For example, font shape does not impact aesthetics significantly if using the chosen font uniformly (Viegas et al., 2009), but users find font shape variations affect readability (Waldner et al., 2013). The major reasons for its detriment are known from the studies that examined the relation of readability to user acceptance. For example, tight groupings of text (Lee et al., 2010), high number of visualized tags within an assigned space (Emerson et al., 2013), small, faded or inappropriate font (Waldner et al., 2013), tag clutter (Chen et al., 2009), and tag overlapping (Weiwei Cui et al., 2010). However little research has focused on tag orientation. This is a surprising finding, since the preliminary study suggested vertical tag orientation yields low user acceptance (Danilovic, 2013; Emerson et al., 2013; Waldner et al., 2013). However, in those few studies employing user surveys, tags set at an angle carried lower readability. One explanation is in most widely used cloud layouts, which often employ tags at different angles to increase the attractiveness, and avert focus from that specific aspect when surveying users, especially since conveying organized information is often not the primary goal. Even though there is little empiric evidence to support it and implies subsequent research, the preliminary conclusion is to avoid using angles in tag placement as a precaution. This is emphasized in domains where semantical and associative relations are important, as it can increase users' cognitive effort. However, aesthetics has a strong influence on user acceptance, and some users may find vertical tags pleasing, therefore this relation needs further investigation to draw solid conclusions.

Clustered Layouts

Clustered layouts was well saturated topic, with greatest usability effect in the application domains where conveying semantical or any associative information is the objective, such as folksonomic hierarchies (Bateman et al., 2008; Skoutas & Alrifai, 2011). These layouts also have medium to high user acceptance, mostly hindered by the size requirement that is necessary to display the relations between tags. Clustering excels at visualizing multiple topics simultaneously, and demands lower cognitive effort for semantical relationships understanding (Weiwei Cui et al., 2010). However, users often find this type of navigation difficult for identifying a specific tag or popular topic (Chen et al., 2009; Lohmann et al., 2009). Gupta et al. (2010) claim that clustered layouts outperform alphabetical lists in a semantical sense, contrary to the findings of two experimental studies (Chen et al., 2009; Schrammel et al., 2009).

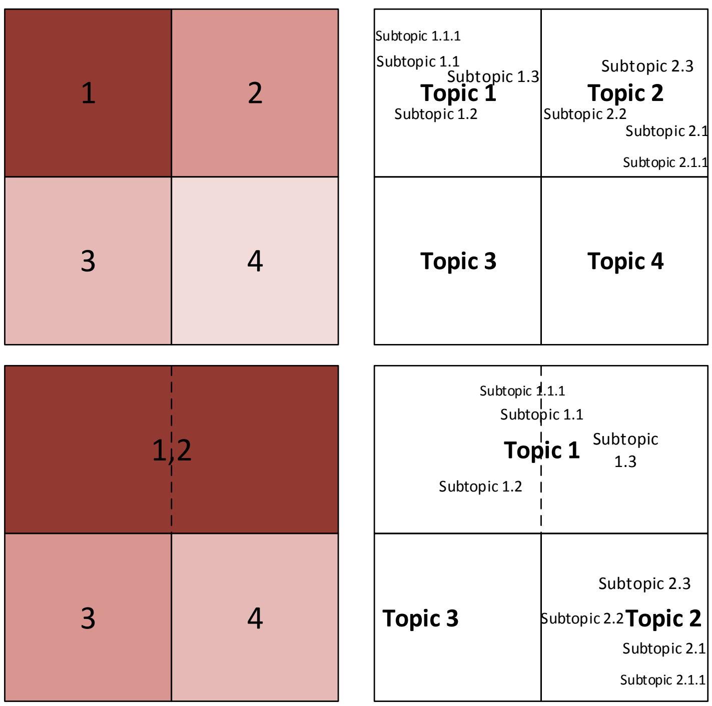
Because of the high spatial distribution, these types of layout are the least affected by the ill effects of certain graphical elements variations, such as large tag size, inter-tag spaces or background colouration. However, constriction to smaller physical boundaries reduces their efficiency, that is, denser coupled tags will reduce the clarity of relationships, while overlapping tags that can disturb the semantical groups' navigational cueing accuracy (Lohmann et al., 2009).

Clustered layouts have the highest tendency to achieve balanced distribution across visual quadrants (Lohmann et al., 2009). They can incorporate lower number of tags compared with other layouts in delivering similar information value, mostly because of transparent relationships. With variations in design, spatial distribution efficiency varies accordingly, mostly influenced by high tag entropy within a quadrant (Chen et al., 2009), especially with large clusters (Oelke & Gurevych, 2014). Although high tag entropy can appear attractive to users, the unavoidable visual clutter and disorderly user interface increase the cognitive effort in discovering their exact relationships (Chen et al., 2009; Forlines & Balakrishnan, 2009; Lohmann et al., 2009; Oelke & Gurevych, 2014). There are some suggestions that introducing an orderly tag grid to correct the problems with tag entropy can further increase layout size because of the wasted inter-tag space (Schrammel, Leitner, et al., 2009; Weiwei Cui et al., 2010).

Contrary to other layouts, inter-cluster spaces make a meaningful navigational cue, by segregating main topical clusters and reflecting users' organizational patterns (Oelke & Gurevych, 2014). However, cluster colouration can also perform a similar role (Deutsch et al., 2011; Lohmann et al., 2009), and it is an approach with high user acceptance (Chen et al., 2009). In that light, one possible way of reducing the cognitive effort during scanning is to align the main topics' clusters equally spread across the four quadrants, and predetermining the number of topics. It is possible to quantize these topics within interface allowance, limiting the number of second-level tags, while retaining equal inter-cluster spacing to emphasize topical segregation (Figure V-10). Such visual arrangement would allow users to scan the tag cloud for topic offered in a typical top-left to bottom-right or zig-zag pattern, without the distractive influence of centrally positioned tag. Introducing orderly grid can equalize the importance of main topics and needs lower cognitive effort during scanning (Sanchez-Zamora & Llamas-Nistal, 2009), however at the cost of topical reduction. Although this could hinder the informational value, many studies show that users are likely to interact with appealing over efficient tag cloud (Chen et al., 2009; Kuo et al., 2007; C. Trattner, 2011). Tag size has the highest positive influence exactly in clustered layouts, as larger tags provide navigational cueing for main topics (Figure V-10), and mostly rely on grouping co-occurring or semantically similar terms around the central tag (Cui et al., 2010).

Figure V-10.

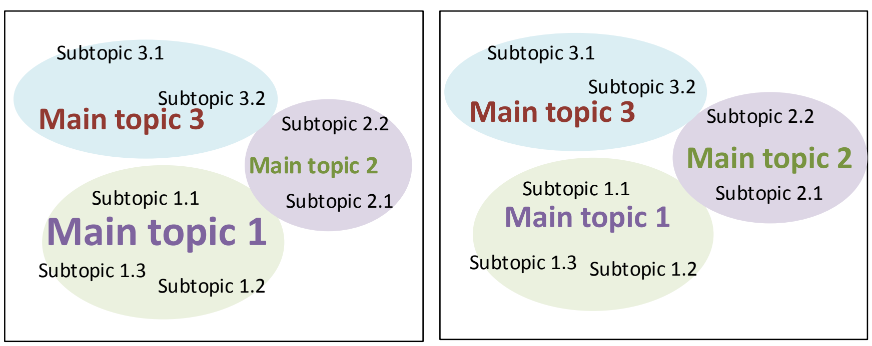
Most frequent scanning pattern in a tag cloud 1-4 (left); the proposed quadrant distribution (right).



Stimulating tag size is favorable since clustered layouts can more clearly distinguish multiple topics, signalled by the largest tags, while smaller tags assume a descriptive role (Schrammel et al., 2009). This layout however can suffer if representing main topics through varying tag sizes, resulting in interface that has the potential to confuse the users (Chen et al., 2009). The effect is present when signalling tag frequency with topical relatedness. A potential approach is predefining tag size based on topical level (e.g. a category), and avoiding varying tag sizes based on weights or frequencies. Using tag sizes between the largest and the smallest to suggest tag frequency or co-occurrence can hinder semantical or ontological groupings, and in those cases reconsidering clustering to other layouts may be fitting. A proposal for resolving tag sizes is presented below (Figure V-10). While setting a background colour is not a good practice, it can segregate topical or hierarchical clusters or point out overlapping subtopics in clustered layouts (Figure V-11, and Figure V-12). This practice can also increase the visual appeal (Chen et al., 2009; Di Caro et al., 2011).

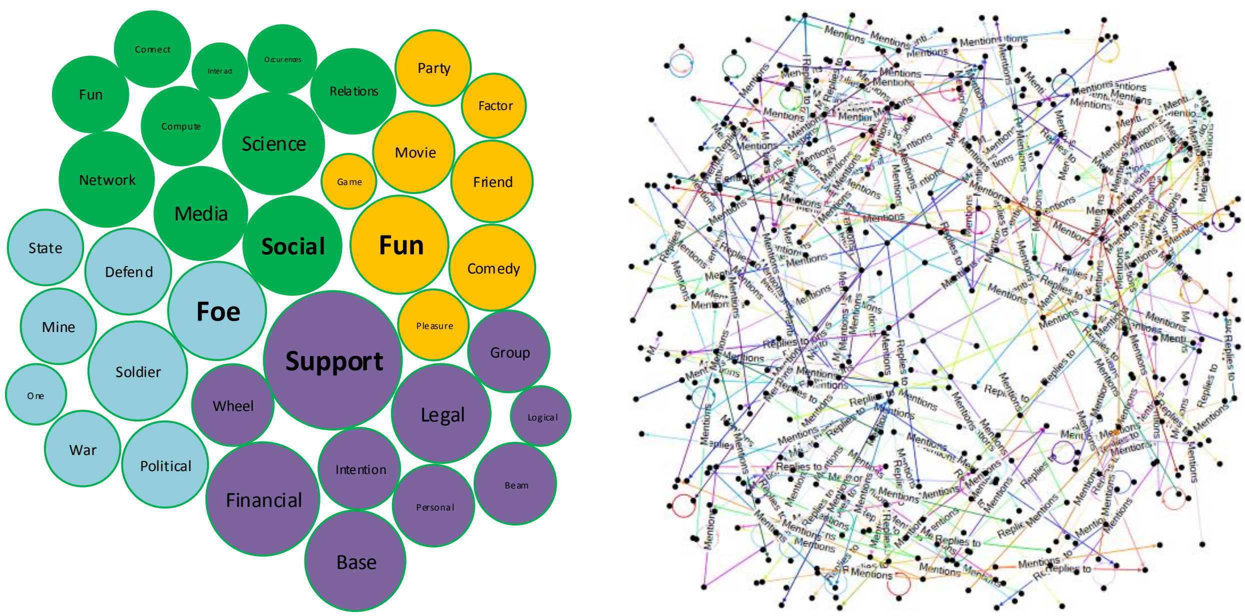
Figure V-11.

Topical segregation by colour, varying tag size (left), same tag size (right).



Proper readability of clustered tag clouds mainly depends on spatial consumption, the number of tags displayed and potentially overlapping topics (Lohmann et al., 2009). High spatial consumption introduces lager area for scanning, and a higher number of tags will produce more items to be scanned, both of which add to cognitive effort and reduced readability (Gou et al., 2010). It is possible to minimize the number of tags to the quantity required for navigation (Oelke & Gurevych, 2014), as opposed to some current practices of displaying as many tags as possible. While this will reduce interface size, larger inter-tag spaces should only separate main topical clusters, and grouping within-cluster tags should at significantly smaller distance (Gambette & Véronis, 2010b).

Figure V-12.
Separating topical clusters by colours (left), displaying all relationships (right).



Cloud layouts

Cloud layouts are the most commonly used layouts, with high effectiveness in application domains focusing on conveying popular topics or an increase in website attractiveness. They have little to no semantical or organizational power, depending on the exact layout variant; however yield high user acceptance in applications domains where popular trends are the nurtured (Viegas et al., 2009). These layouts are easily scalable in size, and have the potential of satisfying the user need for resulting in a discreet user interface. It is also the least suitable choice for information discovery, other than for simple ranked display of web pages that contain the user-selected tag. When compared to tables, cloud layouts will underperform in selection precision, except for finding a minima and maxima in a dataset (Oosterman & Cockburn, 2010), which is a result of signalling those extreme values through tag size. They are also one of the most popular choices for tag cloud representation, mainly because of their ability to easily represent the domain purpose (Schrammel et al., 2009), and impress users visually (Lohmann et al., 2009; Viegas et al., 2009). The navigational cues (tags) of highest significance are often placed in the center, and with largest tag size, while dynamically scattering the less important ones around them. Tag importance is signalled by varying tag sizes or colours (Chuang, Manning, & Heer, 2012; Viegas et al., 2009; Waldner et al., 2013). Using these cues simultaneously without an obvious paradigm, can result in high cognitive stall in any navigation other than serendipitous exploration of a domain. Even when users start scanning in a chaotically, if no desired term is found they switch to more orderly zig-zag pattern (Deutsch et al., 2011), which in this case is not obvious. Since tag displacement is often random, the information discovery is difficult to learn and logically reapply, with a higher potential for longer "*pre-activation paralysis analysis*". By exposing users to a diverse set of navigational cues in a seemingly disorderly tag allocation, a specific cue dominance decreases, resulting in system's reduced ability to guide (Oosterman & Cockburn, 2010). This causes recommendations to lose their significance since some users may select coloured while others larger tags, and often dictating tag placement by space conservation and filling of inter-tag spaces (Figure V-8). Although randomness is detrimental to user acceptance for information findability (Halvey & Keane, 2007; Schrammel et al., 2009), cloud layouts are oppose this trend through attractiveness, which is an important finding. In comparison to other layouts, the cloud ones are however in disadvantage by having little maneuvering space to setting up the visual order.

Figure V-13.

Wordle visual variants (www.wordle.net).

Tag size can be detrimental to the effectiveness of serendipitous navigation since users mostly focus on larger tags; especially because intense colour variations prevent an effective visual cueing and instead focus on increasing the visual appeal (Viegas et al., 2009). However, if the primary goal is to guide users, then presenting the important cues through larger tag size is suitable. Since cloud layouts probably have the smallest navigational power, tag size use can be liberal and serve as a leverage to increasing the attractiveness of the user interface, where larger tags will still have the power to attract user attention to key cues.

Inter-tag spaces are generally well resolved and any extreme values can only affect the aesthetics, with little impact on the quality of information. However, those spaces sometimes can directly affect tag size. Tags' frequent reduction in size not based on their lower significance, but the ability to fill the empty inter-tag area created by the non-overlapping borders of larger (higher significance) tags (Figure V-11). Therefore, sometimes the information importance is not certain if based solely on tag size. Even though it is possible to place a significant amount of tags within a relatively small area, these effects will reduce any navigational cueing to a few selectable in practice (Figure V-8.) Cloud layouts also potentially yield the lowest readability of all layout types, stemming from overlapping (Weiwei Cui et al., 2010), and vertical or angled tags (Waldner et al., 2013). In their eye-tracking study, Lohmann et al. (2009) found cloud layouts to have the highest centrality power, where users give less attention to peripherally placed tags. Chuang et al. (2012) tried to disrupt this behaviour and increase the visibility by centrally placing smaller tags. The attempt was to equalize centrality power with assigning the largest tags across different quadrants, however it resulted in low user acceptance. If cloud layouts retain a rather disorganized form, then the need for smaller tags becomes questionable leading to very sparse tag clouds (Deutsch et al., 2011). One starting solution to this problem would be to place tags with higher e.g. co-occurring frequency closer to the centre, emphasize popular tags with discrete tag size or weights (Sanchez-Zamora & Llamas-Nistal, 2009; Seifert et al., 2008), and use data grids to visualize tag patterns. This would have the potential to divert user attention to other quadrants, especially if limiting the maximum tag size to avoid any extreme visual asymmetries.

Sequential and List Layouts

Sequential layouts perform the best for organizing information, governed by the traditional left-to-right reading pattern (Kaser & Lemire, 2007). It carries the least cognitive effort for users to identify the intended pattern, either frequency-based or alphabetical (Lohmann et al., 2009; Oelke & Gurevych, 2014).

Figure V-14.

Alphabetical list layout (left), frequency-based signalling (right).

Figure V-15.

An example of suggesting a tag using colour.

Figure V-16.

Use of extreme colours (left and right), framed tags (bottom).

The evidence clearly revealed that sequential layouts have the least tolerance for variations in tag size (Figure V-15, top-left), due to increase in inter-tag spaces affecting their aesthetics (Deutsch et al., 2011; Emerson et al., 2013; Gupta et al., 2010; Weiwei Cui et al., 2010). Kaser and Lemire (2007) suggest designing the default line spacing in advance by predicting for the largest tag, thus avoiding variances. While the alphabetical arrangement needs subtle or no emphasis of the starting letters, any other organization of cues undoubtedly does, especially if sorting tags by popularity or frequency (Figure V-14, right). Considering the latter intent, varying tag size in continuous text is a questionable design choice as opposed to other available cues. For example, tag colour and font weights could potentially be as equally powerful if combined (Figure V-15). Unlike cloud layouts, background colourations (Figure V-16, top-left and bottom), and erratic colour variations (Figure V-16, top right and top left) create a significant visual clutter. This is not a good practice for information discovery (Bateman et al., 2008), and affects readability (Carpendale et al., 2012; Waldner et al., 2013). For these reasons, balancing visual properties to gain both high attractiveness and quality information representation can be challenging. Several studies confirmed that users do not read the tag clouds (or even Google results), but rather scan them (Carpendale et al., 2012; Gwizdka & Cole, 2013; Halvey & Keane, 2007).

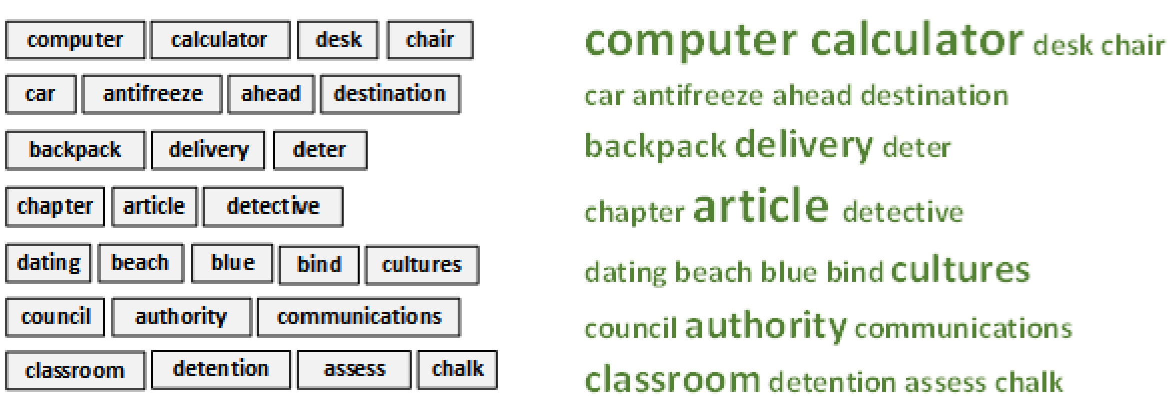
List and sequential layouts carry low cognitive effort when scanned, both in comparison with the traditional interfaces (e.g. tables) and most other layout variations (Carpendale et al., 2012). If sorted alphabetically, the scanning patterns follow top-left to bottom-right trend. However, this is not explicit, because of the user’s understanding of the underlying pattern (Deutsch et al., 2011). The scanning focuses to first letters of a targeted word: as soon as the user locates the group beginning with a targeted letter, the scanning assumes a zig-zag pattern (Figure V-17).

Figure V-17.

Scanning patterns example, alphabetical (left), random (right).

Figure V-18.

Vertically aligned list.



List layouts can be considered a subtype of sequential layouts with vertical alignment used to separate the main topics in a single line. They share visual similarity to bulleted lists in text documents, with the exceptional ability to represent tag relationship or classification, typically in a top-down and left-to-right fashion (Figure V-18).

While tag size is the dominating factor in selecting a specific tag (Cress et al., 2013; Dron, 2008b), the position on the list also has strong influence, where top items have a higher likelihood of selection (Dron, 2008b). This layout type inherits the tag size problem from sequential ones, and extreme visualization can create large inter-tag space unappealing to users. It is therefore less suitable for visualizing frequency-based algorithms and more fitting for reflecting semantical or ontological relationships.

Recommender Algorithm

Although recommender algorithms are not within the scope of this study, a brief outline was necessary since it is impossible to address visualization in an isolation, and these subsystems are inseparable. The intent was assess the relative success of implementing a certain feature, and extrapolate the overall feasibility. Such finding were invaluable, with potential of providing a base for adaptive tag clouds, rich with features that could cohere to various application domains.

The literature selection method partially biased the analysis of *semantical efficiency*, by favouring user acceptance, tag cloud features and interfaces. It got medium coverage by studies pool, (IS 4.9, IF 0.13) on the recommender algorithm. Even if the initially analyzed publications outside the selected pool had been included, it would have only solidified the results. Semantical relations are probably the most influential navigational cue in a tag cloud (Chen et al., 2009; Dron, 2008b; Oelke & Gurevych, 2014), especially when compared to mere keywords (Fu et al., 2010). In social bookmarking systems, users' attention diverts to topics semantically represented by tags. Subsequently, those topics influence the users' information extraction, and in turn they further influence resource's tags (Fu et al., 2010). Similar principle applies to social annotation software. For example, CoREAD uses word signalling states to describe the importance of the document content, proved semantic stability appearing within a group through self-organization (A. Chiarella & Lajoie, 2010). From the recommender perspective, many problems hinder its success, primarily the power of internal associations of a user, which outweigh the influence of the external ones (Fu, Kannampallil, & Kang, 2010). This prevents the uniform application of information relatedness across a wide range of users (Milicevic et al., 2010). The tendency of online communities' vocabulary to converge through crowds' influence (tag adopting) reduces semantical extrapolation abilities. As tags decrease in uniqueness, and the resource entropy increases, they become less meaningful (Fu, Kannampallil, & Kang, 2010; Wash & Rader, 2008). Spamming is also a significant problem that affects semantical precision more than of any other algorithm, e.g. tag co-occurrences, frequency, etc. To design a semantical extraction algorithm, it is necessary to segregate user types (experts versus novices) (Fu et al., 2010; Körner et al., 2010; Strohmaier et al., 2012), which contributes to an increased design effort. Previously mentioned successful applications perform within mostly expert groups or academic environments, which are a narrow subset of a global social group, sharing some similarity. Electrical engineers and medical doctors have low academic similarity, however it significantly grows if contrasted with people from lower academic background. Therefore, stigmergic signs can have different meaning to different users (Dron, 2005). From data visualization perspective, proper display of relationships between tags or explaining the criteria used for construct demands ample amounts of space. It is therefore questionable if semantical relationships have a future in simple environments of traditional tag clouds and widely applicable types of social software. This does not mean abandoning the idea of visualizing semantical relationships, but rather limited to specific application domains that could benefit from it.

Ontologies, or socially generated categories, with item significance of (7.44), and with impact force of (0.41), was well supported topic. These results however are limited to theory, since most of the reviewed publications did not have a completed system to prove its effectiveness. Ontologies share similar problems with semantics, from necessary user segregation analysis, over data sparsity (Helic et al., 2012; Shepitsen et al., 2008), to demanding spatial distribution (Helic & Strohmaier, 2011), and data noise (Liu et al., 2010). The internal semantical association dominates co-occurring relations when users are manually categorizing tags (Oelke & Gurevych, 2014). This implies semantical construct as the base for successful ontologies (Gou et al., 2010; Shepitsen et al., 2008), which differ across users. The associated problems mainly stem from the necessary discovery the efficient mechanisms for dividing users into groups and based on their tagging activities. By assigning roles as describers or categorizers, folksonomy formation depends on tag extraction from the latter group. While some studies used tag co-occurrences for designing tag hierarchies, overall developing a successful implementation had little success, regardless of the approach. Even after identifying categories, it is difficult to determine their exact relation, e.g. is tag B a subcategory of tag A or conversely (Song, Qiu, & Farooq, 2011). Overall, justifying the usability assessment of folksonomies in tagging systems, or its benefits to users, did not receive enough attention. The efforts involved in designing such efficient systems, with other necessary recommendation features, probably outweigh the usability and purpose of tag clouds. But this is where semantical relationships and ontologies part ways: while semantical relationships can perform in smaller scale, offering sparser but still useful cueing, such limitations imposed on ontologies are probably not so productive. In other words, ontologies can successfully operate in large list or clustered layouts, however it averts from a tag cloud philosophy of space conservation. Finally, there is nothing wrong in using tags with an accommodating interface and repurposing it, similar to ChainGraph system (Lohmann et al., 2009).

Vocabulary obtained an IS of (5.78) and an IF of (0.61), pointing out moderately supported topic, with no noticeable distinction between influence on behaviour and recommender topics. Since user behaviour analysis is most often performed to increase the efficiency of recommendations, this can provide an explanation for such effect. In social tagging software, user vocabulary has the tendency of becoming less descriptive and useful. This stems from unwillingness to tag using unique keywords in continuity, especially under the influence of more popular ones (Chi & Mytkowicz, 2007). Most users also assign relatively small number of tags per resource, most commonly averaging up to seven (Chuang et al., 2012; Panke & Gaiser, 2009). As the diversity of resources increases, it results in a dropping trend of unique resource identifiers, losing their distinction when presented by tag-based results (Chi & Mytkowicz, 2007). This affects the navigability as the search results become long and diverse, challenging any recommender algorithm. On the other side of spectrum, deviating from social vocabulary means those unique tags have lower likelihood of being noticed by the recommender algorithm and presented in the results. To prevent spam, many algorithms employ a “lower threshold” technique to remove those unique tags, considered as misspelled or ambiguous (Gupta et al., 2010). The intended use of a resource also plays a major role as some users may tag with personal tags (e.g. “Owned” or “Me”), resulting in limited information value within a wider social system (Sen et al., 2006). A possible solution to these problems is to eliminate the top *n* most popular tags when extrapolating the relevant resources and compare the remaining to similar tag sets. However, this approach can be effective only in narrowing the results, and relies on the informational power of unique tags. The second direction is to increase tagging motivation and take advantage of the overlapping tags to deliver better tailored results, (Chi & Mytkowicz, 2007). This is possible by suggesting both from popular and personal pool of tags, and it would aspire to balance the uniqueness and popularity. Providing suggestions can reduce the problem of cross-domain and multilinguistic vocabulary barriers (Panke & Gaiser, 2009), resulting in increased correct spelling and jargon adoption.

User Motivation

Tag currency, with IS (8.13, IF 0.3), and impact on recommender (0.20) was topic not often visited, even outside the analyzed literature corpora. Visualizing tag currency reduces problems of recommender algorithm discarding popular older tags to newer and conversely. This is useful for increasing the system stability since long tail of popular tags can create a stagnant content, however, limiting to recent popular topics can create a constant flux unsuitable for all application domains (Dron, 2008b). In contrast, occasionally it is necessary to promote newer opinions and subsidize ones not reconfirmed (Chiarella & Lajoie, 2010). For example, long tail is helpful in learning environments since it provides a degree of quality assurance, while on a news portal constant change is more fitting. Tag currency not only can aid in offering more informed choices, but can separate the neighbours from followers in collaborative filtering systems (Au Yeung & Iwata, 2010). This is an example of non-transparent process, which assumes that influencers have more informative tags or resources than the followers, and the user group separation can increase recommendation accuracy. Since the follower faster adopts the recommended item, this temporal pattern can distinguish groups. Another non-transparent aspect is comparing users' tags from a specific period in the past with currently browsed resources to increase the recommendation precision (Farooq et al., 2007). This is useful for connecting to any older and repetitive search, for example, for academic papers or any archived information. The evolving vocabulary does not present a problem for personal knowledge organization, since users simply reduce interaction with the older resources and do not remove or retag them (Panke & Gaiser, 2009). The non-transparent predictions however can become less reflective of the newer vocabulary if including older tags and resources from user's profile, and a transparent time-stamp in this case can aid in differentiation. Transparent time-stamping can have multiple benefits, however at the expense of potential visual clutter or space consumption if made continuously visible to users. This conclusion is drawn from similar implementations that visualized frequency or other graph types (Lee et al., 2010; Lohmann et al., 2009). Tag time-stamping should be discreet, either by employing small underscore colours, or possibly enable a button/mouse-hover for temporary tag colouration or a small chart based on their age.

Tag filtering can be set up transparently as the ability of user to manually filter-out unwanted tags, or as a non-transparent mechanism within a recommender algorithm, to remove spam, misspelling, synonyms, etc. In its relation to motivation, it obtained an IS (6.9) and an IF (0.51). To increase the motivation, the focus was on transparent filtering, which did not achieve a significant topical saturation. Szomszor et al. (2008) found that non-transparent tag filtering can improve the correlation between user profiles across different social domains, with the potential of creating richer description of user interests. They also proposed designing modular filters (spam, misspelling, etc.), in which the output of one would serve as an input for the other. It is an excellent idea from the design perspective as it would allow for better adaption and multi-perspective study. User-controlled (subjective) filtering alone can increase usefulness of a tag cloud when compared to a non-transparent (factual) filter (Vig et al., 2008). Two studies that measured user acceptance found users prefer to manipulate the results in such manner (Christie et al., 2010; Vig et al., 2008), with a clear indication these processes should be more transparent (Christie et al., 2011; Emerson et al., 2013; Oelke & Gurevych, 2014). YouTube upgraded the quality of their recommendations simply by allowing the users to remove the ones they do not like, an effective example of transparent filtering. In tag clouds, user ability to add or remove tags from the search results improves the navigational efficiency and reduces cognitive effort (Gwizdka & Cole, 2013). In addition, Emerson et al. (2013) propose an even wider tag manipulation by allowing the user to adjust the tag cloud by retaining the user-selected tags only or even form a new one. Setting up tag filtering is rather easy for most recommender algorithms therefore designing user-controlled mechanisms is justifiable if the interaction with the filter is not excessively complex. For example, drag-n-drop to a recycle bin area or control+click could make suitable design choices. However, undoubtedly some filtering should unquestionably remain non-transparent, such as spamming or synonyms filtering.

Cold start is the effect of no available tags to display for a new user of a social software (Dron, 2008b), or the delay between a new tag entering a system and time to appear in a tag cloud (Au Yeung & Iwata, 2010). The topic had an IS (6.3), however with lower impact force on the model of (0.36). This is explainable by studies recognizing the problem, but with little corrective solutions. This is a surprising finding considering its effect on user motivation is detrimental, with IF (-0.44). Empty query results will discourage the users, and that empty interfaces should never appear (Carpendale et al., 2012). The differentiation between community-based and personal profiles depends on the number of tags contained within an individual profile, requiring additional information in recommending to domain novices (Schoefegger & Granitzer, 2012). In that light, there is a potential in approach taking advantage of community (popular) tags and displaying them as a starting point for all users. Gradual tag cloud morphing as the user assigns more tags, would better reflect users' individual interests. Goularte and Manzato (2012) proposed an algorithm which would target novice users and extrapolate sets of tags based on the watched resource. Although this solution is proper for mitigating such an issue, it still introduces tag noise to active users, since the algorithm does not distinguish between the two. An often overlooked problem is the social software maturity (Helic et al., 2010b), since most studies use datasets from already mature ones (BibSonomy, del.icio.us). For example, Herlocker et al. (2000) found their recommendation explanation algorithm efficient, based on neighbour similarity. They however have not considered novice users with not enough tag data available for comparison. In other words, the richer the metadata, the more cold start problem becomes pronounced (Dron, 2008b).This imposes the need for further research into solutions for start-up or growing social software solutions in overcoming cold start issues. The more active tagging is, the faster will emerge the "wisdom of the crowds" (Dron, 2008b; Körner et al., 2010), and the motivation for active tagging is lower if the results or tag suggestions are absent. Another approach is to create bi-faceted interface, one with popular community-only tags, and the other with personalized, and potentially the popular tags would be still present as a sufficient baseline for novice users. This however, has the tendency to reduce the available space by segregating the interface, and introduces the problem of an empty space in personal tag cloud.

Tag qualities represent user ability to assign a predefined property or sets of properties to a resource, such as ratings, voting, emotions, or even other tags. This feature got medium value for both item significance (5.57) and an impact force (0.57). While tags aim to describe the resource or its contents, tag qualities can expand it by assigning opinions to those tags (Dron et al., 2000; Sen et al., 2006; Vig et al., 2010). Rating the tags can significantly increase the recommendation precision (Gedikli & Jannach, 2013), having the potential to further reduce the cognitive effort, especially alongside tag suggestions. For example, some users may prefer to rate an existing tag instead choosing from the pool of suggested ones, if allowing resource retagging (e.g. Facebook). Tag qualities can also useful in large social networks, where resources are tagged by a high number of tags. The recommender algorithm can favour ones with the highest rating as opposed to tags assigned most frequently (Heckner et al., 2008; Sen et al., 2006). In educational environments, learners must know beyond simple content description, and tag qualities can help decide how well will the resource fit their current learning needs (Dron, 2008b). One problem that arises when assigning opinions is the personal interpretation of rating (Vig et al., 2010). Users may assign a negative rating although the tag describes the resource accurately, e.g. tag “rape”, thus affecting the recommendation precision. Another problem is low use of this feature (Heckner et al., 2008), probably because the necessary user involvement outweighs the benefits (Dron, 2008b). Users prefer simpler cues as opposed to complex rationalization behind the recommended results (Gedikli et al., 2011), which can also apply to ratings. Therefore, a possible solution would be to make this process transparent when rating or voting, and non-transparent for recommending. It is more efficient to encode the ratings in navigational cues, alongside the co-occurrence, frequency, or semantical relationships, and reduce the visual clutter or confusion. If carried out in this fashion, even if users stop rating, the effect of decreased recommending accuracy would be non-transparent to users. Before any implementation, it is necessary to conduct the preliminary analysis of an application domain taking user motivation in account, since this feature can remain either underutilized and create clutter, or overused for malicious purposes. Users have high motivation to rate movie tags (Gedikli et al., 2011), however less for educational resources, probably because personal needs and perspectives change over time (Dron, 2008b; Panke & Gaiser, 2009). Some social annotation software achieved success in assigning ratings to keywords and even paragraphs (A. Chiarella, 2011). This inspires further investigation on how this motivation translates or adapts to tagging systems. Higher impact force on motivation also confirms there is a more focused research on the potential user benefits in contrast to the underlying principles of operation. Therefore, finding the motivation for users to use this feature is the key to its success.

Employing tag colours as opposed to statistical figures can stimulate emotional response (Viegas et al., 2009). If paired with ratings, colour-coding can be probably as equally powerful cue as the bar charts or numbers, and potentially motivate the users. User reputation through ratings can be helpful as a leverage to increasing user motivation, whether by rating tags or the community members. It would certainly be interesting to set up this feature using different visualization styles and perform empirical testing over time to assess its usability.

Domain models, with an IS (4.46) and an IF (0.69), was a topic with a moderate coverage in both respects. Even though specific domain models were examined often (e.g. Amazon, Del.icio.us, LastFM, etc.), there was an obvious lack of understanding across literature for various application domains. Another reason for moderate values is the dominating quantity of theoretical and survey type studies, with little practical contribution. Some domain (sub)models have been provided, however isolated and specialized around individual features, and this practice cannot give the insight into the design process itself. Some examples are: tagging resource only once or allow resource retagging; permission-based or liberal; sources of resources (users or system); resource grouping or individual; to include folksonomies or not (Ames & Naaman, 2007; Milicevic et al., 2010). While this study recognizes the efforts in conducting posterior domain analysis, the goal was to discover wider social applicability, and few studies offered this insight. In other words, a satisfactory approach for one domain might not work in another, regardless of well the individual functionalities performed.

Since the efficiency of entire system will depend on it, domain model must be a priority when choosing an acceptable search interface. If tag cloud's primary purpose is to find a specific resource, the list layout would be choice that is more suitable. If finding a specific term is to be a pivotal role, perhaps a search engine would provide a better solution (Sinclair & Cardew-Hall, 2008). Tags can describe a resource, provide a summary, express an opinion, or tag other users, which are diametrically different purposes (Heckner et al., 2008). Is semantical relatedness needed or creating ontologies, or is it necessary to be fun and appealing? The list of features is nearly inexhaustible and designers must invest significant time researching the related effects. However, with little guidance on how they interact and whether certain feature combinations can improve or deteriorate tag cloud's performance.

The literature analysis had showed the domain models as rarely noticed at design time and implementations overspecialized, which leads to knowledge parcellation and an uncertain success. It is possible to plan a plethora of features and design a system that is self-learning, gradually removing ones that are conflicting (or their effect is), or to introduce human administrative control. Another direction is to study the most similar domain models and build upon their positive and negative aspects. The latter approach creates lower computational and intellectual effort in designing, and justifying the former only if the final product is applicable across wider range of domains, e.g. a customizable commercial product. Probably the most important design implication is to make a *high abstraction* of an application domain at the beginning of a process, and use it to shape all later decisions. Oosterman and Cockburn (2010) found linear lists to outperform a tag cloud when users are searching for a known item. Therefore, if providing the solution for e.g. self-organization task by other means, it is suitable to exclude such feature from a tag cloud, and devote its resources to expanding and complementing. For example, MovieLens is a social network dedicated to recommending using informative interfaces (Herlocker et al., 2000), and an attractive tag cloud providing social insight would be more fitting than the feature-rich one, thus avoiding competing interfaces. If its intended purpose is to complement the search engine (e.g. for domain novices), then the tag cloud should aim for less visual clutter and more organized information. On more extreme end of the spectrum, some sites have high persuasiveness interfaces which are usually visually both rich and informative (Tintarev & Masthoff, 2012). Therefore it is questionable if tag clouds have any useful role in such domains; for example, shopping sites such as Ebay or Amazon. Facebook can benefit from advanced exploratory features of a tag cloud (tagged photos and posts), and for increased attractiveness of personal pages. This high abstraction decides the intended *purpose* of tag cloud.

Some of the analyzed studies showed almost no differentiation between tag clouds, tag-based recommending, and visually rich systems (e.g. tagged photographs): often approaching the analysis from a tag cloud perspective has a high potential to produce inaccurate conclusions. Relatively recently, Flickr and Delicious removed the tag clouds from their websites, probably because of low use. Perhaps one remedy is to give users broader control over some of its features and appearance (e.g. frequency or co-occurrence sorting), rather than the domain-dictated static representation of tags. Conversely, Facebook expanded the tagging from photos to posts, with wider user control over tags, by which users other than the author can remove the tags associated with them.

Considering the adequate lower-level domain specific features should occur only after performing high abstraction analysis. Once the purpose of a tag cloud is clear, these lower-level features describe its intended *functionality*. In collaborative filtering systems users want to see the rationale behind the explanations (Herlocker et al., 2000). However, if users cannot influence the recommending factors, the transparency alone will add to the cognitive effort since it is necessary to account for another piece of information must. One of the solutions is to design tag aggregation pattern, and decide whether users can tag a resource only once using a single tag (set-model), or allow retagging (bag-model). The former is suitable for tag qualities, and the latter allows the generation of statistical data, i.e. rating tags based on their popularity (Milicevic et al., 2010). Within a photo-sharing social site, it would make sense to apply set-model and employ co-occurrence algorithm to recommend a similar resource to the user, while in video-sharing website promoting popular videos would be more valuable. The choice of tagging permissions can affect the very definition of the social structure within a domain, ranging from author-only, over permission-based, to free-for-all assigning of tags to resources (Milicevic et al., 2010). This is tightly knitted with tag aggregation, as the domain model boundaries are becoming clearer. Tags can describe a resource in several ways depending on the application domain, such as dates, locations, content, etc. (Heckner et al., 2008), and the better the anticipated tag use analysis is, the higher recommendation quality will be from the start. For some application domains this is an easy task to determine (e.g. social annotation software). Others are more difficult to assess, for example, video sharing or shopping websites where different information can relate to a resource. In a concrete case of video sharing sites, author, location and date, apart from content description, are predictable. Therefore the tag cloud recommendations can compensate for accuracy, regardless whether we may want to promote retagged authors and build their community reputation, or reduce their importance and focus on the content quality. In photo-sharing sites users mainly tag the locations and participants' names but we may want to focus on frequency-based recommendation. In other words, the frequency of viewing a specific photo (Ames & Naaman, 2007), or even better, based on retagging frequency or rating. However, many tag types can only are predictable up to a certain level: videos may contain tags associated with hardware reviews (device names, brands, models), or describing music video (artist name, song name, lyrics, live). Since all modern social websites have search engine incorporated, there needs to be a clear distinction of delegating specific functions to tag clouds, to avoid overlaps. One final distinction is necessary: while tag clouds can and should have domain-dependant distinctive functionalities, they also need to remain widely applicable. Users cannot afford steep learning curves when migrating among domains, adapting to significantly different visualizations or interaction styles. The essence therefore should not change, only the supporting features.

User Behaviour in Tagging Systems

Domain expertize distinguishes a subset of user types performing any search as experts or novices, mostly to create better recommendations by favouring the tags applied by experts (Christie et al., 2011). It received high item significance (8.67), and the impact on behaviour is a moderate (0.44), supported by solid conclusions. Regardless of the limited number of studies discussing this topic, the result was satisfying within the literature bias and the scope of this research. Domain experts use more efficiently search engines as opposed to novice users (Christie et al., 2010). They also have higher vocabulary similarity within the same domain, i.e. share common semantical representations of the same topic (Kang & Fu, 2010). Since tagging roots in the user's individual perspective of the resources, considering taggers who tag with higher flexibility and using various strategies more influential (Lin & Chen, 2012). While Kang and Fu (2010) propose favouring high-quality tags created by the domain experts as a navigational cue, there is a problem with converging vocabularies. In wider social systems, as the vocabulary converges. coupled with an increasing number of recommended resources and taggers, it becomes harder to segregate experts from novices (Chi & Mytkowicz, 2007). Secondly, there is no guarantee that experts' explanations (e.g. carried out through tags) will carry enough advising power to novice users and the effort invested in isolating these groups is questionable against the potential benefits (Wu & Bowles, 2010). Although using the expert knowledge is helpful to users with low topical knowledge, allowing excessive relying on navigational signals can lead to absence of an independent engagement with the content (A. F. Chiarella, 2012). Further research therefore must discover whether mining for expert knowledge is justifiable for tag clouds in larger scale social networks. Even if harvesting and embedding the knowledge is possible, another problem would be to settle and automate the balance of expert cues capable of assisting but not disengage independent interaction processes. Fu et al. (2010) proposed experts segregation based on differences in vocabulary convergences, assuming novice users differ from the experts based on the unique tags, where a higher number suggests a novice user. Although this logic is accurate, it is not quite enough since novice users can be early adopters which does not make them experts, since converged vocabularies more often reflect popularity than expertise. Finally, few studies had offered any useful algorithm or behavioural pattern that could aid in extracting the expert knowledge, even outside the analyzed literature corpora.

Expert knowledge had occasionally proven to be valuable and possible, for example, in educational environments where experts are setting the signals deliberately (Chiarella & Chmiliar, 2012; Dron et al., 2000). Whenever the wisdom of the crowds and motivation to contribute are likely to dominate, this approach is justifiable. Because this field lacks solid scientific evidence supporting implementational certainty, one solution is scalar tagging, with the problem of users becoming demotivated to use it after certain period because of its complexity (Dron, 2008b). However, scalar tags should not be dismissed yet, since it is possible to apply simpler mechanisms of rating tags with e.g. like or dislike. By applying those ratings to tagger's user profile, his/her reputation can grow, thus favouring those tags during the recommendation process. A positive feedback is crucial for emergence of self-organizing behaviour, since it amplifies the actions of others (Chiarella, 2011). This way it is possible to use vocabulary convergence since other experts are likely to agree, stimulating the experts' motivation to tag to gain the reputation. Most importantly, novice users can express the opinion on how they relate to the resource description provided by a specific tag. Such system would utilize set-model that would prevent retagging with a same tag to avoid producing false reputation, and selecting experts not only based on the popularity of their tags but the number of resources they had tagged. This solution however requires empirical testing in discovering its efficiency on user motivation over time.

Tag influence topic obtained an IS (8.87) and a moderate IF (0.57), indicating it as an important factor to consider, especially because the literature corpora was intentionally biased towards user acceptance and visual interfaces. A tag cloud has to visualize only a fraction of information available and convey it to users, giving them a social insight and potentially aid in serendipitous exploration or more specific navigational tasks (Allam et al., 2012; Sanchez-Zamora & Llamas-Nistal, 2009). Tag clouds have the potential of a learning tool for influencing novice users in a useful way. They introduce them to relevant information or concepts, allowing their knowledge of the domain to increase more rapidly by adapting their internal concepts to the social ones (Cress et al., 2013; Fu et al., 2010). Depending on recommender algorithm's design, that information can carry more or less persuasiveness, largely depending on the application domain. Amazon will continuously persuade the potential shoppers to buy items of interest to them, based on the likeness with other users. YouTube founds their recommendation in popularity, by weighing the ratings, subscriptions, etc. By overestimating the persuasiveness level it is possible to direct users to specific resources by manipulating their perception of its value, while underestimating is suitable for pointing to a range of resources, by decreasing the importance of a cue (Gedikli et al., 2014). Although this process can appear undesirable and suffering from a "dishonest car dealer" syndrome, high persuasiveness can be valuable if implemented through a transparent social insight. Users may want to be persuaded into selecting a resource based on peoples' opinions (Christie et al., 2011; Gedikli et al., 2014). This effect is even more highlighted in the learning environments, where users need a justification that goes beyond accepting simple "a priori" recommendation (Dron, 2014). Suggesting a useful resource can be helpful as users could have difficulty finding it through serendipitous browsing (Hotho et al., 2006). Although persuasiveness does help imposing a specific resource, it also reduces the efficiency of decisions and possible choices (Herlocker et al., 2000). While some websites use more aggressive visual means than tag clouds to influence users, the principles of operation are nearly identical – influence in tag clouds is the tag's ability to visually signal its importance based on the criteria set in in the recommender algorithm. By understanding how users respond to different navigational cues, both better tag visualization and recommending patterns can and should cohere to user expectations. Users respond to signals and likely to select the visually emphasized word over one that is not, even if they are have the same importance levels (A. Chiarella, 2011).

A topic on the opposite side of spectrum is *tagging imitation*, IS (9.49) and IF (5.67). No matter how strong the tag influence is, it is uncertain whether the users will respond in an anticipated way. In CoREAD, more users responds to high-signalling words, implying agreement with the collective opinion (Chiarella, 2012; Chiarella & Lajoie, 2010). It is possible the environment plays an important role in imitation. Education environments are inherently not malicious and relying on cues provided by others is logical. Media sites have users with varying tastes and views (even political), and imitation may not occur as easy. Besides, tagging imitation does not occur explicitly: users do not follow cues blindly, but use the power of internal associations against the cued ones (Chiarella & Lajoie, n.d.; Held et al., 2012). Most users, when presented with the possibility, will employ both personal and suggested tags, depending on the purpose (Panke & Gaiser, 2009). Undoubtedly, tagging imitation accelerates the domain's vocabulary adoption (Hotho et al., 2006). The biggest implication however is providing tag sets by the designer, to steer communities in meaningful direction (Hotho et al., 2006). Although tagging imitation alone does not provide implementation mechanisms, it is useful to understand how to align other features in providing domain-relevant guidance, for example, tag suggestions or recommender algorithms.

Communities of interest got a high item significance (9.18), and a moderate impact force (0.61) on user behaviour and (0.13) on recommender algorithm. This was a frequently analyzed topic, however, the results rarely become embedded in practical implementations. It is possible to view collectives on many scales, from animals versus plants, humans versus the rest of the animal kingdom, over to human-only sub-collectives, such as national, racial, religious, sports and music, etc. Whenever those collectives share some common interest in a particular ideology, they can form more or less distinct groups, which is the very principle of operation in social networking – we join, follow or befriend what we consider of importance. In nature, the actions of individuals are a foundation for group behaviour, classified in two broad families: direct collectives, stimulated directly by the actions of others, and stigmergic collectives, stimulated by the surrounding environment (Dron, 2014). Translated to social networking software, some users may prefer a direct influence by other users (followers, neighbours), and some will better respond to popular trends (news, music). If designing tag cloud or any social networking software in a deterministic fashion, there is a possibility of repelling one of these general groups. Up to this point, by adapting the recommendations through the application domain specific set of criteria circumvented this problem. For example, Amazon is interested in matching similar items bought, YouTube will promote popular videos, while Facebook will hybridize both approaches, depending on the content. Although this domain-dependent implementation is suitable for general recommending, in tag clouds the situation is different. Tags not only describe a resource or its content, they can add emotion, opinion, timestamp or geographical information, and any single one (or more) of these information can of interest to user, but not the other. This add complexity since there is a feedback loop – users input tags into the system, processed and aggregated by a central authority (algorithm), and fed back to users (Dron, 2014). For example, a user may want to view recent photographs of Spain, but he or she is not interested whether other users liked it or not, as long as they are recent. Another user may want the opposite, photographs of Spain with highest rating. A tag cloud that uses popularity biased weighing algorithm will not cater to the first user, while the co-occurrence algorithm will fail in providing the satisfactory results for the second one. Another group of users may wish to view their neighbours' resources, the rationale behind recommendations in a form of neighbours' ratings, or even engage in a discussion (Christie et al., 2011; Herlocker et al., 2000). Some users want to follow topics associated with a specific tag or a tagger (Wash & Rader, 2008), all of which is useful in e.g. learning environments. Hotho et al. (2006) propose extraction of communities of interest from folksonomies, by favouring top contributors' tags and resources, and in that way increase the overall community interaction. However, the extraction procedures of those needed elements have weak definition and constitute an ambiguous base for a concrete implementation.

Long tail is acceptable for emphasizing quality, but it tends to become stagnant (Dron, 2008b). In communities of interest not only vocabularies converge, but opinions as well, which decreases information value over time as it rarely changes. A possible partial solution to these problems could be to provide a limited user control over recommendations type, however, as the number of choices grows, less control is provided (Dron, 2008b), also inducing a high design overhead with questionable returns. On another end of the spectrum, some systems emerge due to a natural stigmergic self-organization, such as CoREAD (A. F. Chiarella & Lajoie, n.d.), or CoFIND (Dron et al., 2000). Such examples open the possibility of using well-motivated expert groups, gathering the necessary cueing and at the least imbue the necessary domain spirit. However, in wider social environments, this user recruitment would prove challenging. Based on this analysis one distinction emerges: the first question that must be asked is whether a tag cloud is necessary for a specific domain. The question especially relates to exploring if design effort could outweigh the benefits, in contrast to deploying a visually rich however different interface. For example, Amazon would have lower sales if they left out the product pictures and rely on a tag cloud, and the general popularity of resources does not concern them as the user interests are highly parcellated. In addition, the stigmergic development of online social networking systems dictates recommending algorithms to follow that trend as well, by being more organic, and cater to variable user interests and navigational goals. Since tag clouds have been “clouded” with an array of theoretical proposals, the following chapter will address steps that may help in mitigating some of these problems.

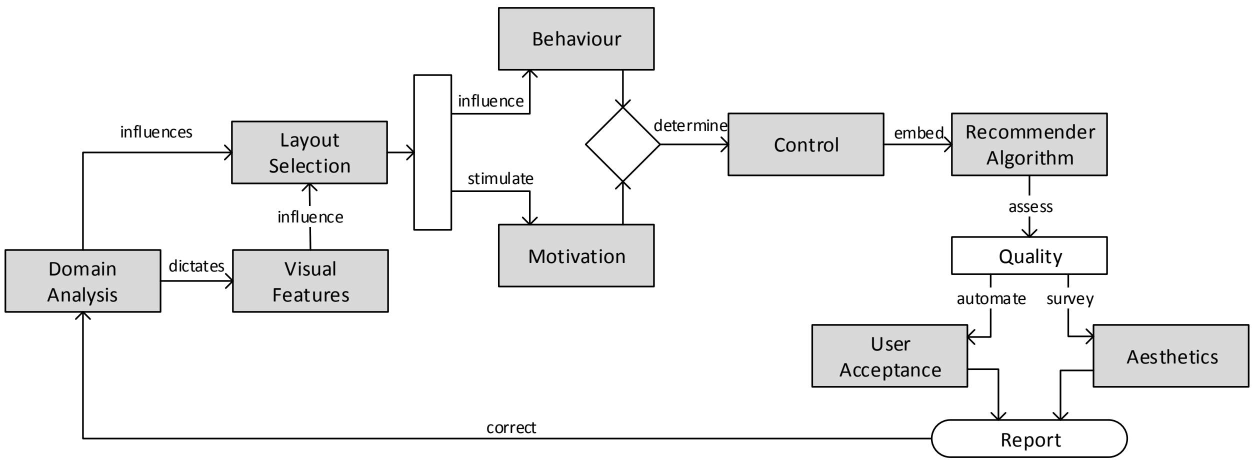
Chapter VI - DESIGN PRINCIPLES AND DISCUSSION

The theoretical model described in the previous chapter is unsuitable for any manipulation other than further knowledge enrichment. This was its intended purpose, and by summarizing the findings revolving around practical implications it opened an insight into the design practices and factors that had more or less success. As the later research develops, embedding any additional finding becomes an easy task that can potentially change the implementation odds of an individual feature or a topic.

Since it was impossible to manipulate the independent variables, the theoretical model in its abstract state provides little value for design choices, due to lack of fitting publications to discover the ratios or likelihoods at which the variables affect one another. However, translating the results into qualitative values mitigated this problem by reducing high abstraction level, and describing the basic principles for tag cloud implementation. All the relations in this chapter are presented using action diagrams, and in cases where a certain variable had a stimulating effect, it was assigned "*increases*" quality, and conversely. Before explaining the relations, a logical course of action was to form a linear framework. It entitles six distinct steps, using top-bottom approach and sequentially narrowing the design choices by using the outputs (design choices) of previous step as a guideline for the next (Figure VI-1). This approach was named S.E.C.U.R.E., or **S**crutinize, **E**nvision, **C**onvey, **U**ser control, **R**ecommend, **E**valuate. Considering many uncertainties surrounding tag cloud implementations, this method at the very least provides a secure outcome, and while the resulting tag cloud may not be the best performing or visually appealing, it will however not suffer from many acknowledged drawbacks. Beyond SECURE method, domain assessment, and visualisation results, all proposals represent only a *partial* solution of the current issues with tag clouds. These should be perceived as an example of the entire process, since no approach can claim achivement until empirically tested.

Figure VI-1

Proposed design principle (SECURE)



Domain Analysis (Scrutinize)

One of the most overlooked steps when deploying a tag cloud is domain assessment. For example, is another visually rich interface competing with a tag cloud, or should it increase user attraction through its beauty versus the ability to enable serendipitous exploration. Visualizing photographs or videos can be a primary tool in providing recommendations, however tag clouds' comprehensiveness can provide a different perspective into site's activities. For websites that recommend using rich visual stimuli (pictures, videos), tag cloud should not compete with those cues and remain discreet, best in a form of on-demand feature or assuming an orderly and informative layout. The second applicable role within these domains could be to increase the attractiveness of specific subdomains. For example, Amazon has little attractiveness on sellers' store page, and a tag cloud could give a quick and attractive insight into relevant activities, collected from tags the seller assigned to own resources. In this case, the recommendation type would not be crucial whether using co-occurrence or other weighing algorithm. The third alternative may exclude tag clouds, if the existing recommendations fulfill their purpose. For domains that do need a tag cloud, the main question is whether to deploy primarily for increased attractiveness or information organization. Do users need guidance or to serendipitously explore the domain? The former imposes sequential or list layouts as a logical choice, and the latter can use any layout. Unfortunately, there is no defined process in structuring these decisions, but perhaps the best guideline is weighing the implementation overhead with the potential benefits.

While domain analysis is suitable for any software design, most domains employ deterministic tag cloud implementations. If any product is too narrowly defined, then the danger of complete failure emerges, since there is no room for adaptability. The literature revealed the lack of supporting evidence that any form of tag cloud so far has performed notably well. Firstly because of strict design rules applied against uncertain online social behavioural patterns, and secondly, it always used a computational approach as a starting point. The question “What is the best way to visualise data?” should be replacing “What is the best way to visualise **my** data?” Despite users’ preference for attractiveness over performance, there should be a delicate balance of both. Too much attractiveness transforms a tool into a toy, and high performance needs visualizations that can be dull. Because of these findings, a conclusion imposed itself: it is imperative to perform domain analysis using high abstraction, mainly to answer the question whether to employ tag clouds and in what capacity.

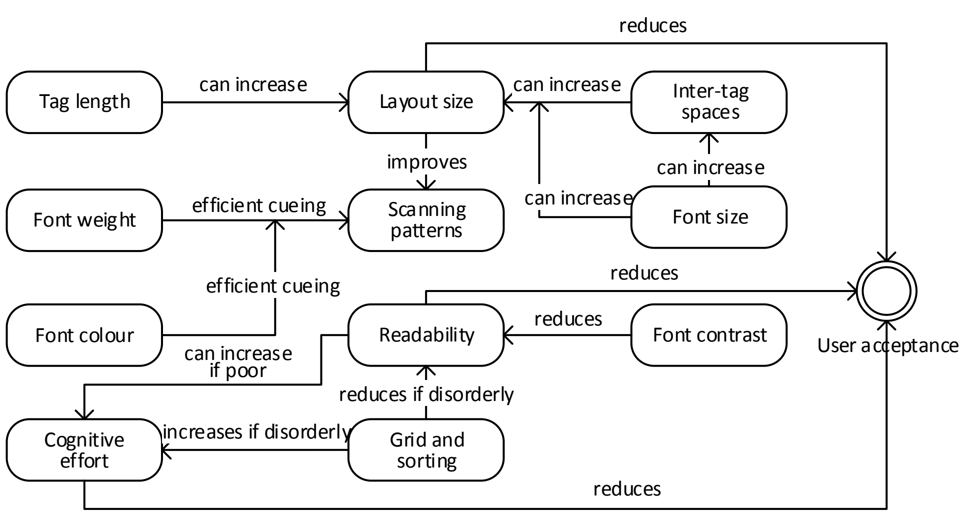
Layout Choice and Tag Visualisation (Envision)

Although the second step of the design framework determines the correct layout, it is important to understand the rationale behind such a choice, indirectly imposed by the associated visualization problems and benefits (Figure VI-2).

Any significant increase in *tag size* (size, length), will produce an increase in inter-tag spaces and layout size, which will result in reduced user acceptance. With an increase in layout size, the efficiency of scanning patterns increases as well. This is not of obvious benefit to users, and although the suitable layout size is unknown, there are indications it should occupy only a smaller portion of the overall user interface. Tag size is the most powerful navigational cue, as univocally found in the literature, and using extreme values can question the purpose of smaller tags or other navigational cues.

Figure VI-2

Visual factors' known relations.



With tag size increase, the chances of selecting those tags nearly proportionally decreases, not to neglect users facing a strong persuasion as opposed to recommendation.

Principle 1: Tag size variance in a tag cloud should not be significant, unless designed solely for attractiveness.

Principle 2: Inter-tag spaces should not assume any of the extreme values, neither too wide nor too small to produce tag overlaps.

When compared with tag size, combining font weight and tag colour provides at least as equally strong navigation cue, however more subtly, resulting in an increased scanning pattern efficiency. This effect is observable in sequential layouts, yielding the highest score. Tag contrast should be used with extreme caution as it reduces readability and directly affects user acceptance, however it can be potentially useful as an indirect navigational cue. For example, when a user hovers with a mouse over a specific tag, fading the unrelated tags to highlight the relevant ones.

Principle 3: Do not employ tag contrast as a guiding navigational cue since it reduces readability and user acceptance.

Principle 4: Proper readability is imperative to user acceptance, reduces the cognitive effort and requires evaluation through the entire design process.

The first two principles do not apply to clustered layouts, since inter-tag spaces and tag size variance do not carry any adverse effects, if using them to signal clusters (through clear separation), and main topics (larger tag size). Although the preliminary study favoured clustered layouts as informative and organized, the analysis revealed that all the implementations with relative success were performed on either large screen formats or consuming most of the user screen. This practice out of alignment with tag cloud environment, supposing to leave most of the screen for listing resulting resources, a role that is similar in essence to a search engine. This does not imply disregard for clustered layouts, but only in special cases when occupying most of the screen is justifiable. For example, browsing tagged photos, setting the primary cues on visually rich media that can provide faster exploration than semantical descriptions.

Principle 5: Use clustered layouts only in cases when sacrificing most of the user interface is justifiable to achieve goals. Clustered layouts do no perform well in small, constricted areas.

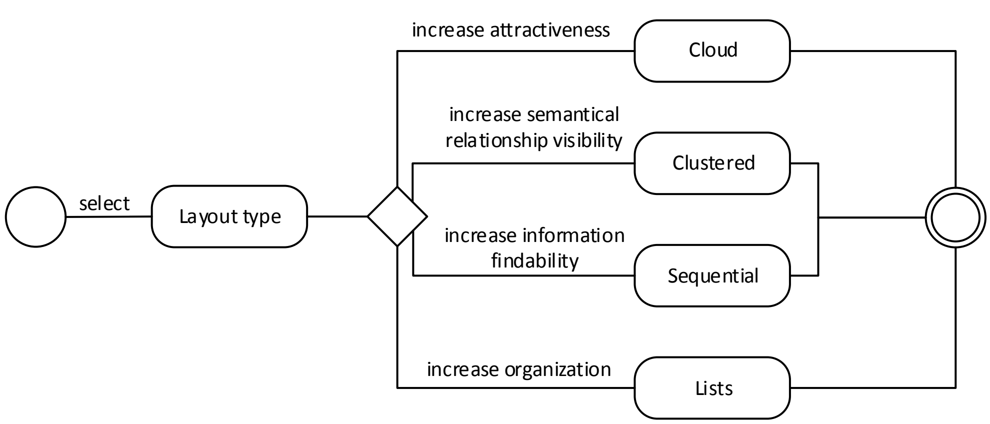
The analysis clearly proved tag clouds' inability to provide self-contained navigation other than pointing to popular or shared resources. Browsing the tag cloud has a high *exploratory* and a low *explanatory* value to users, therefore its primary purpose is serving as a pointer to more informative resources. With the increase in research efforts to attach more complex features to tag clouds, there is an obvious decline in their success, mainly because of the well planned recommending computation "meeting" a constricted user interface.

Principle 6: Because tag clouds have to perform in rather constricted portions of the user interface, their role in navigation is supportive as opposed to imperative.

When selecting domain-appropriate layout type, the choice is relatively easy to make since clearly defined benefits of most used layout types align with domain's intended purpose. Sequential and list layouts are adequate for information structuring, whether using alphabetical, weighted, frequency-based or ontological algorithms, however, often lacking attractiveness because of well-organized interface, resembling a typical spreadsheet software. Cloud layouts do have a high visual appeal, but are inefficient for any structured navigation, especially because of their dynamic and unpredictable tag placement. If preventing users from establishing some browsing pattern, the cognitive effort will increase and can be repelling for exploration. This does not mean such tag cloud is useless, in fact, it contributes to site's visual appeal, especially if used with personal set of tags; it has been proven that users like to have their own description in such a manner. Having personalized and at the same time attractive tag cloud has the potential to increase user motivation in assigning tags to resources. Finally, selecting recommender algorithm is less relevant in case of cloud layouts, since visual cueing cannot support any recommendation complexity, resulting in lower design effort.

Figure VI-3

Layout type selection rationale.



Figures VII-4 through VII-6 describe the visual implications presented in the previous chapter (Figure VI-4, Figure VI-5, and Figure VI-6).

Figure VI-4

Clustered layout visual design implications

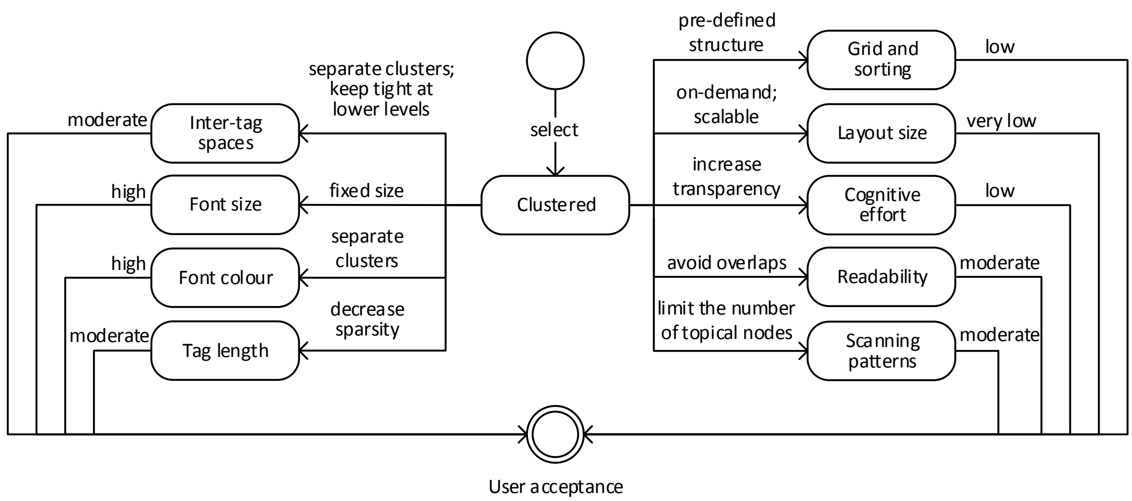
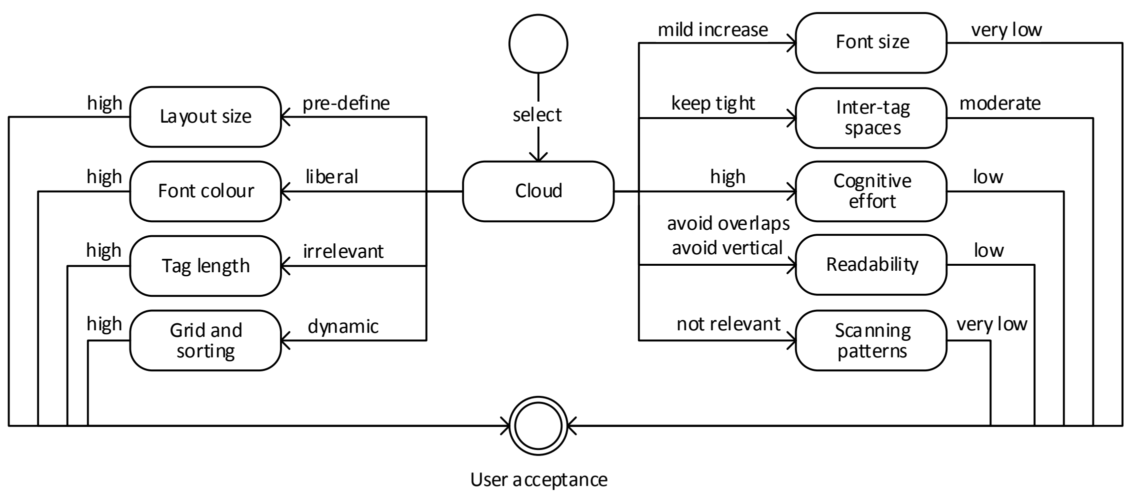


Figure VI-5

Cloud layout visual design implications

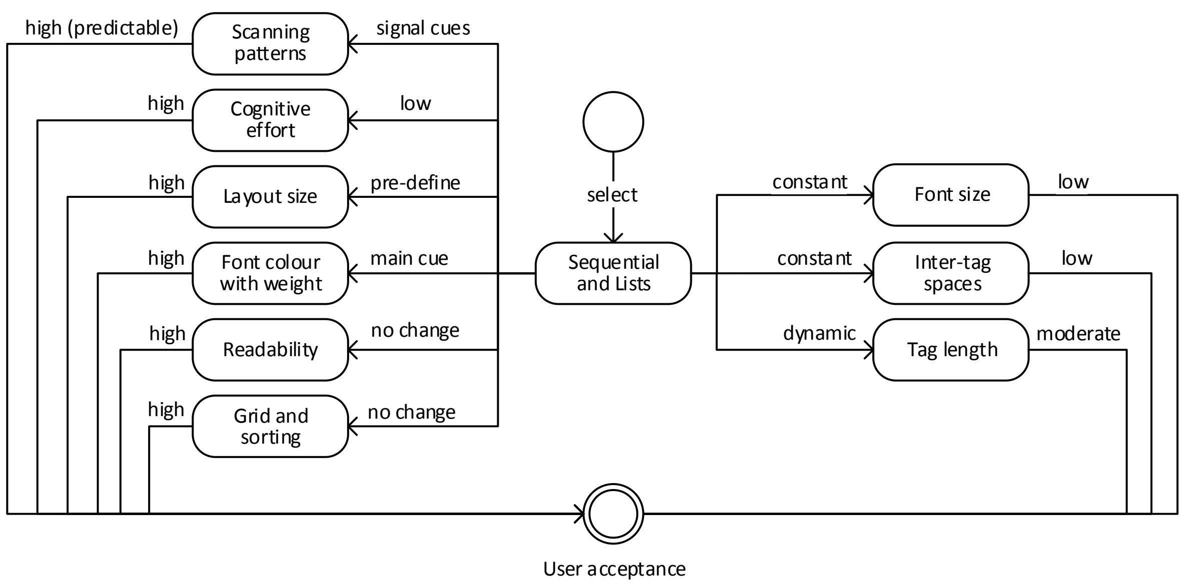


Principle 7: Deploy either sequential or list layouts when domain's purpose can benefit from conveying or organizing information.

Principle 8: Deploy cloud layouts when domain can benefit from an increased visual appeal, and exploration and recommendation is provided by other means or not essential.

Figure VI-6

Sequential and lists layouts visual design implications



Motivation and Behaviour (Convey)

While direct user type segregation using statistical analysis has proven to be moderately effective, it is necessary to assure the design will adequately serve the domain users. Tag clouds have the power to *influence* crowd behavior up to an extent, and *stimulate* motivation to assign tags to resources and use tag clouds for navigation. Although many studies discussed these areas, the focus was on discovering technical features that can aid in these tasks. From the behavioural aspect, three promising areas emerged: communities of interest, domain expertize, and tagging imitation, when combined having the power to cater to a wider range of users.

Studies on communities of interest have been extensive and with great success within visually rich recommender systems (e.g. Amazon), however the opposite logic applies to tag clouds. Because of limited user interface, both spatially and cognitively, it is impractical to visualize neighbour ratings or follow other users. There is no need to increment the complexity of an innately simple recommendation system, especially when other social networking software (e.g. Twitter) nowadays offers such features in a form of the Web plug-ins, which can complement a social domain as opposed to competing. Since in communities interests are parcellated, users need more focused descriptions. For example, a music domain is topically distinct from a movie one, however with different genres, and not all are applicable to user tastes. Offering subscriptions to topics/genres (or by liking them) can help in shaping the tag cloud, and *informing* users of any topical updates. In educational domains this is not necessary as users need to guidance to a specific (sets) resource, therefore resource relations need to be in focus. Regardless of the application domain, by following memberships in different communities it is possible to recommend potentially interesting topics or even new communities, thus influencing the user when making choices.

Domain expertize is useful when users need guidance to a specific sets of resources, by shortening the navigational path. If considering expert knowledge in design, the associated quality of tagged resources is likely to provide an increase in results accuracy, influencing the selection of certain navigational path to one that system thinks is better for a user. However, domain expertize is difficult to harvest outside motivated crowds willing to volunteer the knowledge. If such user guidance is needed, then the best realization is through motivational mechanisms, for example, personalized tag clouds that grow with versatility and involvement with tagging. Another approach is to increase the social status of those taggers through reputation, which proved to be effective in online forums.

User imitation of others' tags results in faster domain adoption, specifically transparent through vocabulary adoption. That behaviour can be influenced to either further accelerate this process or increasing the quality of assigned tags, with autocomplete and tag suggesting being the best performing mechanisms. Autocomplete reduces spelling errors and prevents those tags removal by spam filters, and reduces lingual barriers. Tag suggestion reduces tagging cognitive paralysis and can increase the quality of resource description. However, if suggesting tags solely from the popular pool, then the result may be an irrelevant resource's high ranking in search results. This can be resolved by either blending personal and popular tags, or from expert knowledge pool. The latter solution could be desirable in education domains as it would allow novice users to promote content with higher quality. Considering that educational domains mainly entitle non-original content, e.g. works from various authors, and suggesting from expert pool could enable a novice user to effectively start contributing.

Five factors emerged from as having the highest potential to affect user motivation in tagging systems: tag currency, tag filtering, personalization, tag qualities, and cold start (Figure VI-7). Tag currency is a feature able to distinguish tags based on their life timeline in the system. This can be either by automatically filtering tags based on their pre-set time range, or allowing users to gain such an insight on-demand. For example, an educational environment will benefit from older tags with longer tail, while a technology-oriented domain will normally favour recent ones, because of the continuous changes in popularity trends. Since this practice is deterministic in recommending, perhaps a better approach is to allow limited user control over this process, by implementing on-demand overview or allowing for selection of custom time ranges. If on-demand, this feature can contribute to visual clutter and negatively affect user acceptance, therefore a mouse click-and-hold action would be fitting for two reasons: the necessary visual clutter is temporary, and user is present with a simple option that reduces the overall complexity of the interface.

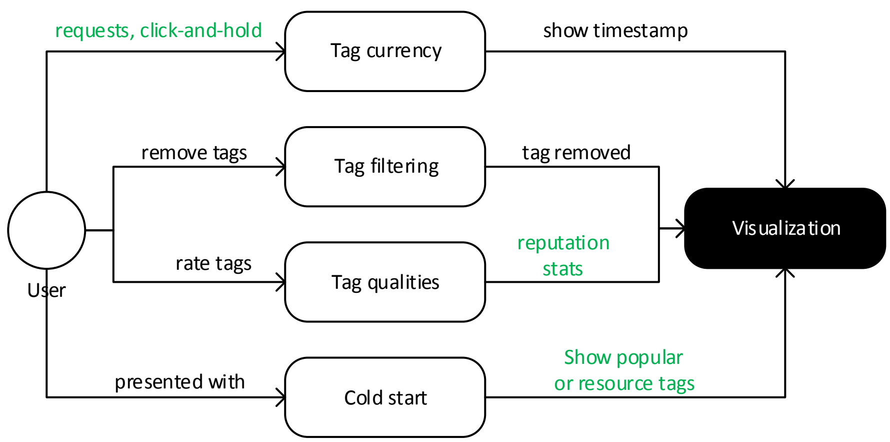
Similar logic applies to tag filtering, where a user can remove the unwanted tags, either by employing drag-and-drop move to a disposal field, or check-marking the tags and clicking on a “delete” button. By removing unwanted tags, the *personalized* recommendation precision will increase, and so will the trust in the system. Besides, tag filtering has the ability to increase this precision far more efficiently that any recommender algorithm, since certainty replaces estimation. As a precaution, there also should be a small database with removed tags, allowing the user to restore an accidentally removed tag or is necessary for other search type.

Personalization can additionally motivate users to tag by providing a user with insight into quality and versatility of own tags. If populating the personal tag cloud exclusively with tags a user assigned to resources, it provides an incentive for richer tagging both in quality and quantity. This functionality has the potential to overturn the cold start problem into an increased motivation to tag, since populating personal tag cloud becomes a fun goal.

Cold start is detrimental to user acceptance, with no operational solution so far. However, cold start can avoidable by applying previously mentioned social insight, similar to recommendations on Amazon or YouTube. In other words, as soon as the novice user begins interacting with a domain, presenting her/him with the most popular tags, instead an empty cloud. As user assigns tags to resources, those apply against popular ones, and the recommendation starts reflecting user preferences. This approach is acceptable for two reasons: a user gains an insight into domain's activities, enforces the vocabulary adoption, and with increase in interaction the popularity-biased effect becomes less emphasized.

Figure VI-7

Mechanisms for increasing user motivation.



All the mentioned motivational features were selected because of their proven benefits and the relative ease of setting up, however considering other choices is advisable if new evidence supports their effectiveness.

Administrative and User Control Planning (Control)

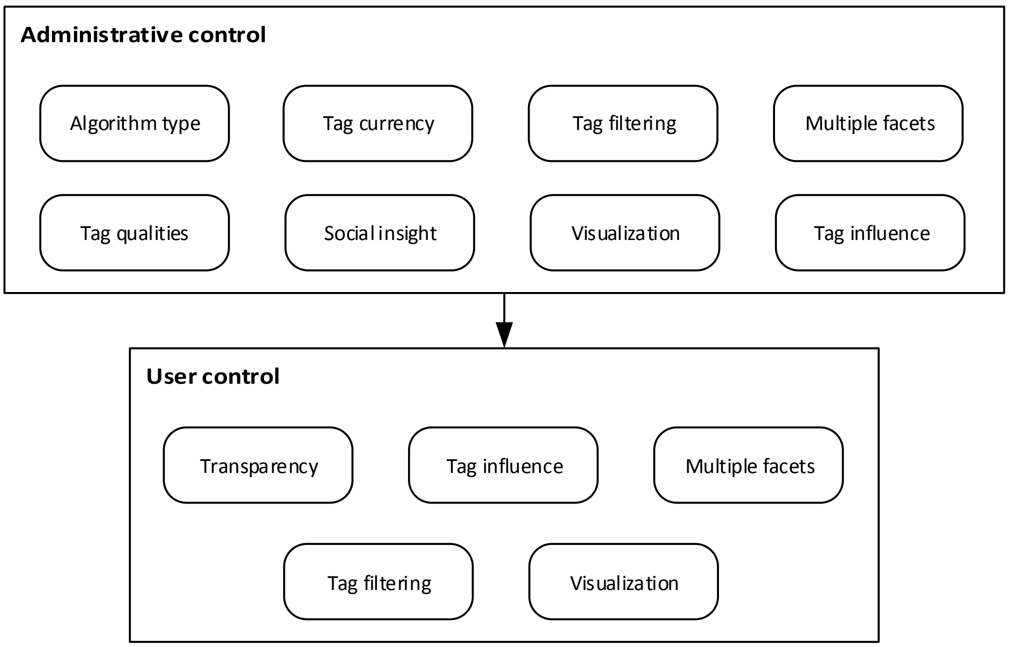
The analysis of user control for potential benefits is important, since too much control over a simplistic interface can lead to confusion and reduced user acceptance. Many newer studies however expressed the need for broader user control over tags clouds. Therefore, it was necessary to select a subset of features for the task since implementing all would create a high overhead. Those features introduce the lowest possible complexity and potentially provide highest benefits, directed at catering to diverse user preferences.

Setting up the administrative control introduces a human factor in determining the amount of control users should exercise. For example, the main characteristic of Amazon is high persuasiveness and too much control could impact their sales, especially with impulsive buyers, while YouTube could benefit as their earning system does not depend on specific products. For example, commercials from nearly any video increases their profits, and better tailored recommendations are favourable. Almost no recommender system will employ a single algorithm, and with domain growth in both resource and users there might be a need for recommendation refinement, driven by the changing goals or technology improvements. When adding or improving algorithms, a designer can estimate but not know the exact effects, which is difficult to correct in the post-implementation stages. Providing the administrative choice to engage or disengage the algorithm at will reduces this effect significantly. Furthermore, the added control, regulating the ratio at which algorithms influence one another can aid in active overseeing, refining, and acting on results. Considering that tag clouds are usually not an essential feature for many domains, this ability is optional, but necessary in the experimental environments.

Many studies suggested multifaceted browsing as desirable, but with few implementations or the intent of catering to differing user preferences for visualization. There is a plethora of circumstantial evidence supporting the need for this feature – in studies that compared significantly different layout styles or visual sets, there never was an absolute user preference, but rather minor differences. Contrary to this, studies that deployed more supporting features directly to user interface, e.g. frequency bar chart under a tag, yielded low user acceptance. Setting up multifaceted browsing is achievable by either on-demand views change, or through dual-view interface, and possibly combining both. For example, layout changes from cloud to lists, or personal versus public tag views.

Figure VI-8

Administrative control transition to user control



Principle 9: User controls and different visualizations are desirable, however using subtle values and features. Any visual overlaying with the supporting functionalities contributes to visual clutter and will negatively affect user acceptance. User controls should be clear and simple to avoid any increase in complexity.

The controls for this feature should not be obscure and hidden in some menus, but easily accessible and transparent if user accidentally selects it. In Figure VI-8 transparency is not present in administrative controls, since this is not a negotiable feature: no administrator privilege should influence it (Figure VI-8). On the other hand, users should be able to disable help features once their interaction skill grows, or integrate it using button that pictures explanations for the entire interface, e.g. using callout boxes. The latter solution is perhaps a better one, as hover-on dialogs can interfere with the tag cloud visualization and static explanations would consume an already small space allowance.

Options that cannot be delegated to users are the ones integrated into domain, in this case tag reuse and tag qualities (rating system), however this presents no problem as users simply opt not to use tag suggestions or rating system. Any features not affecting tag cloud navigation are not perceived as intrusive if not aggressively embedded to affect the overall user experience in interacting with the domain; for example, a popup screen asking users to rate a resource or tags.

Some users prefer to follow popular trends while others seek resources similar to their own taste, resulting in differentiation that is the core problem in any recommending. *Tag influence* is an optional feature that allows users to adjust the recommendations towards personalized or public bias, with a goal better tailor to user preferences and a potential increase of trust in the system. The change does not have to be overall significant and it can prevent from employing extreme values, especially if dual-view interface is the primary choice, depending on the previously set administrative control and domain’s purpose. This feature also depends on a suitable recommender algorithm, capable of dynamically changing its results. Unlike some simpler user options, this one should not be accessible directly from the interface, as inexperienced users are unlikely to understand the implications. Whether the administrator will enable it depends on the application domain. For example, in educational environments this would be useful and users with academic background should be able to understand it. In wider social domains may cause confusion, justifiable if information discovery is the primary objective of a tag cloud. The relevant labels should not be obscured by technical terms, but downgraded to a comprehension level of an inexperienced user, such as “Sort by popular tags” or “Sort by related tags”.

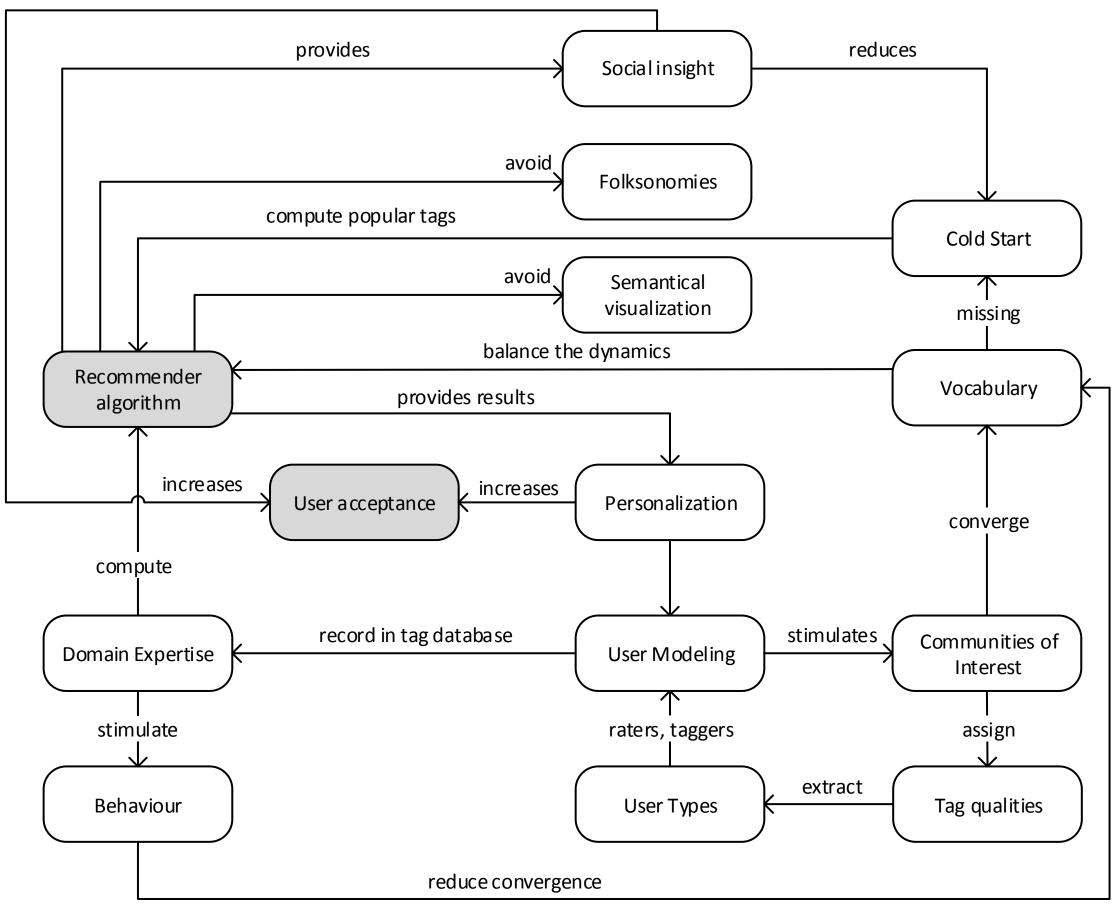
Tag reuse has a positive influence on user interaction with a tag cloud, because it reduces the cognitive effort associated with tagging. By suggesting popular tags, the system can empower the users to increase the popularity of their resources or their motivation to frequently take part in tagging. Although tag reuse is sometimes pointed out as a potential contributor to spamming, this scenario is unlikely. Malicious users would have to invest more effort to click and drag the tags compared with current copy-paste action, whereas inserting a whole set of tags is much easier. Its accuracy is however questionable since proposing the popular tags does not necessarily imply their relevance for the given resource. An interesting approach would be to design for a supporting algorithm that calculates a few top *n* popular tags and the ones from personal collection. For novice users the suggested tags would be entirely popular, and an increase in interaction would blend in the personal tags, which partially resolves tagging cold start.

Recommender Algorithm (Recommend)

By following previous steps, the recommender algorithm’s main functions should naturally emerge, in defining what the requirements are with what the designer aspires to deliver. Simply by starting from a user perspective, the types of recommendations become less important, and whether employing weighted, co-occurrence or hybridized algorithm, the success is likely to be higher than approaching it simply from a computational perspective (Figure VI-9). The algorithm choice should strive for simplicity, since some of the more complex solutions failed to produce satisfying results, whether because of limited visual space or inability to extract the necessary user data to build upon.

Figure VI-9

Recommender algorithm perspective



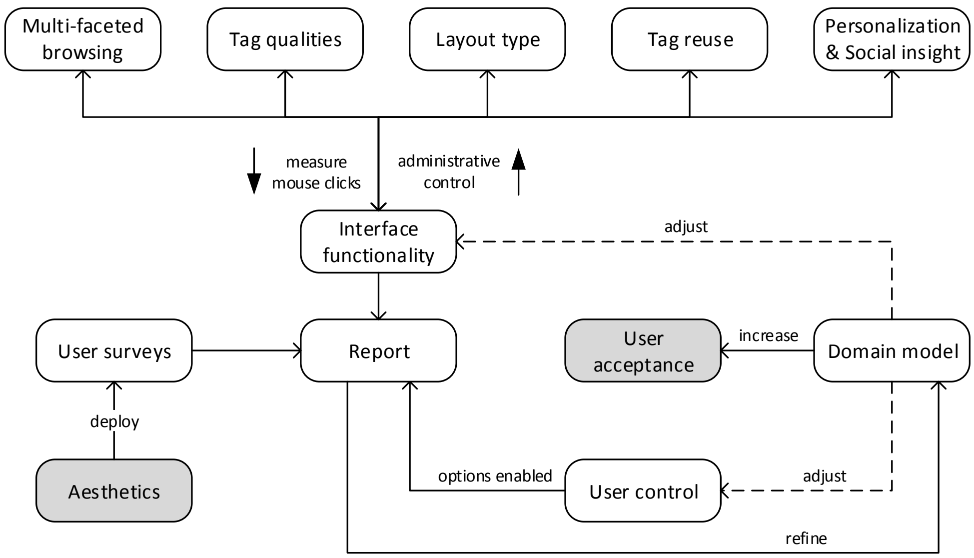
There is no record that any ontological implementation was successful, and therefore should be avoided unless an alternative approach brings a creative spark. The only current solution is in human-generated taxonomy providing a base for algorithmic extrapolation, but there was no evidence of such systems in operation. Attempts at semantical extraction are countless, without a promising base that could serve as a starting point. The vocabulary convergence is the foundation for extracting both ontologies and semantics, but it becomes stagnant and so do the results. This means that in current state semantical extraction creates a barrier for the natural dynamics of tag clouds. A firm finding of this study is the early testing of these approaches should occur within a socially richer environment that has fewer visual limits. For example, extracting a forum vocabulary and presenting the semantical relations in full-screen, and later adapt the gained knowledge to tag clouds' spatial constrictions. This also affects clustered layouts, best suited for exactly representing semantical and categorical relationships, and considering their hunger for space, the prospect is not promising now.

Step 6 – Quality Analysis (Evaluate)

Quality analysis should focus on iteratively assessing and fine-tuning any performance issues that may hinder user acceptance, accompanied by domain model evaluation. For example, whether the user habits or domain purpose have changed direction or evolved. Compared with a traditional desktop software, social networking is more dynamic and dictating, with rather unpredictable outcomes imposing a bigger burden when striving for quality. It is possible to divide this assessment in two categories: *automated*, for example, counting the number of mouse clicks against certain feature, and *human*, e.g. using surveys. While it possible to assess certain feature's frequency of use and afterwards disable or improve it if the results are not favourable, evaluating aesthetics is only possible by interacting with users and collecting opinions. Reaper, a digital audio workstation, although not so powerful in features as more expensive competitors, gained enormous popularity through forums in which users may suggest changes and witness those changes incorporated. In a similar fashion, it is important to collect the wisdom of the crowds and instead of constantly changing the approaches, persistently attempting to improve the existing solution. Low user acceptance does not necessarily imply a low quality tag cloud, but it can point to its unsuitability for the domain's purpose.

Figure VI-10

Quality monitoring and control



Principle 10: A tag cloud, like any other software, must be assessed and improved, since no software fulfills its expectations immediately outside-of-the box. Whenever possible, try to align the available features with the application domain's purpose.

` Administrative and user controls allow for a broad range of combinations, and continuous refining could bring positive results. For example, a news portal will have little value from a personalized tag cloud since its social aspect is open and dynamic. However, in a tightly knitted social environment a personal tag cloud can describe the tag(s) owner. Wordle users print out their personal tag clouds on coffee cups and shirts, which can be motivating. Even the smallest feature set could be a good starting point, but once introduced it needs continuous assessment to achieve favourable user acceptance.

Reducing to Realistic Boundaries

It may not seem practical for all designers to engage in deep analyzing process when designing tag clouds, however the presented S.E.C.U.R.E. method can reduce the planning effort. Not all features require implementing, and it is not always necessary to conduct quality analysis, for example, with personal blogs, where such effort would be unjustifiable against the benefits. However, this method conveys an important message: choose tag cloud visualization according to planned domain purpose. Users may not have choices available at all, but transparency must be present, unless a tag cloud serves as a decoration reflecting domain's purpose. The entire analysis and the proposed method was conducted and designed in support of larger (wider) social network software, where the associated design effort is justifiable. A tag cloud that would use all the proposed features can only have two sources, either commercial or academic, since the associated overhead in designing the entire solution is high. The full set of features serves experimentation purposes, and ideally in the future it will contribute to full interpretation of factors worth carrying out and in what capacity. Only when there is enough evidence stemming from practice, it will become possible to discuss more deterministic models with higher certainty of success.

Answering the Research Questions

All the research questions were answered within initially set scope, mainly because the weighing system for literature selection and the statements provided a critical perspective that would not be otherwise possible. However, the confidence level for most navigational factors is low, resulting from conflicting and parcellated methodologies and perspectives across the literature. The grounded theory alone would have presented a challenge in separating the facts from educated speculations and the quality of this work would be repetitious in following the settled paradigms. At the least, this study provides a fresh perspective into designing tag clouds, and although some claims were exclusive or restrictive it was to only advise on operational and promising solutions and approaches.

1. *Which significant factors the relevant research suggest affect the likelihood of an individual choosing a particular navigation path in a tag cloud?*

Most individuals are likely to select the path dictated by the visual cue; novice users are likely to follow them regardless of the recommender algorithm suggestion and even if it contradicts the power of internal association to topic. For intermediate to expert users this is not quite the case, since they match the cues provided with internal associations, and choose more informed paths. The semantical construct, if present, is likely to dominate the visual cues. The most influential **visual cues** are (in the order of list):

1. Tag size; strongest influence on tag selection.
2. Tag contrast; negative influence on aesthetics.
3. Tag colour; positive influence.
4. Tag length; negative influence, difficult to control.

5. Tag weight; positive influence.

Layout properties:

1. Layout size (if increased); positive influence on navigational paths structuring; negative influence on user acceptance.

2. Inter-tag spaces (if increased); no influence on navigability; negative influence on aesthetics, except in clustered layouts.

3. Sequential and list layouts; positive influence on navigational paths structuring; moderately negative influence on aesthetics.

4. Clustered layouts; positive influence on navigational paths structuring; negative influence on user acceptance due to necessary size.

5. Cloud layout; negative influence on navigational paths structuring; positive influence on aesthetics and user acceptance.

Recommender algorithms (limited list):

1. Tag co-occurrences; suitable for resource discovery.

2. Semantical structuring; theoretical but promising.

3. Tag weighing; biased without the support of another algorithm type.

4. Ontologies; theoretical; no operational system;

2. *What evidence does the relevant research provide on the interrelationship properties of these factors in respect to structuring the navigational paths?*

One of the most important contributions of this study is a clear segregation of visual elements based on a specific layout, since not all visual properties have uniform impact or role. Figure VI-4 through Figure VI-6 describe the visual interrelations within layout context, and chapter V provides a detailed analysis of each factor (Figure VI-4, Figure VI-5, and Figure VI-6). Recommender algorithms largely influence interrelationship properties, both visual and computational, however that topic was not in focus of this study.

3. *Which factors primarily define a tag cloud's successful implementation?*

This entire chapter is dedicated to answering this question, by introducing ten principles of tag cloud design and proposing several solutions. Answering this question resulted in S.E.C.U.R.E. method that incorporates six concrete steps to successful implementation. This approach did not address recommender algorithms, awaiting subsequent research, however visual, motivational and behavioural analysis provided evidence sufficient to influence the current practices in tag cloud implementation.

Chapter VII - DEVELOPING THE SOFTWARE ARTEFACT

The artefact was designed using Microsoft Visual Studio developing environment and implemented as a platform-independent Silverlight plugin. The benefits of this approach are high portability and programming environment, transferrable from one developer to another. Two goals guided the development: proving the validity of S.E.C.U.R.E. method in practical design, and visualizing the theoretical results of this study.

Scrutinize (Domain Analysis)

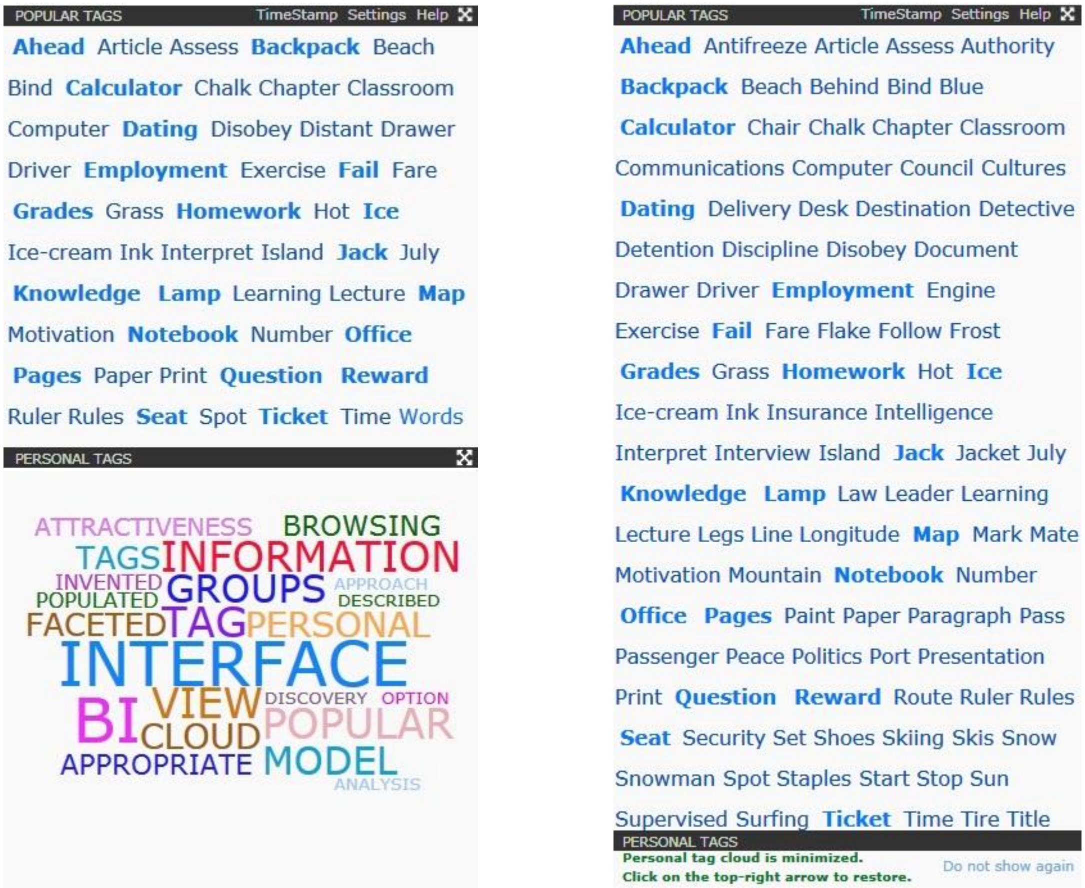
Further research and experimentation served as a rationale for building the software interface, therefore incorporating most of the previously mentioned features and visualizations. Since the targeted domain will potentially be of wider scale, all the functionalities were enabled and managed based on the user acceptance analysis.

Envision (Layout and Visual Elements Selection)

The interface in this tag cloud needed to satisfy bi-faceted model, which would allow users to see both popular and personal tag sets (Figure VII-1, left). Although that approach would have catered to both user groups, it was necessary to consider that certain ones would like to have only one of these views. This may not be a continuing requirement, but the need to browse for current information or print out/view their persona as described by tags, and called for the ability to turn the dual to single view (Figure VII-1, right). Since personal tags are used for browsing in minor number of cases, this was a suitable role for cloud layout contributing to the attractiveness of the interface. The opposite logic applied to popular view, where information browsing and discovery should be the primary purpose.

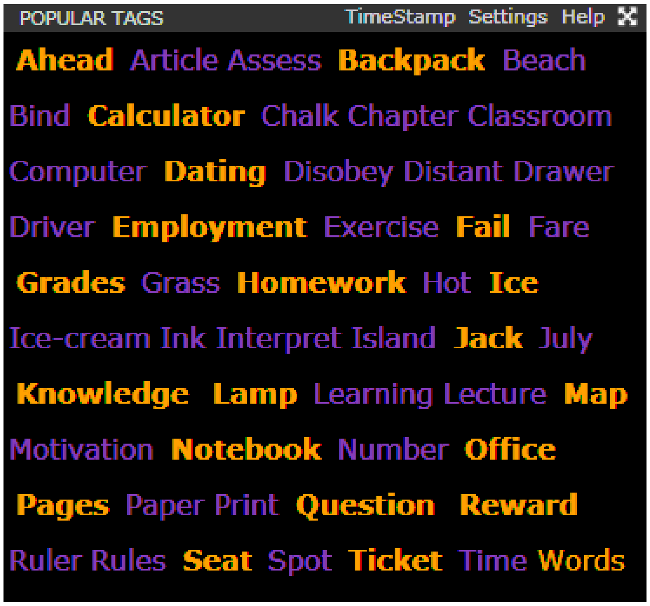
Figure VII-1

Dual and single view interface



In this setup, the popular tags assume alphabetical ordering by default, while presenting the personal tags using cloud layout. By acknowledging the alphabetical layout as the most efficient for information retrieval, and cloud layout choice as visually appealing to users, opened the possibility for even wider experimentation. In single view mode, a popup warns the user of the current layout state, prompting for action. In this example, clicking on the self-explanatory notification returns to dual-view mode and minimizes personal tags, which is useful for novice users.

Figure VII-2
Visual cues



A discreet spacing around keywords and weighed font make an equally strong navigational cue as the increased tag size, especially if coupled with colour (Figure VII-2). Although the background colour is probably best left matching the domain visual theme by default, allowing limited control over it can be valuable for personalization. The changes in colouration not only increase attractiveness, but certain combinations can aid colour blind people in providing effective cueing. In comparison to white background (Figure VII-1), the black background with or without coloured keywords (Figure VII-2) provides a significantly better contrast. Although this is an extreme contrast, it explains the value of colour alone as a navigational cue, without the ill effects that tag size can sometimes cause.

Conduct (User Motivation and Behaviour)

Although there are many possible approaches to increase user motivation, the one proposed in the previous chapter involved rating the tags. The need for high user involvement is detrimental to success of scalar tags, however its potential is not dismissible. Instead of rating individual tags, a potentially better approach is rating an entire tag set.

Social insight can be rather easily drawn from the popular tags and personalization from user profiles, but harvesting domain expertize has proven to be challenging and theoretical, mainly because it depends on a statistical user type segregation. It is however possible to *stimulate* users to *volunteer* domain expertise as opposed to *extracting*, by employing scalar tags in a form of simple ratings, such as “like” or “do not like” (Figure VII-3). If allowed to rate tags, and applying those ratings against the taggers profile, then it is possible to motivate further tagging through a reputation increase. Higher rated tags will consequently have recommending priority, leading to more valuable resources or influential taggers, depending on the algorithm type and domain’s purpose. Ratings also have a higher motivations incentive than e.g. tag tagging, especially if the rating itself contributes to an increased reputation in smaller proportion. This opens the possibility to collect domain expertise in a manner that better reflects a wider range of users’ comprehension levels. For any content described by tags, the quality of those descriptions is as important as the quality of the content itself. This especially applies to novice users, whom have less knowledge and ability to assess the resource quality, for easier understanding of resource content descriptions (provided by tags). While there is no real problem in rating resources’ quality, voting for tags can add another dimension. Votes for content quality can promote it higher in search engine's rating, but not in tag clouds, whereas voting for tags can perform both roles.

Figure VII-3

Domain expertize proposed solution

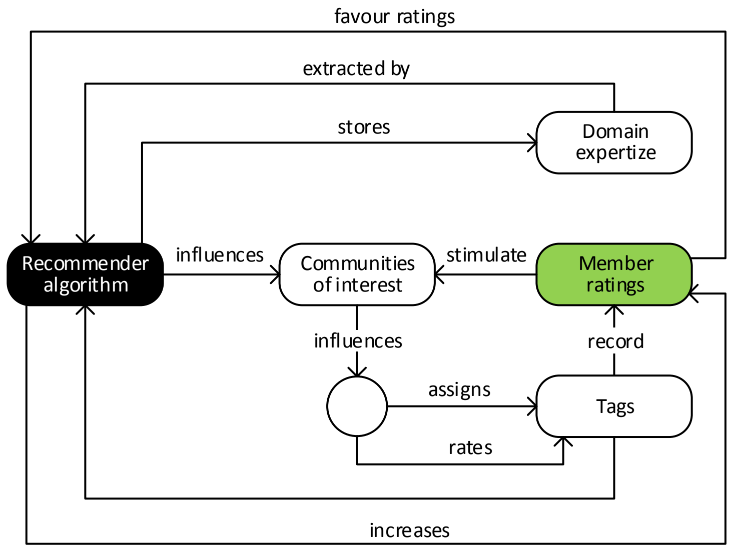
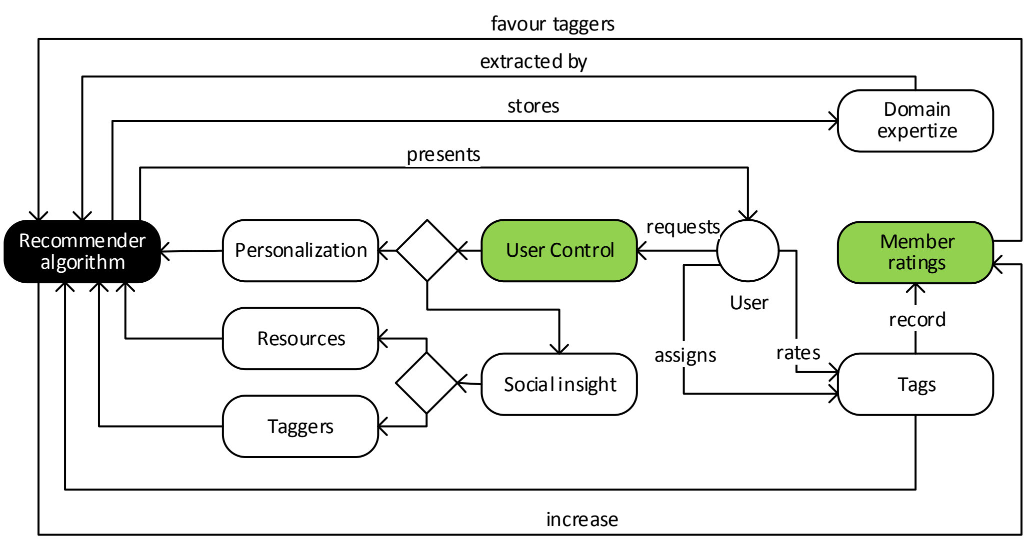


Figure VII-4

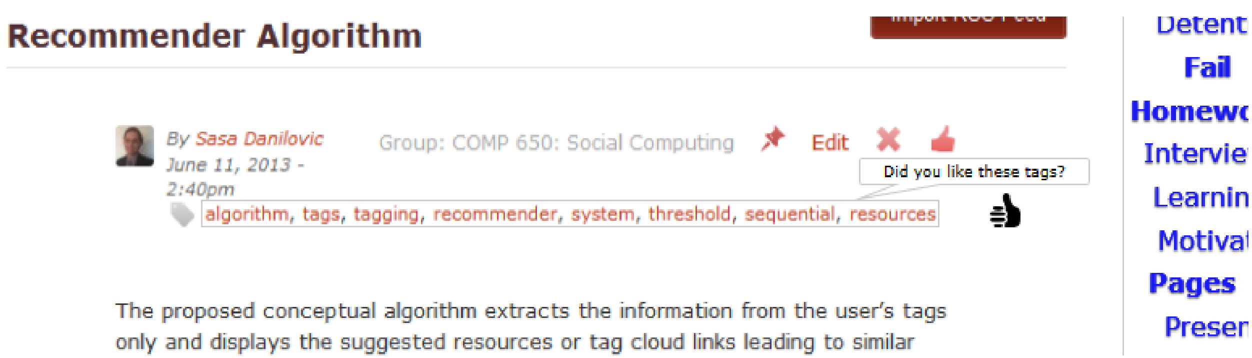
Harvesting expertize, user perspective



Naturally, the recommender algorithm's design would to account for the resources with higher rated tags, and display them on top of the list. The "dislike" button could cause unwanted effects in socially wider domains (stupidity of mobs), therefore omitted by design. Diverting user attention to tags can reduce spam, increase motivation to better describe a resource, along with the quality of descriptions across a wide range of users. The rated tag sets are applied against personal tag cloud, reflecting the most successful ones, by using weighing and co-occurrence algorithms. In this scenario, the resource's "thumbs up" (in red) would have to be removed, since it can create confusion with users (Figure VII-5). Considering that providing social insight is tag clouds' primary purpose, rating system has the potential to adhere to crowds' wide preferences, contrary to rather unsuccessful and mechanical semantical extraction. However, this design goes beyond a tag cloud, and the domain's entire philosophy would need to follow, since the resource recommendation relies on taggers and tags, which needs further empirical research.

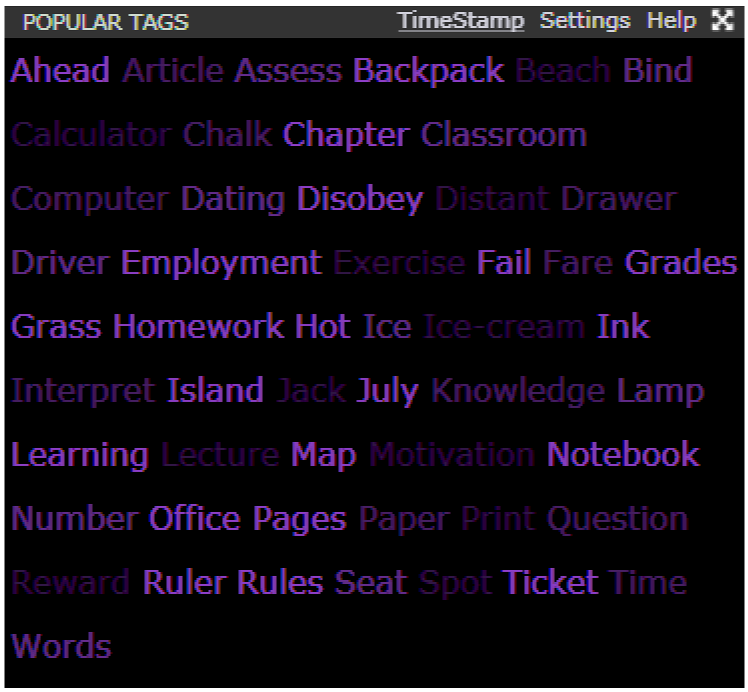
Figure VII-5

Rating the entire tag set



Visualizing tag currency should be discreet and triggered only on demand, especially since it can compete with other cues. This not only creates confusion with novice users, but it had been proven to reduce user acceptance. An idea was to have a button on a title bar that can visualize the age of tags by press-and-hold mouse action. This would allow the function to be temporary and serve only as an added pointer in search, while not competing with other cues or producing confusion (Figure VII-6).

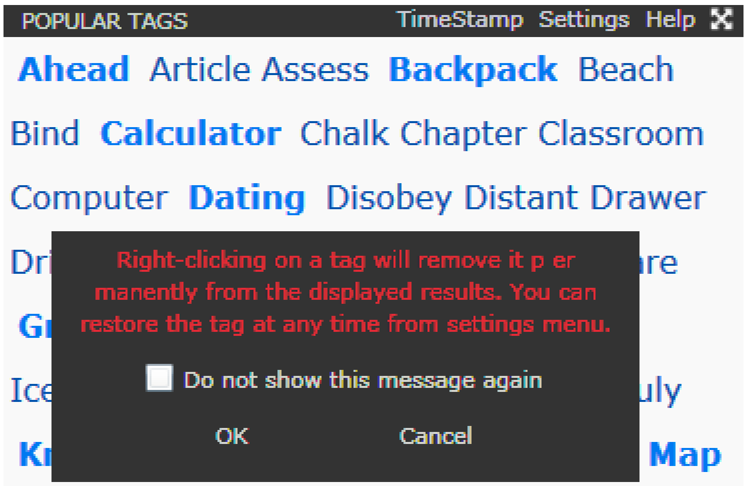
Figure VII-6
Timestamp



Although faded tags are not suitable for cueing, the goal was to distinguish tag age without using charts or diagrams, since this visualisation can only temporarily affect aesthetics, and the emphasis of newer or older tags depends on the domain purpose. For example, in stagnant domains newer tags can convey current popular topics, while in dynamic attest for stable ones. The title bar transparently notifies user of the selected option, and the background colour changed to increase the contrast. Although the user must remember the desired tag, which increases the cognitive effort, this feature is nonessential and even better left out if in danger of competing with cues that are more important.

Figure VII-7

Tag filtering



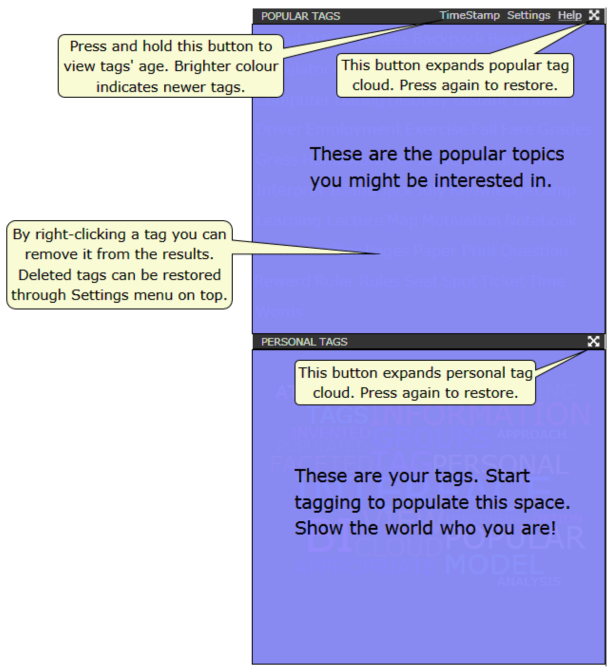
Tag filtering was resolved by using click-and-hold mouse motion, and offering a small menu dedicated to tag removal (Figure VII-7). As a precaution, another dialog notifies user of the action and the outcomes, considering that action is easily accessible. It is also possible to restore the deleted tag through user menu, to distinguish tag deletion function from tag filtering. The help function is accessible from the title bar using press-and-hold mouse action and transparently describes the user interface. Hiding tags during help avoids visual clutter, however restored on releasing the mouse button (Figure VII-8).

Control (Administrative and User Control)

The first step was to design administrative controls for managing tag cloud, but also with ability of restricting or delegating features to user (Figure VII-9). The administrative panel design was following two guidelines: apart from controlling the scope of available user options, the administrator needs an easy understandable set of controls, since tag clouds are only a domain's supporting feature. Since significant administrative overhead would need a deeper knowledge of tag clouds' operation, designing these controls had a typical webmaster in focus, needing minimum guidance.

Figure VII-8

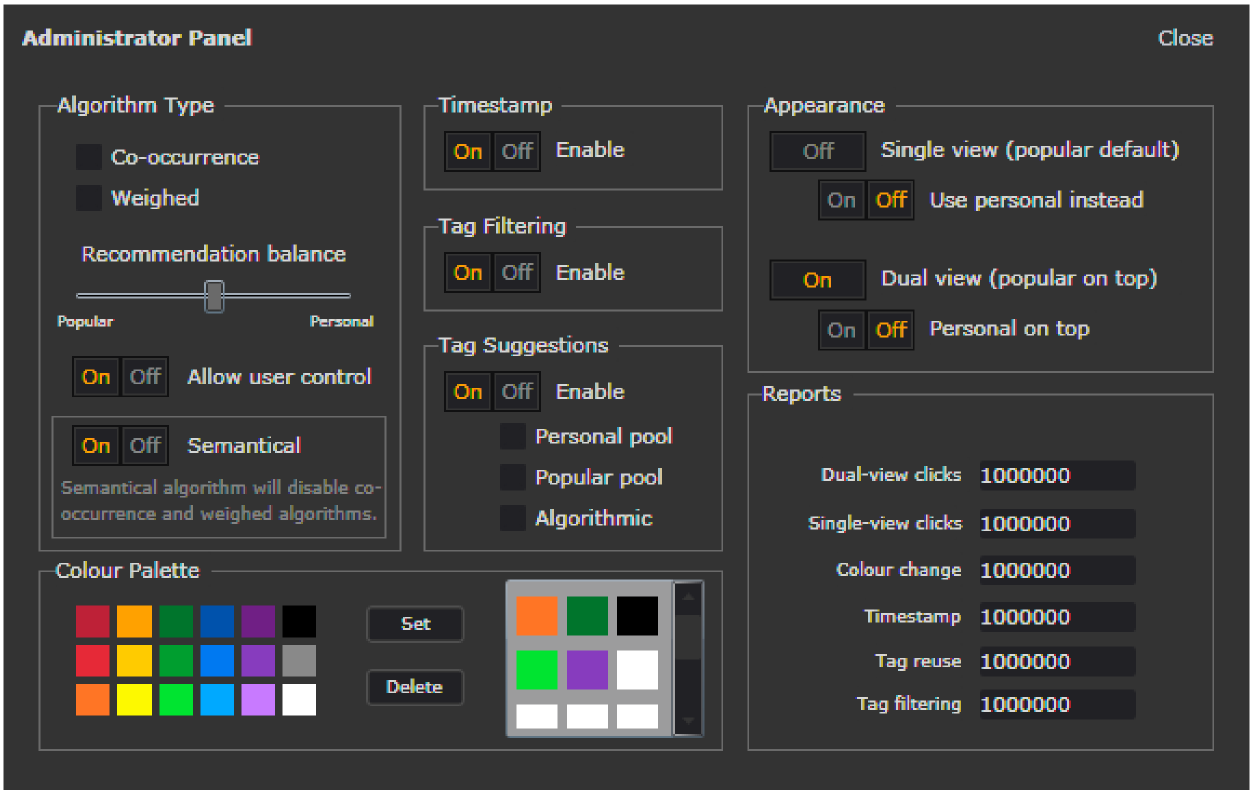
Transparency (help function)



The algorithm choice stems from the preliminary study's concept, where combining weighing and co-occurrences ones is the principle for recommendation. Occasionally, weighing algorithm can provide a cleaner insight into popular tags, while co-occurrences provide a better description of resources. This is optional depending on domain's purpose. Although semantical algorithms are not successful in practice, it is likely the theory will advance to the point where such inclusion is possible. The slider provides control over the influence of personal and popular tags, another recommendation concept from the preliminary study.

Figure VII-9

Administrator panel

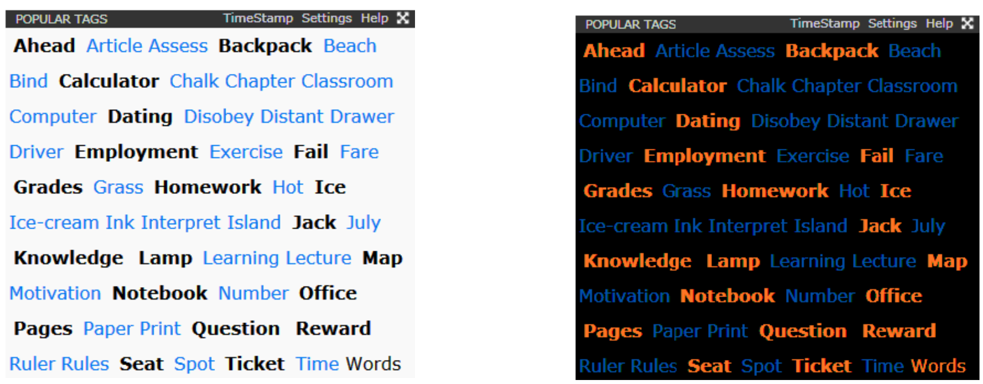


The middle of the interface logically grouped for motivational features, except for the tag rating proposal, which cannot be switched on or off since it is embedded into the entire system. By checking-marking tag timestamps and filtering, these controls are passed on to the user interface. In instances of start-up domains, tag filtering could cause sparse tag clouds, and timestamp would not be able to visualize noticeable differentiation, therefore these options best suit mature domains.

Tag suggestions are possible to set from personal, popular, or hybridised pool of tags, with latter being useful in wider social domains where using popular tags can produce wrongly described resources. By limiting the number of popular tags, can reduce the problems with overtagging and spamming. The appearance is resolved in two ways: by deciding whether the users can have single or dual views (based on domain maturity), and the colour patterns. If allowing users broad control over colouration, the result could avert from the domain theme.

Figure VII-10

Colour pattern change



The analysis results also proved tag colour to have a significant role on navigational cueing and unsuitable colour combinations could lead to unintentional self-induced detriment to user satisfaction. Therefore, the administrator controls fitting colour combinations, even leaving room for consulting a professional designer to increase the attractiveness. User control uses stack panel, therefore mitigating any changes to user interface dynamically, leaving no empty fields for the disabled option (Figure VII-11). Avoiding grayed-out sections encourages fluent workflow and reduces any potential confusion (Figure VII-11). The administrator has the control over four user-controllable features in user menu: colour patterns, recommendation type, restoration of deleted tags, and semantical or alphabetical sorting. Since the colour do not change in real-time, the right side of visualization section shows a sample of the resulting colour pattern appearance.

Figure VII-11

User settings window

Colour pattern selection was at first placed on the title-bar, however it was moved inside the user menu, since it can motivate users' familiarity with other options. The change in background colour of a popular view automatically changes the background of the personalized one (Figure VII-10). The tactics is suitable if menu is richer in user-selectable functionalities as it flattens the learning curve of the advanced ones.

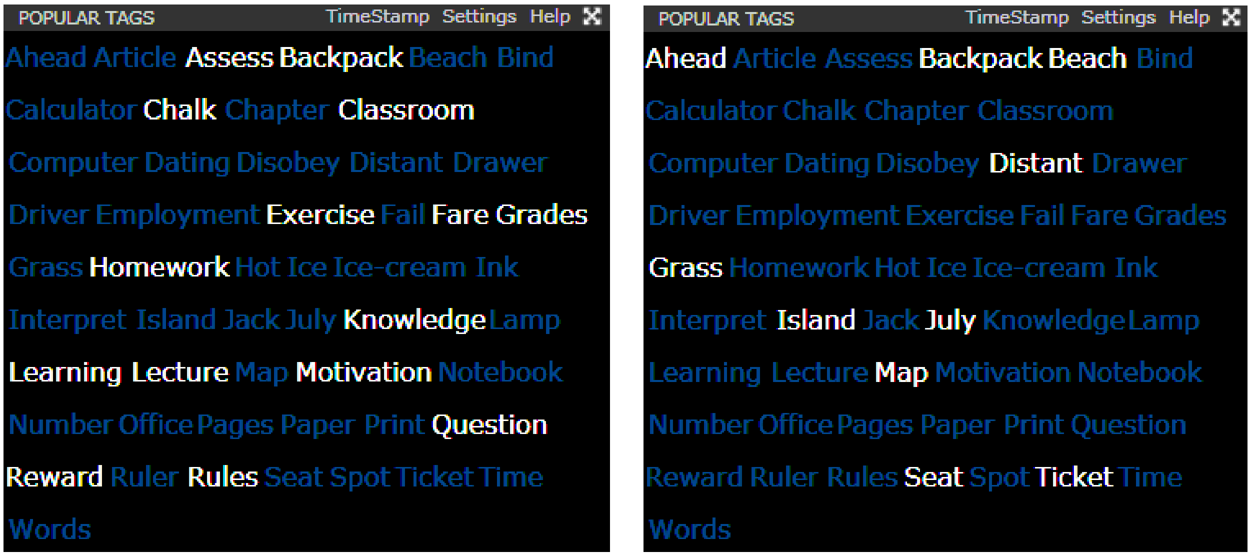
The recommendation type only affects the recommender algorithm in small proportion, and does not give the user a complete control over it. This is driven by dual-view tag cloud's existing popular and personalized tags views. By moving slider towards popular side, favouring those tags will in the popular view, however always with a blend of personal tags. For easier understanding, label "*You know what's best for me*" substitutes "*Personalized*" one. Tag undelete is a simple feature listing tags that user deleted, and by selecting a tag and Restore button, again including the tag in search results.

Recommend (Recommender Algorithm)

Although this software artefact does not have a recommender algorithm, for most part it is recommender-ready. The concept for the recommender algorithm from preliminary study however required visualization beyond alphabetical sorting or personal view. By accommodating this need, the visual interface became even more versatile in visualizing tag relations (Figure VII-12). When user uses mouse to hover over a tag, the weighing and co-occurrence algorithms activate and highlight the correlated tags. This logic is not limiting to those algorithm types, but extends to semantical ones as well, visualizing the relevant relations. This also eliminates the need for tag sorting or changing views, since the ordering can remain alphabetical and still have the power to indicate relevant tags. Although tag contrast is detrimental to aesthetics, used in this dynamical manner can transform into a useful navigational cue. In other words, using a cue not well received by users to signal the tags not supposed to be selected. The importance of this visualization becomes more emphasized in cloud layout, with the potential of introducing a level of organization.

Figure VII-12

Tag importance on mouse-hover



Evaluate (Quality Assessment)

This interface had some drawbacks from logical perspective: it included neither clustered nor list layouts. Since clustered layouts are space-demanding, it averted from discreet role of a tag cloud, secondly, the research efforts must currently focus on discovering the effects of simpler navigational cues. The literature analysis showed that with increase in complexity, the cognitive effort grows, reducing user acceptance (pp. 73-74, 81, 83, 92-93, 95-97, 99). Once establishing the foundations, it is possible to move forward to visualizations that are more comfortable. Although efficient, the list layouts were not included because they perform the best for visualizing semantical and ontology relationships, which the planned algorithm does not support. Therefore, it would be pointless to visualize a layout simply because it is possible. These layouts also carry high visual cueing similarity to sequential ones, and if there is a later need, the same concepts apply. The S.E.C.U.R.E. method works well in discovering the correct visualizations, sets of features and recommending type, and significantly lowers the design overhead. Starting from the visualization point and user acceptance, the method revealed the possible algorithms that can be employed, and in what capacity. For example, these layouts are not suitable for ontologies, therefore not considering that algorithm. This however does not mean fulfilling ontologies in tag clouds is impossible, only that efforts should focus on finding a solution within user acceptance boundaries, which start with visualization.

The goal of this artefact is continuous experimentation and assessment, by integrating the recommender algorithm and monitor its performance. The administrator panel offers some means of performing this task, by counting the number of clicks a certain feature got. This is not a definitive approach to settling user preferences theory, but one step closer to understanding it. Once the recommender module is built, and the whole system deployed, it will be possible to evaluate features' use levels coupled with user surveys to recognize both the flaws and what solutions did work.

Conclusions and Further Research

By analyzing findings of hundred publications, this study attempted at providing a better understanding of visual navigation cues commonly set in in different variants of tag clouds. The result was a segregation of the relevant visual factors, and the most common layouts, since relative significance of cues is dependent on layout variants. The interrelation properties of visual elements can either increase or decrease their influence on selection of a particular navigation path. For example, tag size undoubtedly has the highest impact on user selecting a tag, however other combinations can be as equally strong, such as tag colour and weight. However, if combining all three, the role of differently marked tags becomes questionable, reducing the chances of their selection significantly. On the other side of the spectrum, certain cues have a low user acceptance if continuously presented, for example, using faded, various styles or italic fonts. In other words, employing any visualization type contributing to the reduced readability has negative connotations and it is a primary reason for user dissatisfaction. The effect of competing cues has the greatest power of creating confusion while selecting a path. By visualizing several roles, such as tag popularity and their semantical relationship (or a timestamp) is likely to be counterproductive, especially to novice users. All of these implications, realized through ten principles of tag cloud design and S.E.C.U.R.E. method presented in this study, aim to provide basic implementation guidelines.

The results aim to aid designers in forming effective visual interfaces able to improve user experience and acceptance, and allowing them to perform the associated tasks more effectively. The ideas and potential solutions presented are not finite, but a subject to further iterative empirical testing and refinement, since tag clouds are an integral part of social systems that need continuous user approval.

The theoretical model in its current state provides a solid starting point for subsequent theory building, especially in the insufficiently supported areas, distinguishing well saturated topics from the ones that are not. By adding more evidence, the likelihood of lowering its high abstraction level increases, and in the future potentially serve as a practically applicable design framework. The model’s flexibility allows for examining different topical relations, even the ones not related in its present state, or not researched yet.

The future work will address two major areas: in the first stage, expanding the user interface presented in this study with the administrative and visual choices, to provide a tool for empirical testing. By including multiple tag cloud layouts, colourations, orderings and sorting alternatives, it will be possible to observe and survey the users in discovering the most effective and used option sets. This attempt strives to provide the researchers with the necessary focus, as opposed to continuous inventing and reinventing. The user interface will also be developed to the single “entry point”, accepting and visualizing the data provided by the recommender algorithm without making any compromises. In other words, any recommender algorithms’ results will have to conform to the rules of the user-accepted visualization and not conversely. Once we develop a deeper understanding of users’ browsing habits, it is possible to discuss guidance down a longer path, fulfilling the intended role of a tag cloud. The second stage will entitle developing the recommender algorithm, from the idea emerged in the preliminary study, resulting in a fully functional tag cloud.

REFERENCES

Allam, H., Blustein, J., Bliemel, M., & Spiteri, L. (2012). Exploring Factors Impacting Users' Attitude and Intention towards Social Tagging Systems. In *2012 45th Hawaii International Conference on System Sciences* (pp. 3129–3138). IEEE. doi:10.1109/HICSS.2012.267

Ames, M., & Naaman, M. (2007). Why we tag. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07* (p. 971). New York, New York, USA: ACM Press. doi:10.1145/1240624.1240772

Au Yeung, C., & Iwata, T. (2010). Capturing implicit user influence in online social sharing. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia - HT '10* (p. 245). New York, New York, USA: ACM Press. doi:10.1145/1810617.1810662

Axelsson, K., & Goldkuhl, G. (n.d.). *Theory Modelling - Action Focus when Building a Multi-Grounded Theory*. Linköping.

Bateman, S., Gutwin, C., & Nacenta, M. (2008). Seeing things in the clouds: the effect of visual features on tag cloud selections. *Proceedings of the Nineteenth ACM Conference on Hypertext and Hypermedia*. Pittsburgh, PA, USA: ACM. doi:10.1145/1379092.1379130

Bateman, S., Muller, M. J., & Freyne, J. (2009). Personalized retrieval in social bookmarking. In *Proceedinfs of the ACM 2009 international conference on Supporting group work - GROUP '09* (p. 91). New York, New York, USA: ACM Press. doi:10.1145/1531674.1531688

Berlocher, I., Lee, K., & Kim, K. (2008). TopicRank. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '08* (p. 703). New York, New York, USA: ACM Press. doi:10.1145/1390334.1390460

Borenstein, M., Hedges, L. V, Higgins, J. P. T., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. John Wiley & Sons.

Candan, K. S., Di Caro, L., & Sapino, M. L. (2008). Creating tag hierarchies for effective navigation in social media. In *Proceeding of the 2008 ACM workshop on Search in social media - SSM '08* (p. 75). New York, New York, USA: ACM Press. doi:10.1145/1458583.1458597

Carpendale, S., Dörk, M., & Williamson, C. (2012). Navigating tomorrow's web. *ACM Transactions on the Web*, *6*(3), 1–28. doi:10.1145/2344416.2344420

Chen, Y.-X., Santamaría, R., Butz, A., & Therón, R. (2009). *TagClusters: Semantic aggregation of collaborative tags beyond TagClouds*. (A. Butz, B. Fisher, M. Christie, A. Krüger, P. Olivier, & R. Therón, Eds.) (Vol. 5531). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-02115-2

Chi, E. H., & Mytkowicz, T. (2007). Understanding navigability of social tagging systems. In *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia - HT '08* (p. 10). New York, New York, USA: ACM Press. doi:10.1145/1379092.1379110

Chiarella, A. (2011). How do readers respond to social text signals? In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (Vol. 2011, pp. 2331–2336). Retrieved from /p/39077/

Chiarella, A., & Chmiliar, L. (2012). Reading Together: Indirect Collaboration through a Social Software Application. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (Vol. 2012, pp. 1772–1776). Retrieved from /p/41865/

Chiarella, A. F. (2012). The influence of self-organized, social text signals on readers' behaviour collective intelligence or unchecked imitation. *Annual Meeting of the American Educational Research Association, Vancouver, British Columbia, Canada*.

Chiarella, A. F., & Lajoie, S. P. (n.d.). Collaboratively signalling text with CoREAD. *McGill University, Montreal, Canada*.

Chiarella, A., & Lajoie, S. (2010). Social Annotation: Emergent Text Signals through Self-Organization. In *EdMedia: World Conference on Educational Media and Technology* (Vol. 2010, pp. 1456–1462). Retrieved from /p/34828/

Chiarella, A. F. & Lajoie, S. P. (n.d.). Dynamically Modifying Text Signals: A Self-Organising Systems Approach to Collaboration. *Annual Meeting of the American Educational Research Association, San Diego, CA.*

Christie, T., Lueg, C., & Baghaei, N. (2010). Tag clouds as social signallers. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction - OZCHI '10* (p. 360). New York, New York, USA: ACM Press. doi:10.1145/1952222.1952302

Christie, T., Lueg, C., & Baghaei, N. (2011). PETAC. In *Proceedings of the 12th Annual Conference of the New Zealand Chapter of the ACM Special Interest Group on Computer-Human Interaction - CHINZ '11* (pp. 73–80). New York, New York, USA: ACM Press. doi:10.1145/2000756.2000766

Chuang, J., Manning, C. D., & Heer, J. (2012). "Without the clutter of unimportant words." *ACM Transactions on Computer-Human Interaction*, *19*(3), 1–29. doi:10.1145/2362364.2362367

Crawford, G. . (1987). The geometric mean procedure for estimating the scale of a judgement matrix. *Mathematical Modelling*, *9*(3-5), 327–334. doi:10.1016/0270-0255(87)90489-1

Cress, U., Held, C., & Kimmerle, J. (2013). The collective knowledge of social tags: Direct and indirect influences on navigation, learning, and information processing. *Computers & Education*, *60*(1), 59–73. doi:10.1016/j.compedu.2012.06.015

Danilovic, S. (2013). ICARUS – Adaptive Tag Cloud Navigational and Recommender System. Unpublished manuscript, School of Computing and Information Systems, Athabasca University, Athabasca, Canada.

De Vita, E., Marcialis, I., Marotto, V., & Paddeu, G. (2011). Pag cloud. In *Proceedings of the 9th International Conference on Advances in Mobile Computing and Multimedia - MoMM '11* (p. 231). New York, New York, USA: ACM Press. doi:10.1145/2095697.2095743

Derntl, M., Hampel, T., Motschnig-Pitrik, R., & Pitner, T. (2011). Inclusive social tagging and its support in Web 2.0 services. *Computers in Human Behavior*, *27*(4), 1460–1466. doi:10.1016/j.chb.2010.09.014

Deutsch, S., Schrammel, J., & Tscheligi, M. (2011). Comparing Different Layouts of Tag Clouds: Findings on Visual Perception. In A. Ebert, A. Dix, N. Gershon, & M. Pohl (Eds.), *Human Aspects of Visualization SE - 3* (Vol. 6431, pp. 23–37). Springer Berlin Heidelberg. doi:10.1007/978-3-642-19641-6_3

Di Caro, L., Selçuk Candan, K., & Luisa Sapino, M. (2011). Navigating within news collections using tag-flakes. *Journal of Visual Languages & Computing*, *22*(2), 120–139. doi:10.1016/j.jvlc.2010.11.001

Diaz, J., Hu, K., & Tory, M. (2009). An exploratory study of tag-based visual interfaces for searching folksonomies, 410–417. Retrieved from http://0-dl.acm.org.aupac.lib.athabascau.ca/citation.cfm?id=1671011.1671064

Dörk, M., Feng, P., Collins, C., & Carpendale, S. (2013). Critical InfoVis. In *CHI ’13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13* (p. 2189). New York, New York, USA: ACM Press. doi:10.1145/2468356.2468739

Dron, J. (2005). Discovering the complex effects of navigation cues in an e-learning environment. In *Proceedings of E-Learn 2005* (pp. 2026–2033). Vancouver.

Dron, J. (2008a). Designing the Undesignable: Social software and control. *The International Review of Research in Open and Distance Learning; Vol 9, No 1 (2008): Special Issue - Role of Distance Learning in the Right to Education*. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/532/1006

Dron, J. (2008b). The trouble with tags: an approach to richer tagging for online learning. In C. J. Bonk, M. M. Lee, & T. Reynolds (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2008* (pp. 3660–3667). Las Vegas, Nevada, USA: AACE. Retrieved from http://www.editlib.org/p/30195

Dron, J. (2014). Social Navigation for Learning in Immersive Worlds. In *Synthetic Worlds* (pp. 119–136). Springer.

Dron, J., Mitchell, R., Siviter, P., & Boyne, C. (2000). CoFIND—An experiment in N-dimensional collaborative filtering. *Journal of Network and Computer Applications*, *23*(2), 131–142. doi:10.1006/jnca.2000.0105

Dumais, J., & Foy, P. (1996). *Calculation of Sampling Weights*. Retrieved from http://timssandpirls.bc.edu/timss1995i/TIMSSPDF/TR2chap4.pdf

Emerson, J., Churcher, N., & Deaker, C. (2013). From Toy to Tool: Extending Tag Clouds for Software and Information Visualisation. In *2013 22nd Australian Software Engineering Conference* (pp. 155–164). IEEE. doi:10.1109/ASWEC.2013.27

Farooq, U., Kannampallil, T. G., Song, Y., Ganoe, C. H., Carroll, J. M., & Giles, L. (2007). Evaluating tagging behavior in social bookmarking systems. In *Proceedings of the 2007 international ACM conference on Conference on supporting group work - GROUP ’07* (p.

351). New York, New York, USA: ACM Press. doi:10.1145/1316624.1316677

Feng, Y. (2004). Note on "sums of independent fuzzy random variables." *Fuzzy Sets and Systems*, *143*(3), 479–485. doi:10.1016/S0165-0114(03)00137-4

Filho, F. F., de Albuquerque, J. P., Resende, A., de Geus, P. L., & Olson, G. M. (2009). A visualization interface for interactive search refinement.

Filho, F. F., Olson, G. M., & de Geus, P. L. (2010). Kolline. In *Proceedings of the 28th ACM International Conference on Design of Communication - SIGDOC '10* (p. 89). New York, New York, USA: ACM Press. doi:10.1145/1878450.1878465

Forlines, C., & Balakrishnan, R. (2009). Improving visual search with image segmentation. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09* (p. 1093). New York, New York, USA: ACM Press. doi:10.1145/1518701.1518868

Fu, W.-T., Kannampallil, T. G., & Kang, R. (2010). Facilitating exploratory search by model-based navigational cues. In *Proceedings of the 15th international conference on Intelligent user interfaces - IUI '10* (p. 199). New York, New York, USA: ACM Press. doi:10.1145/1719970.1719998

Fu, W.-T., Kannampallil, T., Kang, R., & He, J. (2010). Semantic imitation in social tagging. *ACM Transactions on Computer-Human Interaction*, *17*(3), 1–37. doi:10.1145/1806923.1806926

Fujimura, K., Fujimura, S., Matsubayashi, T., Yamada, T., & Okuda, H. (2008). Topigraphy: Visualization for Large-scale Tag Clouds. In *Proceeding of the 17th international conference on World Wide Web - WWW '08* (p. 1087). New York, New York, USA: ACM Press. doi:10.1145/1367497.1367669

Gambette, P., & Véronis, J. (2010a). Visualising a Text with a Tree Cloud. In H. Locarek-Junge & C. Weihs (Eds.), *Classification as a Tool for Research SE - 61* (pp. 561–569). Springer Berlin Heidelberg. doi:10.1007/978-3-642-10745-0_61

Gambette, P., & Véronis, J. (2010b). *Visualizing text with a tree cloud*. (H. Locarek-Junge & C. Weihs, Eds.)*IFCS'09: International Federation of Classification Societies Conference*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-10745-0

Gedikli, F., Ge, M., & Jannach, D. (2011). Understanding Recommendations by Reading the Clouds. In C. Huemer & T. Setzer (Eds.), *E-Commerce and Web Technologies SE - 17* (Vol. 85, pp. 196–208). Springer Berlin Heidelberg. doi:10.1007/978-3-642-23014-1_17

Gedikli, F., & Jannach, D. (2013). Improving recommendation accuracy based on item-specific tag preferences. *ACM Transactions on Intelligent Systems and Technology*, *4*(1), 1–19. doi:10.1145/2414425.2414436

Gedikli, F., Jannach, D., & Ge, M. (2014). How should I explain? A comparison of different explanation types for recommender systems. *International Journal of Human-Computer Studies*, *72*(4), 367–382. doi:10.1016/j.ijhcs.2013.12.007

Gibbs. (2011). Core Elements Part 1. Grounded Theory. Retrieved from http://www.youtube.com/watch?v=4SZDTp3_New

Gou, L., Zhang, S., Wang, J., & Zhang, X. (Luke). (2010). TagNetLens. In *Proceedings of the 3rd International Symposium on Visual Information Communication - VINCE '10* (p. 1). New York, New York, USA: ACM Press. doi:10.1145/1865841.1865865

Gupta, M., Li, R., Yin, Z., & Han, J. (2010). Survey on social tagging techniques. *ACM SIGKDD Explorations Newsletter*, *12*(1), 58. doi:10.1145/1882471.1882480

Gwizdka, J. (2010). Of kings, traffic signs and flowers. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia - HT '10* (p. 167). New York, New York, USA: ACM Press. doi:10.1145/1810617.1810646

Gwizdka, J., & Cole, M. (2013). Does interactive search results overview help? In *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13* (p. 1869). New York, New York, USA: ACM Press. doi:10.1145/2468356.2468691

Halvey, M. J., & Keane, M. T. (2007). An assessment of tag presentation techniques. In *Proceedings of the 16th international conference on World Wide Web* (pp. 1313–1314). Banff, Alberta, Canada: ACM. doi:10.1145/1242572.1242826

Han, X., Shen, Z., Miao, C., & Luo, X. (2010). Folksonomy-based ontological user interest profile modeling and its application in personalized search, 34–46. Retrieved from http://0-dl.acm.org.aupac.lib.athabascau.ca/citation.cfm?id=1886192.1886201

Hearst, M. A., & Rosner, D. (2008). Tag Clouds: Data Analysis Tool or Social Signaller? *Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences*. IEEE Computer Society. doi:10.1109/hicss.2008.422

Heckner, M., Neubauer, T., & Wolff, C. (2008). Tree, funny, to_read, google. In *Proceeding of the 2008 ACM workshop on Search in social media - SSM '08* (p. 3). New York, New York, USA: ACM Press. doi:10.1145/1458583.1458589

Held, C., Kimmerle, J., & Cress, U. (2012). Learning by foraging: The impact of individual knowledge and social tags on web navigation processes. *Computers in Human Behavior*, *28*(1), 34–40. doi:10.1016/j.chb.2011.08.008

Helic, D., Körner, C., Granitzer, M., Strohmaier, M., & Trattner, C. (2012). Navigational efficiency of broad vs. narrow folksonomies. In *Proceedings of the 23rd ACM conference on Hypertext and social media - HT '12* (p. 63). New York, New York, USA: ACM Press. doi:10.1145/2309996.2310008

Helic, D., & Strohmaier, M. (2011). Building directories for social tagging systems. In *Proceedings of the 20th ACM international conference on Information and knowledge management - CIKM '11* (p. 525). New York, New York, USA: ACM Press. doi:10.1145/2063576.2063655

Helic, D., Trattner, C., Strohmaier, M., & Andrews, K. (2010a). On the Navigability of Social Tagging Systems. In *2010 IEEE Second International Conference on Social Computing* (pp. 161–168). IEEE. doi:10.1109/SocialCom.2010.31

Helic, D., Trattner, C., Strohmaier, M., & Andrews, K. (2010b). On the Navigability of Social Tagging Systems. In *2010 IEEE Second International Conference on Social Computing* (pp.

161–168). IEEE. doi:10.1109/SocialCom.2010.31

Herlocker, J. L., Konstan, J. A., & Riedl, J. (2000). Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work - CSCW '00* (pp. 241–250). New York, New York, USA: ACM Press. doi:10.1145/358916.358995

Hong, L., Chi, E. H., Budiu, R., Pirolli, P., & Nelson, L. (2008). SparTag.us. In *Proceedings of the working conference on Advanced visual interfaces - AVI '08* (p. 65). New York, New York, USA: ACM Press. doi:10.1145/1385569.1385582

Hotho, A., Jäschke, R., Schmitz, C., & Stumme, G. (2006). *Information retrieval in folksonomies: Search and ranking*. (Y. Sure & J. Domingue, Eds.) (Vol. 4011). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/11762256

Jackson, H. (1997). User Interface Design - How Can It Be Improved? *Imperial College London*, *2*. Retrieved from http://www.doc.ic.ac.uk/~nd/surprise_97/journal/vol2/hafj/

Kang, R., & Fu, W.-T. (2010). Exploratory information search by domain experts and novices. In *Proceedings of the 15th international conference on Intelligent user interfaces - IUI '10* (p. 329). New York, New York, USA: ACM Press. doi:10.1145/1719970.1720023

Kaser, O., & Lemire, D. (2007). Tag-Cloud Drawing: Algorithms for Cloud Visualization. Data Structures and Algorithms. Retrieved from http://arxiv.org/abs/cs/0703109

Keele, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Technical report, EBSE Technical Report EBSE-2007-01. Retrieved from https://www.cs.auckland.ac.nz/~norsaremah/2007 Guidelines for performing SLR in SE v2.3.pdf

Khan, K. S., Kunz, R., Kleijnen, J., & Antes, G. (2003). Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine*, *96*(3), 118–21. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=539417&tool=pmcentrez&rendertype=abstract

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University*, *33*(2004), 1–26.

Klaisubun, P., Kajondecha, P., & Ishikawa, T. (2007). Behavior Patterns of Information Discovery in Social Bookmarking Service. In *IEEE/WIC/ACM International Conference on Web Intelligence (WI'07)* (pp. 784–787). IEEE. doi:10.1109/WI.2007.146

Körner, C., Benz, D., Hotho, A., Strohmaier, M., & Stumme, G. (2010). Stop thinking, start tagging. In *Proceedings of the 19th international conference on World wide web - WWW '10* (p. 521). New York, New York, USA: ACM Press. doi:10.1145/1772690.1772744

Kuo, B. Y.-L., Hentrich, T., Good, B. M. ., & Wilkinson, M. D. (2007). Tag clouds for summarizing web search results. In *Proceedings of the 16th international conference on World Wide Web - WWW '07* (p. 1203). New York, New York, USA: ACM Press. doi:10.1145/1242572.1242766

Lee, B., Riche, N. H., Karlson, A. K., & Carpendale, S. (2010). SparkClouds: visualizing trends in tag clouds. *IEEE Transactions on Visualization and Computer Graphics*, *16*(6), 1182–9. doi:10.1109/TVCG.2010.194

Lei, P.-W., & Wu, Q. (2007). Introduction to Structural Equation Modeling: Issues and Practical Considerations. *Educational Measurement: Issues and Practice*, *26*(3), 33–43. doi:10.1111/j.1745-3992.2007.00099.x

Liang, H., Xu, Y., Li, Y., Nayak, R., & Weng, L.-T. (2009). Personalized Recommender Systems Integrating Social Tags and Item Taxonomy. In *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology* (Vol. 1, pp. 540–547). IEEE. doi:10.1109/WI-IAT.2009.89

Lin, C.-S., & Chen, Y.-F. (2012). Examining social tagging behaviour and the construction of an online folksonomy from the perspectives of cultural capital and social capital. *Journal of Information Science*, *38*(6), 540–557. doi:10.1177/0165551512459826

Liu, K., Fang, B., & Zhang, W. (2010). Ontology emergence from folksonomies. In *Proceedings of the 19th ACM international conference on Information and knowledge management - CIKM '10* (p. 1109). New York, New York, USA: ACM Press. doi:10.1145/1871437.1871578

Lohmann, S., Heim, P., Tetzlaff, L., Ertl, T., & Ziegler, J. (2009). Exploring Relationships between Annotated Images with the ChainGraph Visualization. In T.-S. Chua, Y. Kompatsiaris, B. Mérialdo, W. Haas, G. Thallinger, & W. Bailer (Eds.), *Semantic Multimedia SE - 4* (Vol. 5887, pp. 16–27). Springer Berlin Heidelberg. doi:10.1007/978-3-642-10543-2_4

Lohmann, S., Ziegler, J., & Tetzlaff, L. (2009). Comparison of Tag Cloud Layouts: Task-Related Performance and Visual Exploration. *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part I*. Uppsala, Sweden: Springer-Verlag. doi:10.1007/978-3-642-03655-2_43

Manzato, M. G., & Goularte, R. (2012). A multimedia recommender system based on enriched user profiles. In *Proceedings of the 27th Annual ACM Symposium on Applied Computing - SAC '12* (p. 975). New York, New York, USA: ACM Press. doi:10.1145/2245276.2245466

Melenhorst, M., Grootveld, M., van Setten, M., & Veenstra, M. (2008). Tag-based information retrieval of video content. In *Proceeding of the 1st international conference on Designing interactive user experiences for TV and video - uxtv '08* (p. 31). New York, New York, USA: ACM Press. doi:10.1145/1453805.1453813

Meline, T. (2006). Selecting studies for systematic review: Inclusion and exclusion criteria. *Contemporary Issues in Communication Science and …*. Retrieved from http://www.asha.org/uploadedFiles/asha/publications/cicsd/2006SSelectingStudiesforSystematicReview.pdf

Mezghani, M., Zayani, C. A., Amous, I., & Gargouri, F. (2012). A user profile modelling using social annotations. In *Proceedings of the 21st international conference companion on World Wide Web - WWW '12 Companion* (p. 969). New York, New York, USA: ACM Press. doi:10.1145/2187980.2188230

Milicevic, A. K., Nanopoulos, A., & Ivanovic, M. (2010). Social tagging in recommender systems: a survey of the state-of-the-art and possible extensions. *Artificial Intelligence Review*, *33*(3), 187–209. doi:10.1007/s10462-009-9153-2

Müller, K. F., Briel, M., D'Amario, A., Kleijnen, J., Marusic, A., Wager, E., … Bassler, D. (2013). Defining publication bias: protocol for a systematic review of highly cited articles and proposal for a new framework. *Systematic Reviews*, *2*(1), 34. doi:10.1186/2046-4053-2-34

Nielsen, J. (1993). *Usability Engineering*. Morgan Kaufmann. Retrieved from http://books.google.ca/books?id=95As2OF67f0C

Nielsen, J. (1994). Usability inspection methods. In *Conference companion on Human factors in computing systems - CHI '94* (pp. 413–414). New York, New York, USA: ACM Press. doi:10.1145/259963.260531

Oelke, D., & Gurevych, I. (2014). A study on human-generated tag structures to inform tag cloud layout. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces - AVI '14* (pp. 297–304). New York, New York, USA: ACM Press. doi:10.1145/2598153.2598155

Oosterman, J., & Cockburn, A. (2010). An empirical comparison of tag clouds and tables. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction - OZCHI '10* (pp. 288–295). Brisbane, Australia: ACM Press. doi:10.1145/1952222.1952284

Panke, S. (1), & Gaiser, B. (2). (2009). "With my head up in the clouds": Using social tagging to organize knowledge. *Journal of Business and Technical Communication*, *23*(3), 318–349. doi:10.1177/1050651909333275

Pérez García-Plaza, A., Zubiaga, A., Fresno, V., & Martínez, R. (2012). Reorganizing clouds: A study on tag clustering and evaluation. *Expert Systems with Applications*, *39*(10), 9483–9493. doi:10.1016/j.eswa.2012.02.108

Ravendran, R., MacColl, I., & Docherty, M. (2012). A comparative usability study of a tag-based interface in the mobile banking context. In *Proceedings of the 13th International Conference of the NZ Chapter of the ACM's Special Interest Group on Human-Computer Interaction - CHINZ '12* (p. 80). New York, New York, USA: ACM Press. doi:10.1145/2379256.2379269

Rivadeneira, A. W., Gruen, D. M., Muller, M. J., & Millen, D. R. (2007). Getting our head in the clouds: toward evaluation studies of tagclouds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 995–998). San Jose, California, USA: ACM. doi:10.1145/1240624.1240775

Sanchez-Zamora, F., & Llamas-Nistal, M. (2009). Visualizing tags as a network of relatedness. In *2009 39th IEEE Frontiers in Education Conference* (pp. 1–6). IEEE. doi:10.1109/FIE.2009.5350530

Schoefegger, K., & Granitzer, M. (2012). Overview and analysis of personal and social tagging context to construct user models. In *Proceedings of the 2nd Workshop on Context-*

awareness in Retrieval and Recommendation - CaRR '12 (pp. 14–21). New York, New York, USA: ACM Press. doi:10.1145/2162102.2162106

Schrammel, J., Deutsch, S., & Tscheligi, M. (2009). *Visual search strategies of tag clouds - results from an eyetracking study*. (T. Gross, J. Gulliksen, P. Kotzé, L. Oestreicher, P. Palanque, R. O. Prates, & M. Winckler, Eds.) (Vol. 5727). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-03658-3

Schrammel, J., Leitner, M., & Tscheligi, M. (2009). Semantically structured tag clouds. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09* (p. 2037). New York, New York, USA: ACM Press. doi:10.1145/1518701.1519010

Seifert, C., Kump, B., Kienreich, W., Granitzer, G., & Granitzer, M. (2008). On the Beauty and Usability of Tag Clouds. In *Information Visualisation, 2008. IV '08. 12th International Conference* (pp. 17–25). IEEE. doi:10.1109/iv.2008.89

Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., … Riedl, J. (2006). tagging, communities, vocabulary, evolution. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work - CSCW '06* (p. 181). New York, New York, USA: ACM Press. doi:10.1145/1180875.1180904

Shapiro, A. F. (2009). Fuzzy random variables. *Insurance: Mathematics and Economics*, *44*(2), 307–314. doi:10.1016/j.insmatheco.2008.05.008

Shepitsen, A., Gemmell, J., Mobasher, B., & Burke, R. (2008). Personalized recommendation in social tagging systems using hierarchical clustering. In *Proceedings of the 2008 ACM conference on Recommender systems - RecSys '08* (p. 259). New York, New York, USA: ACM Press. doi:10.1145/1454008.1454048

Sinclair, J., & Cardew-Hall, M. (2008). The folksonomy tag cloud: when is it useful? *Journal of Information Science*, *34*(1), 15–29. doi:10.1177/0165551506078083

Skoutas, D., & Alrifai, M. (2011). Tag clouds revisited. In *Proceedings of the 20th ACM international conference on Information and knowledge management - CIKM '11* (p. 221). New York, New York, USA: ACM Press. doi:10.1145/2063576.2063613

Song, Y., Qiu, B., & Farooq, U. (2011). Hierarchical tag visualization and application for tag recommendations. In *Proceedings of the 20th ACM international conference on Information and knowledge management - CIKM '11* (p. 1331). New York, New York, USA: ACM Press. doi:10.1145/2063576.2063769

Soulliere, D., Britt, D. W., & Maines, D. R. (2001). CONCEPTUAL MODELING AS A TOOLBOX FOR GROUNDED THEORISTS. *The Sociological Quarterly*, *42*(2), 253–269. doi:10.1111/j.1533-8525.2001.tb00033.x

Statistical Significance. (n.d.). Retrieved from http://help.surveymonkey.com/articles/en_US/kb/Significant-Differences

Strohmaier, M., Körner, C., & Kern, R. (2012). Understanding why users tag: A survey of tagging motivation literature and results from an empirical study. *Web Semantics (Online)*, *17*(C), 1–11. doi:10.1016/j.websem.2012.09.003

Szomszor, M. N., Cantador, I., & Alani, H. (2008). Correlating user profiles from multiple folksonomies. In *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia - HT '08* (p. 33). New York, New York, USA: ACM Press. doi:10.1145/1379092.1379103

Tintarev, N., & Masthoff, J. (2012). Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction*, *22*(4-5), 399–439. doi:10.1007/s11257-011-9117-5

Tomashevskii, I. L. (2014). Geometric mean method for judgement matrices: Formulas for errors. Optimization and Control. Retrieved from http://arxiv.org/abs/1410.0823

Trattner, C. (2011). Improving the navigability of tagging systems with hierarchically constructed resource lists: A comparative study.

Trattner, C., Körner, C., & Helic, D. (2011). Enhancing the navigability of social tagging systems with tag taxonomies. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies - i-KNOW '11* (p. 1). New York, New York, USA: ACM Press. doi:10.1145/2024288.2024310

Treadwell, J. R., Singh, S., Talati, R., McPheeters, M. L., & Reston, J. T. (n.d.). Methods/Approaches. Agency for Healthcare Research and Quality (US). Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK56660/

Ullman, J. B., & Bentler, P. M. (2003). *Structural equation modeling*. Wiley Online Library.

Venetis, P., Koutrika, G., & Garcia-Molina, H. (2011). On the selection of tags for tag clouds. In *Proceedings of the fourth ACM international conference on Web search and data mining - WSDM '11* (p. 835). New York, New York, USA: ACM Press. doi:10.1145/1935826.1935855

Viégas, F. B., & Wattenberg, M. (2008). TIMELINESTag clouds and the case for vernacular visualization. *Interactions*, *15*(4), 49. doi:10.1145/1374489.1374501

Viegas, F. B., Wattenberg, M., & Feinberg, J. (2009). Participatory Visualization with Wordle. *Visualization and Computer Graphics, IEEE Transactions on*, *15*(6), 1137–1144. doi:10.1109/tvcg.2009.171

Vig, J., Sen, S., & Riedl, J. (2008). Tagsplanations. In *Proceedingsc of the 13th international conference on Intelligent user interfaces - IUI '09* (p. 47). New York, New York, USA: ACM Press. doi:10.1145/1502650.1502661

Vig, J., Soukup, M., Sen, S., & Riedl, J. (2010). Tag expression - tagging with feeling. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology - UIST '10* (p. 323). New York, New York, USA: ACM Press. doi:10.1145/1866029.1866079

Waldner, M., Schrammel, J., Klein, M., Kristjánsdóttir, K., Unger, D., & Tscheligi, M. (2013). FacetClouds: exploring tag clouds for multi-dimensional data, 17–24. Retrieved from http://0-dl.acm.org.aupac.lib.athabascau.ca/citation.cfm?id=2532129.2532134

Walker, E., Hernandez, A. V., & Kattan, M. W. (2008). Meta-analysis: Its strengths and limitations. *Cleveland Clinic Journal of Medicine*, *75*(6), 431–439. doi:10.3949/ccjm.75.6.431

Ware, C. (2012). Information Visualization: Perception for Design. Retrieved from http://0-dl.acm.org.aupac.lib.athabascau.ca/citation.cfm?id=2285540

Wash, R., & Rader, E. (2008). Public bookmarks and private benefits: An analysis of incentives in social computing. *Proceedings of the American Society for Information Science and Technology*, *44*(1), 1–13. doi:10.1002/meet.1450440240

Weiwei Cui, Yingcai Wu, Shixia Liu, Furu Wei, Zhou, M. X., & Huamin Qu. (2010). Context-preserving, dynamic word cloud visualization. *IEEE Computer Graphics and Applications*, *30*(6), 42–53. doi:10.1109/MCG.2010.102

Wu, M., & Bowles, C. T. (2010). Principles for Applying Social Navigation to Collaborative Systems. In *Proceedings of the 4th Symposium on Computer Human Interaction for the Management of Information Technology* (pp. 2:1–2:10). New York, NY, USA: ACM. doi:10.1145/1873561.1873563

Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—II. *Information Sciences*, *8*(4), 301–357. doi:10.1016/0020-0255(75)90046-8

Zanardi, V., & Capra, L. (2008a). Social ranking. In *Proceedings of the 2008 ACM conference on Recommender systems - RecSys '08* (p. 51). New York, New York, USA: ACM Press. doi:10.1145/1454008.1454018

Zanardi, V., & Capra, L. (2008b). Social ranking. In *Proceedings of the 2008 ACM conference on Recommender systems - RecSys '08* (p. 51). New York, New York, USA: ACM Press. doi:10.1145/1454008.1454018

Appendix A - LITERATURE USED IN GROUNDED THEORY CODING

Identifier	Title
Ravendran2012	A Comparative Usability Study Of A Tag-Based Interface In
Manzato	A Multimedia Recommender System Based On Enriched User Profiles
Oelke2014	A Study On Human-Generated Tag Structures To Inform Tag Cloud Layout
Mezghani2012	A User Profile Modelling Using Social Annotations
Halvey2007	An Assessment Of Tag Presentation Techniques
Oosterman2010	An Empirical Comparison Of Tag Clouds And Tables
Diaz2009	An Exploratory Study Of Tag-Based Visual Interfaces For Searching Folksonomies
Klaisubun2007	Behavior Patterns Of Information Discovery In Social Bookmarking Service
Helic2012	Building Directories For Social Tagging Systems
Bischoff2008	Can All Tags Be Used For Search?
Yanbe2007	Can Social Bookmarking Enhance Search In The Web?
AuYeung2010	Capturing Implicit User Influence In Online Social Sharing
Dron2000	Cofind—An Experiment In N-Dimensional Collaborative Filtering
Deutsch2011	Comparing Different Layouts Of Tag Clouds: Findings On Visual Perception
LohmannZiegler2009	Comparison Of Tag Cloud Layouts: Task-Related Performance And Visual Exploration
WeiweiCui2010	Context-Preserving, Dynamic Word Cloud Visualization.
Szomzor2008	Correlating User Profiles From Multiple Folksonomies
Candan2008	Creating Tag Hierarchies For Effective Navigation In Social Media
Dork2013	Critical Infovis
Chuang2012	Descriptive Keyphrases For Text Visualization
Dron2005	Discovering The Complex Effects Of Navigation Cues In An E-Learning Environment
Gwizdka2013	Does Interactive Search Results Overview Help?
Chiarella2009	Dynamically Modifying Text Signals: A Self-Organising Systems Approach To Collaboration
Chiarella2006	Enabling The Collective To Assist The Individual: Coread, A Self-Organising Reading Environment

Trattner2011a	Enhancing The Navigability Of Social Tagging Systems With Tag Taxonomies
Golub2005	Entag: Enhancing Social Tagging For Discovery
Farooq2007	Evaluating Tagging Behavior In Social Bookmarking Systems
Tintarev2012	Evaluating The Effectiveness Of Explanations For Recommender Systems
Lin2012	Examining Social Tagging Behaviour And The Construction Of An Online Folksonomy From The Perspectives Of Cultural Capital And Social Capital
Herlocker2000	Explaining Collaborative Filtering Recommendations
Kang2010	Exploratory Information Search By Domain Experts And Novices
Allam2012	Exploring Factors Impacting Users' Attitude And Intention Towards Social Tagging Systems
Lohmann2009	Exploring Relationships Between Annotated Images With The Chaingraph Visualization
Waldner2013	Facetclouds: Exploring Tag Clouds For Multi-Dimensional Data
Fu2010	Facilitating Exploratory Search By Model-Based Navigational Cues
Emerson2013	From Toy To Tool: Extending Tag Clouds For Software And Information Visualisation
Rivadeneira2007	Getting Our Head In The Clouds: Toward Evaluation Studies Of Tagclouds
Song2011	Hierarchical Tag Visualization And Application For Tag Recommendations
Chiarella2011	How Do Readers Respond To Social Text Signals?
Gedikli2014	How Should I Explain? A Comparison Of Different Explanation Types For Recommender Systems
Danilovic2013	ICARUS Adaptive Tag Cloud Navigational And Recommender System
Gedikli2013	Improving Recommendation Accuracy Based On Item-Specific Tag Preferences
Derntl2011	Inclusive Social Tagging And Its Support In Web 2.0 Services
Rader2008	Influences On Tag Choices In Del.Icio.Us
Hotho2006	Information Retrieval In Folksonomies - Search And Ranking
Filho2010	Kolline
Held2012	Learning By Foraging: The Impact Of Individual Knowledge And Social Tags On Web Navigation Processes
Carpendale2012	Navigating Tomorrow's Web
DiCaro2011	Navigating Within News Collections Using Tag-Flakes
Helic2012	Navigational Efficiency Of Broad Vs. Narrow Folksonomies
Korner2010a	Of Categorizers And Describers
Gwizdka2010	Of Kings, Traffic Signs And Flowers: Exploring Navigation Of Tagged

	Documents
Seifert2008	On The Beauty And Usability Of Tag Clouds
Helic2010	On The Navigability Of Social Tagging Systems
Venetis2011	On The Selection Of Tags For Tag Clouds
Liu2010	Ontology Emergence From Folksonomies
Schoefegger2012	Overview And Analysis Of Personal And Social Tagging Context To Construct User Models
DeVita2011	Pag Cloud
Viegas2009	Participatory Visualization With Wordle
Heckner2009	Personal Information Management Vs. Resource Sharing: Towards A Model Of Information Behaviour In Social Tagging Systems
Shepitsen2008	Personalized Recommendation In Social Tagging Systems Using Hierarchical Clustering
Liang2009	Personalized Recommender Systems Integrating Social Tags And Item Taxonomy
Bateman2009	Personalized Retrieval In Social Bookmarking
Christie2011	Petac
Wu2010	Principles For Applying Social Navigation To Collaborative Systems
Wash2008	Public Bookmarks And Private Benefits: An Analysis Of Incentives In Social Computing
Chiarella2012	Reading Together: Indirect Collaboration Through A Social Software Application
Bateman2008	Seeing Things In The Clouds - The Effect Of Visual Features On Tag Cloud Selections
Fu2010a	Semantic Imitation In Social Tagging
Schrammel2009a	Semantically Structured Tag Clouds
Chiarella2010	Social Annotation: Emergent Text Signals Through Self-Organization
Dron2014	Social Navigation For Learning In Immersive Worlds
Zanardi2008	Social Ranking
Milicevic2010	Social Tagging In Recommender Systems: A Survey Of The State-Of-The-Art And Possible Extensions
Lee2010	Sparkclouds: Visualizing Trends In Tag Clouds.
Hong2008	Spartag.Us: A Low Cost Tagging System For Foraging Of Web Content
Korner2010	Stop Thinking, Start Tagging
Gupta2010	Survey On Social Tagging Techniques
Viegas2008	Tag Clouds And The Case For Vernacular Visualization
Christie2010	Tag Clouds As Social Signallers
Skoutas2011	Tag Clouds Revisited

Vig2010	Tag Expression: Tagging With Feeling
Kaser2007	Tag-Cloud Drawing: Algorithms For Cloud Visualization
Chen2009	Tagclusters: Semantic Aggregation Of Collaborative Tags Beyond Tagclouds
Sen2006	Tagging, Communities, Vocabulary, Evolution
Gou2010	Tagnetlens: Multiscale Visualization Of Knowledge Structures In Social Tags
Vig2008	Tagsplanations: Explaining Recommendations Using Tags
Cress2013	The Collective Knowledge Of Social Tags: Direct And Indirect Influences On Navigation, Learning, And Information Processing
Sinclair2008	The Folksonomy Tag Cloud: When Is It Useful?
Chiarella2012	The Influence Of Self-Organized, Social Text Signals On Readers' Behaviour: Collective Intelligence Or Unchecked Imitation?
Dron2008ar	The Trouble With Tags: An Approach To Richer Tagging For Online Learning
Heckner2008	Tree, Funny, To Read, Google
Chi2008	Understanding Navigability Of Social Tagging Systems
Gedikli2011	Understanding Recommendations By Reading The Clouds
Strohmaier2012	Understanding Why Users Tag: A Survey Of Tagging Motivation Literature And Results From An Empirical Study
Schrammel2009	Visual Search Strategies Of Tag Clouds - Results From An Eyetracking Study
Zamora2009	Visualizing Tags As A Network Of Relatedness
Gambette2009	Visualizing Text With A Tree Cloud
Ames2007	Why We Tag
Panke2009	With My Head Up In The Clouds": Using Social Tagging To Organize Knowledge"