HAL
open science

# Development of an Intuitive User-Centric Font Selection Menu

Girish Dalvi

## To cite this version:

Girish Dalvi. Development of an Intuitive User-Centric Font Selection Menu. Second IFIP WG 13.6 Conference on Human Work Interaction Design: Usability in Social, Cultural and Organizational Contexts (HWID), Oct 2009, Pune, India. pp.144-153, 10.1007/978-3-642-11762-6_12 . hal-01056246

**HAL Id: hal-01056246**

**https://inria.hal.science/hal-01056246v1**

Submitted on 18 Aug 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

# Development of an intuitive user-centric font selection menu

Girish Dalvi1

[1] Industrial Design Centre, Indian Institute of Technology Bombay, Powai, Mumbai, India.

**Abstract.** The font selection menu in most application software's is arranged alphabetically; in recent years one can also see the split menu approach being used. An alphabetical arrangement presupposes that the users are aware of the font characteristics and usage scenario through its name. Unless the font name specifies it; the scheme does not in any which way indicate the morphological features or the usage scenarios of a given font. In order to address these issues, a set of systematic experiments was conducted to capture the typeface classification strategies of users with a modified card sorting technique. The collected data was subjected to hierarchical clustering algorithms to come up with a collective user-centric classification system for Devanagari fonts. The classes created by this method appears to address the problems raised by alphabetical orderings by allowing the groupings to be more intuitive (and realistic) while retaining statistical validity over large user sets.

**Keywords:** Devanagari, Font Classification, Card Sorting, Menu Organization.

## 1 Introduction

The font selection menu is one of the most frequently used user interface components in today's software applications. Designing the structure of this selection menu hence becomes an important activity in the user interface design. A considerable amount of research has already been done on the design and evaluation of general menu driven systems. Lee and Raymond [1] have consolidated a considerable number of research findings in this area. Unfortunately there is no direct research on font selection menus and it is practically difficult to transfer these findings to the on to the creation of a type selection scheme.

It is nonetheless essential to look at the significant features of research in generalized menu design. Menu navigation is generally classified as broad (many options on one level) and narrow (less options). The number of options available at a level is called the 'breadth' of a level, and the amount of sublevels available is referred to as the 'depth' of the menu. In deep menu structures, users usually have to take many decisions before making a choice; this increases the risk of users losing their way within the menu hierarchies. In broad menu systems, users have too many options to choose from, which increases their search time. Hence, formulating the optimal balance between the depth and breadth of a menu structure becomes and an important factor in menu design. Earlier studies suggest that in terms of both time and

errors, users perform better in boarder hierarchies as against narrow (deep) hierarchies.

### 1.1 Menu Organization

Previous research has also focused on evaluation of organizational techniques for menu driven systems. Three popularly studied menu organizational techniques are alphabetical, frequency based and categorization. The organization of menus is largely determined by the users' search strategy. Some of the search strategies are: random (where users scan a given set of items randomly), redundant (one or more items are scanned more than once), sequential (items are scanned sequentially), exhaustive (items are scanned exactly once) etc. Empirical as well as eye-movement research strongly suggests that most searches are systematic with a certain amount randomness in them. Searching seems to be knowledge driven, knowledge allows users to focus search to a narrower region of the menu after which search is random or non-systematic [2] [3] [4].

Earlier studies have shown alphabetical ordering to be initially superior for non-experts; however its advantages deteriorate rapidly with practice [5]. In general search tasks, alphabetical organization performs slightly better than random organization, but is quite inferior when compared with any form of categorization. A study using two organizations—alphabetical and (functionally) categorized menus; suggests that a functional menu is more effective than an alphabetical menu, and users make fewer errors with the functionally organized menu [6]. Within categories alphabetical organization has been found equivalent to random ordering, however its efficiency is much worse than that of categorization within categories [5] [7] .

Eleanor Rosch proposed, categories created by mapping the perceived world structures are the most efficient ones [8]. This suggests that menu categories should reflect the cognitive structure of non-expert users, as they are the ones who use the applications. There is strong evidence that suggests categories created by non-experts perform better than the ones created by experts [9] [10] [11]. The reliability of the created categories increases when consolidated with the judgements of many users. Search performance can also be ameliorated by using improved categorization techniques [5] [9] [12]. It is also not surprising note, that users have performed better with categorized menus than with randomly organized menus [5] [7] [12] [13].

It is evident from the stated studies that broad menus generally perform better than narrow ones; and (functionally organized) categorized menus are more efficient than alphabetically or randomly organized menus. In categorized menus itself, categories created by non-experts (novices) are more efficient than the ones created by experts.

### 1.2 Card Sorting

Card sorting is a common generative technique used by researchers to capture how users arrange a set of objects. The captured information can then be used to create an information structure for the given group of items. Largely two variations of card sorting are used: Open card sorting, where participants are asked to create classes

from a given set of items and closed card sorting where participants are given a set of items and a set of classes. Open card sorting is used to generate classes whereas closed card sorting is used more to verify classes created by open sorting. Authors [14] [15] have written introductions to card sorting and the differences between open and closed card sorts. However, within these two techniques if the given data contains many connections amongst its elements, participants often express the need to place a single item in two or more classes. In order to incorporate this need; in this study participants were allowed to place a single item into more than one category if they so wished, so as to create fuzzy classes.

Card sorting data can be analyzed by a variety of methods; the selection of the method depends upon the goal or intent of the study. If the intent of the study is to create a highly structured setting, cluster analysis has at times used to statistically analyze card sorting data and to visualize its results. Factor analysis has also been used [16] to analyze card sort data where the intent of the study was to identify attributes or properties used by participants to assess categories. Visual examination or "eyeballing" [17] has also been used to gauge the overall organizational structure of card sorts. One of the goals of this study was to create a menu structure for application software's (a more intuitive typeface selection scheme); cluster analysis was hence used to come up with distinct clusters for Devanagari typefaces.

## 2 Goals of the Research

The overall aim of this research was to examine how various individuals classify a set of Devanagari fonts. An experiment was designed to capture the different Devanagari font classification strategies used by people through a modified card sorting methodology. The primary objectives of this experiment were:

1. To understand the overall strategies that participants use to classify Devanagari fonts.
2. To identify the significant parameters used to classify Devanagari fonts.
3. To record the nomenclature given by participants to various font classes as well as parameters.
4. To discover which are the parameters that are given preference while creating classes and which are overlooked.
5. To create a user-friendly classification system for menu structures for software applications.

The first four objectives listed above are part of a larger study and hence have not been discussed in this paper. This paper focuses mainly on the accomplishment of the fifth objective viz. creation of a intuitive font classification system for software menu structures.

## 3 Experiment

### 3.1 Stimulus

The stimulus size for the experiment was a 25cm x 4cm cardboard sheet (see Fig.1) with one line of non-sense Devanagari text set at 72 point size (normalized kana height 1.25cm). Thirty font samples were used for the study. Sampling of the font samples was partially random (from a list of typefaces published by CDAC) and partially purposive (fonts in popular use and distribution were selected). The sample card sheets were randomly numbered, with the number being placed on the lower right corner of the sample sheet.

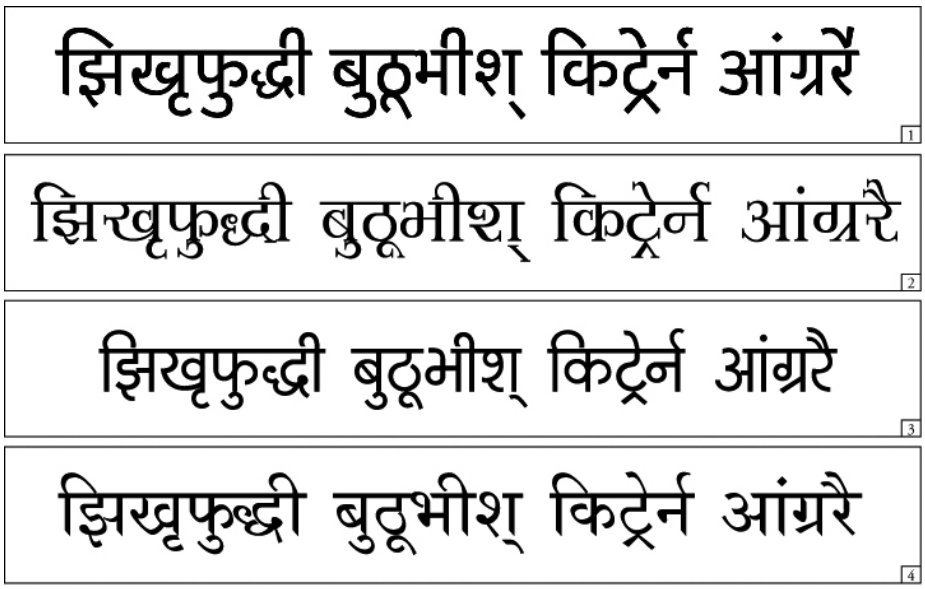


**Fig. 1.** The above figure shows four sample stimulus cards used in the experiment.

### 3.2 Participants

The exercise was administered on thirty-eight participants (twenty males and eighteen females). Their age ranged from 21 years to 42 years, with an average of 25 years (S.D. = 2 Years). All the participants had completed a minimum level of education and had formally studied Hindi, Marathi or Sanskrit languages till the $10^{th}$ grade.

### 3.3 Experimental Procedure

The experiment began with the participants being given a warm up exercise. In the warm up exercise, each participant was given a set of thirty three small cards. On each

of the card was a name of an animal. Participants were told to classify them into as many categories as they wished; and that they could have multiple categories within a category. It was also stated that one animal could belong to more than one category. Since most participants implicitly assumed strict classification within the task, the warm up exercise ensured that participants cleared up all their doubts about the classification exercise and that it was permissible to create fuzzy classes.

The main procedure of the experiment was conducted after a warm up exercise. Participants were given thirty cardboard mounted font sample cards. They were then asked to classify them according to any logic that they pleased; with no restriction on the number of categories that they could make. It was also stated that they could have as many sub-categories as they wanted within the main categories that they created. The classification procedure took around 30 to 120 minutes to complete. After they had completed classifying, participants were asked to give names to each of the classes which had been created; and elucidate the basis on which each of the groups was made. They were also asked to state all the properties of the classes that they had created. When the participants had completed classifying and naming the groups they were asked to reconfirm their classification scheme, and see if they wanted to make any changes (to see if their classification was consistent with the applied logic). Once the classification scheme was confirmed, the categories and the properties attributed to them were recorded on a data collection sheet.

It was essential here that user's were given the option of placing one element into two or more classes. There is sufficient proof [8] [18] that mental categorization is a graded and fuzzy phenomenon. In fonts, a particular font can be categorized considering several properties, and a typical mental model can have multiple rationales, towards its grouping. Many users combined multiple viewpoints for font classification; this necessarily demands most of the fonts to be placed in two or more groups. For example it was common to see users grouping fonts partially based on formal elements (such as written with a traditional Devanagari pen) and partially on the basis of usage (like decorative text or headline text) or even subjective keywords (such as energetic text, professional text etc.).

### 3.4 Data Recording

Data was recorded in a tabular fashion, with the categories and subcategories being written in a top to bottom manner. After the participants had completed classifying the given samples, the groups were recorded with the help of the numbers assigned to each of the samples. The top most categories were recorded first; the name and the description provided by the participants was then recorded besides the recorded numbers along with the associated properties of each class.

### 3.5 Treatment of Data

Clustering algorithms require data to be specifically formatted before it can be processed. Converting card sort data into these formats is not an obvious process and is not usually discussed in literature on card sorts [16].

In this study the recorded data was used to create a proximity matrix for the thirty samples. The matrix here is a similarity matrix rather than a dissimilarity matrix. The proximity matrix $p$ is an $m$ by $m$ matrix ($m$ is sample size and is thirty in our case) containing all the pair wise similarities between the samples considered. If $x_i$ and $x_j$ are the $i^{th}$ and $j^{th}$ objects respectively, then the entry at $i^{th}$ row and $j^{th}$ column of the proximity matrix is the similarity $s_{ij}$ between $x_i$ and $x_j$. Here in our case, the number of times a typeface sample was grouped together at any level by the participant increments the similarity in the proximity matrix. Such an operation was carried out for each category created by the participants. The resulting matrix was then used as an input for the cluster analysis algorithm.

# 4 Experimental analysis

## 4.1 Cluster Analysis:

Cluster analysis groups objects based on information found in the data describing the objects or their relationships [19]. The goal is that the objects in a group should be similar to one another and different or dissimilar from the objects in the other groups. The greater the similarity within the group and the greater the dissimilarity between the outside groups, the better the clustering. Since the objective of this study was to create distinct typographic categories for practical use; hierarchical cluster analysis was used.

*Hierarchical Methods:*

A hierarchical clustering method works by grouping data elements into a tree of clusters. Hierarchical clustering methods are of two kinds: agglomerative clustering where the hierarchical decomposition is formed in a bottom up merging manner and divisive clustering where the hierarchical decomposition is formed in a top-down (splitting) fashion. In agglomerative hierarchical clustering [20] the bottom-up scheme is initiated by placing each element in its own cluster and then merging these atomic clusters into larger clusters, until all of the elements part of a single cluster. We used Ward's minimum-variance to calculate the distance between two clusters. Ward's minimum-variance uses the ANOVA sum of squares between the two clusters added up over all the variables [21]. While creating each cluster the within cluster sum of squares is minimized over all partitions obtainable by combining two clusters from the previous creation. Ward's method tends to join clusters with a small number of observations, and strongly tends towards producing clusters with the same shape and with roughly the same number of observations.

For the $i^{th}$ cluster, the Error Sum of Squares is defined as $\mathrm{ESS_i}$ = sum of squared deviations from the cluster centroid. If there are C clusters, the Total Error Sum of Squares is defined as:

$$\mathrm{ESS}(\mathrm{C_i}) = \sum_{a=1}^{n_i} |x_a - \frac{1}{n_i}\sum_{b=1}^{n_i} x_b|^2$$

Consider the union of every possible pair of clusters. The two clusters would then be combined whose combination results in the smallest increase in ESS. The distance between clusters can hence be calculated as:

$$d_{\text{ward}}(C_i, C_j) = ESS\,(C_i, C_j) - [ESS(C_i) + ESS(C_j)]$$

For this study Ward's minimum variance method was chosen over the others, because:

1. It is good at recovering cluster structure, and yields unique and exact hierarchy.
2. It does not leave any "loose ends". No clusters with only one or a few elements. All data is grouped in bite size chunks, which can be studied further.
3. Aberrant points are also grouped together, which might not have anything in common with each other except for the fact that they are dissimilar from the other objects.

The clusters generated (dendrogram) by using Ward's method on the generated proximity matrix, are shown in figure 2.

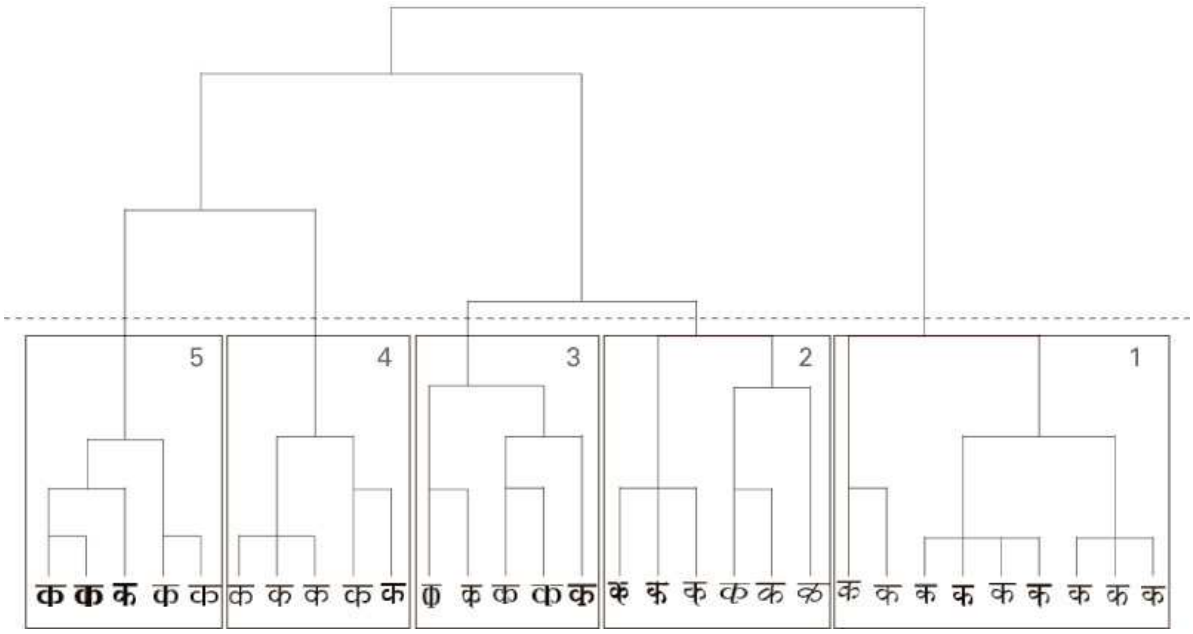


**Fig. 2.** Resulting tree, cutoff, and final clusters.

### 4.2 Cluster Validation

Validating the results of cluster analysis requires subjective decision making [22]. Besides primary validity, which measures how valid the cluster analysis is overall,

there are measures of secondary validity which assess whether the clusters have certain desirable properties. The following cluster validation technique was used to evaluate this study.

*Agreement of Classifications Based on Split Samples of Data:* Sinha [23] has suggested that the objects of the original data matrix be randomly split into two subsamples and separate cluster analyses be run on each to produce two separate classifications. To be judged valid, the two classifications should agree: the number of classes should be the same and their defining attributes should be the same. As part of the cluster validation scheme, the resulting clusters were validated incrementally, first on a set of twelve participants then on a set of twenty-four participants and then on the final set of thirty eight participants. During each of these stages the clusters remained fairly stable and only one or two members changed their position within the lowest sub-categories.

## 5 Results

The resulting dendrogram was subjected to the cutoff line (shown by the dashed line in fig 2.). The cut-off line was chosen by looking at the major (graphical) groups within the clusters, since "the cut-off points to identify clusters … are a matter of subjective judgment by the researchers" [24]. After the cut-off point the resulting clusters can be seen in Fig 2.

Hierarchical cluster analysis yielded five stable and distinct clusters.

Fonts from the first cluster from the left (Fig.2 (1)) have been labelled by many participants as "Traditional text", "Standard Text", "Textbook or Newspaper Text". The formal properties of the whole group indicate that all of them have been drawn with the help of a canted pen (8 right canted, 1 left canted), due to which one sees canted vertical terminals, at inclined axis and high contrast among its letters. This group is further classified into three subgroups; based on their visual characteristics one can notice that the first two groups have been differentiated according to their counters (open vs. closed). Some participants have labelled the closed counters as "Marathi Text" as against open counters as "Hindi or Other Text".

The second cluster (Fig.2 (2)) in the resulting set is that of the "scripts", the only common visual feature that all the elements share is that their vertical terminals have a "swoosh" which either goes to the right or to the left in a smooth manner. Two of them have an inclination towards the right. They are further bifurcated in to two other subgroups, the first one seems to be the one where the fonts have a darker grey value and have a narrower appearance than the regular fonts. The second group contains fonts which are wider in their appearance and have a lighter grey value. Two fonts in this group have an inclination towards the right. One font instead of having a "swoosh" as the vertical terminal has a rounded (felt-tip pen) terminal.

In the third group of clusters (Fig.2 (3)), the significant common visual feature seems to be the serif. These are fonts influenced by Latin typefaces. There is a considerable amount of variation in the serifs of the fonts; from single sided to two sided, thin, thick and inclined. There also is a considerable difference in the width of the fonts; two narrow fonts form a sub group, while three broad fonts form the other

sub group. There is further differentiation in this group as fonts with medium grey value and fonts with dark grey value.

In the fourth cluster (Fig.2 (4)) the prominent visual feature of this group is the fact that all the fonts in this group are mono-linear i.e. the strokes are of uniform thickness. All the elements have a horizontal vertical terminal except in one font sample which has a traditional Devanagari inclined terminal. This group is further subdivided into two categories based on the typographic colour; the medium weights form one group while the darker weights form the other group.

The fifth cluster and final cluster (Fig.2 (5)) contains typefaces which are either dark in their grey value, and/or are very broad or tall. The description that some of the participants have given is that of "display typefaces" or typefaces that one would use for headlines. They are again sub grouped according to their weights; the first subgroup is that of the thick, dark or bold fonts; while the other is group of varying widths i.e. narrow and broad fonts.

## 6 Conclusion

Although this study did not explicitly set out capture cultural preferences among participants while categorizing fonts; certain cultural traits were inevitable reflected in the font groupings. For example, the subgroups of the first major category have been labelled by some participants as "Marathi" versus "Other" text. This is a typical reflection of the fact that closed counter fonts; especially a closed counter "क" (ka) was considered by some participants as belonging to the Marathi language as against the open counter forms which were considered "Hindi, or Other" texts.

Earlier research has shown categorized menus to be more efficient than alphabetical or randomly organized menus. Within categorized menus categories created by non-experts (novices) have yielded better results than the ones created by experts. We have carried out a thorough empirical experiment where the Devanagari font classification schemes created by a group of non-experts were captured. The captured data was then analysed by cluster analysis to create five broad categories. The five broad categories with the most popular titles are: Traditional Text, Script, Serifs, Equal Stroke and Headline Fonts.

Understanding how people perceive a set of objects and the relationships that they associate between those objects is a complex phenomenon which cannot be addressed by quantitative analysis alone. In order to have a complete understanding of the users' mental model a combination of qualitative and quantitative techniques must be used. Our study has produced a large amount of data—in this paper we have presented the initial results and analysis. Further tests on the usability of the created categories and on the given titles in an application environment will be needed to come up with conclusive solutions. This method nonetheless can be used to generate categorized menu structures for usability evaluation. Capra has shown how Fuzzy Card Sort data can be treated with factor analysis to give fuzzy classes [16]. This is practical and appropriate for websites but not software applications. Our method of using hierarchical cluster analysis on Fuzzy Card Sort data yields rigid classes which seem appropriate for software application menus.

## References


1. Lee, E., Raymond, D.: Menu-driven systems. Encyclopedia of Microcomputers, 101–127 (1993)
2. MacGregor, J., Lee, E.: Menu search: random or systematic? International Journal of Man-Machine Studies 26(5), 627-631 (1987)
3. Vandierendonck, A., Hoe, R., Soete, G.: Menu search as a function of menu organization, categorization and experience. Acta Psychologica 69(3), 231–248 (1988)
4. Hendrickson, J.: Performance, preference, and visual scan patterns on a menu-based system: implications for interface design. ACM SIGCHI Bulletin 20(SI), 217-222 (1989)
5. Hollands, J., Merikle, P.: Menu organization and user expertise in information search tasks. Human Factors 29(5), 577-586 (1987)
6. Mehlenbacher, B., Duffy, T., Palmer, J.: Finding information on a menu: linking menu organization to the user's goals. Human-Computer Interaction 4(3), 231-251 (1989)
7. McDonald, J., Stone, J., Liebelt, L.: Searching for items in menus: The effects of organization and type of target. In : Proceedings of the 27th annual meeting of the Human Factors Society, p.834–837 (1983)
8. Rosch, E.: Natural categories. Cognitive Psychology 4(3), 328-350 (1973)
9. Lee, E., Whalen, T., McEwen, S., Latremouille, S.: Optimizing the design of menu pages for information retrieval. Ergonomics 27(10), 1051–1069 (1984)
10. Fischhoff, B.: Creating categories for databases. International Journal of Man-Machine Studies 27(1), 33-63 (1987)
11. Hayhoe, D.: Sorting-based menu categories. International Journal of Man-Machine Studies 33(6), 677 - 705 (1990)
12. Parkinson, S., Sisson, N., Snowberry, K.: Organization of broad computer menu displays. International journal of man-machine studies 23(6), 689–697 (1985)
13. Card, S.: User perceptual mechanisms in the search of computer command menus. In : Proceedings of the 1982 conference on Human factors in computing systems, Gaithersburg, Maryland, United States, pp.190-196 (1982)
14. Robertson, J.: Information Design Using Card Sorting. In: Step Two Design. http://www.steptwo.com.au/papers/cardsorting/
15. Maurer, D.: Card sorting: a definitive guide. In: Boxes and Arrows. http://www.boxesandarrows.com/archives/card_sorting_a_definitive_guide.php
16. Capra, M.: Factor analysis of card sort data: an alternative to hierarchical cluster analysis. In: Proceedings of the Human factors and Ergonomics society 49th annual meeting. Blacksburg: VA (2005)
17. Nielsen, J.: User Interface Design for Sun Microsystem's Internal Web. (1994)
18. Lakoff, G.: Women, Fire, and Dangerous Things. University Of Chicago Press (1990)
19. Wille, L.: New Directions in Statistical Physics: Econophysics, Bioinformatics, and Pattern Recognition. Springer (2004)
20. Han Jiawei: Data Mining: Concepts and Techniques. Morgan Kaufmann (2006)
21. Clark, V.: Sas/Stat 9.1: User's Guide. SAS Publishing ( 2004)
22. Romesburg, C.: Cluster Analysis for Researchers. Lulu.com (2004)
23. Sinha, R. N.: Uses of Multivariate Methods in the Study of Stored-grain Ecosystems. Environmental Entomology 6, 185-192 (April 1977)
24. Mukherjee, R.: Classification in social research. SUNY Press (1983)