technical note technical note

Information Organization for a Portal Using a Card-Sorting Technique

Vicki Ahlstrom, ATO-P

Kenneth Allendoerfer, ATO-P

September 2004

DOT/FAA/CT-TN04/31

U.S. Department of Transportation
Federal Aviation Administration

William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents or use thereof. The United States Government does not endorse products or manufacturers. Trade or manufacturer's names appear herein solely because they are considered essential to the objective of this report. This document does not constitute FAA certification policy.

This report is available at the Federal Aviation Administration, William J. Hughes Technical Center’s full text, technical reports web site: http://actlibrary.tc.faa.gov in Adobe Acrobat portable document format (PDF).

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/CT-TN04/31		

4. Title and Subtitle	5. Report Date
Information Organization for a Portal Using a Card-Sorting Technique	September 2004
	6. Performing Organization Code ATO-P

7. Author(s)	8. Performing Organization Report No.
Vicki Ahlstrom and Kenneth Allendoerfer, ATO-P	DOT/FAA/CT-TN04/31

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Federal Aviation Administration NAS Human Factors Group, ATO-P William J. Hughes Technical Center Atlantic City International Airport, NJ	11. Contract or Grant No.

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Federal Aviation Administration Office of Knowledge Management, ACK-1 William J. Hughes Technical Center Atlantic City International Airport, NJ 08402	Technical Note
	14. Sponsoring Agency Code ACK-1

15. Supplementary Notes

16. Abstract

The Office of Knowledge Management at the Federal Aviation Administration's William J. Hughes Technical Center asked researchers from the National Airspace System Human Factors Group to provide human factors input for a new enterprise information portal. The portal would contain a large amount of information. In order for the portal to be successful, the portal needed to provide information so that the users could easily find what they needed. Researchers from the William J. Hughes Technical Center used a simple method called a card sort to organize the information in a way that would reflect the mental model of the users. This document describes the card-sorting methodology. The researchers developed a set of information that potentially could be included in the portal through a combination of items identified by the sponsor as high priority items, information provided on the current employee intranet, and structured interviews with potential portal users. The researchers placed the information items on index cards. Participants representing various types of end users sorted the cards into categories and then named the categories. The researchers grouped similar categories and analyzed the data to determine categories and strength of association between items. The analyses led to the development of a taxonomy for the organization of information items within the portal. This report describes the card-sorting methodology and provides lessons learned on the application of this technique. Although the card-sorting technique alone will not produce an effective information architecture, it is a valuable technique that can lead to a more usable end product.

17. Key Words	18. Distribution Statement
Card sorting Human Computer Interface Portal Taxonomy	This document is available to the public through the National Technical Information Service, Springfield, Virginia, 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	22	

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

Table of Contents

List of Illustrations

Executive Summary

The Office of Knowledge Management (ACK-1) at the William J. Hughes Technical Center (WJHTC) is sponsoring the development of a web-based portal for employees. The portal will centralize and standardize numerous information systems currently used by WJHTC employees in their work. To increase the usability of the system, the office tasked the National Airspace System Human Factors Group to conduct several user-centered design and usability testing activities. This technical note describes the methodology used in one of those activities.

In order for the portal to be successful, the information needs to be organized so that users can find what they need quickly and accurately. The challenge for the portal designers is how to organize the information to achieve this goal for portal users. Different users or user communities may have very different views and priorities regarding the information contained in the portal. It is not at all clear which items belong together or what hierarchical categories would make sense to the users.

Card sorting is a technique used to gain insight on how users mentally categorize information. By performing a card-sorting task, researchers can create a taxonomy and organize information in a way that corresponds to how intended users of a system think about the information the system contains. By basing the organization on the results of the card sort, the system is more likely to match the mental model of the users, facilitate their ability to find information, and achieve good usability.

This technical note describes the card-sorting methodology and the techniques for analyzing results. We use the employee portal project as a running example to discuss the methodology and the benefits and drawbacks of different analysis techniques.

Researchers from the WJHTC created a list of information items to include in the portal based on high priority items provided by the sponsor, items currently in the intranet, and structured interviews with potential end users. The researchers created sets of cards where each card contained one of the items from the list. Participants sorted the cards into groups based on their own ideas of where items belong and then named each group. Once all of the participants completed the card sort, the data were analyzed using a set of converging methods to determine overall categories and organization. The first method examined the categories and identified patterns that were similar across participants. The researchers followed this by two more quantitative methods: cluster analysis and factor analysis. The researchers found an agreement between the methods.

The result of the card-sorting task was a taxonomy that provided an initial structure to the portal. The taxonomy will be verified and refined through iterative usability testing.

1. INTRODUCTION

An information system such as a web-based portal succeeds when users find the information they want quickly, accurately, and with a minimum of effort. Success depends on the content and structure of the information contained in the system, sometimes called the information architecture. Organizing a large system with many types of information presents a significant engineering challenge. This is especially true when the system affects multiple user communities, each with its own tasks, abilities, and requirements.

Navigation effectiveness in a system such as a portal is dependent on how the information is structured. To organize information, developers often rely upon a taxonomy. A taxonomy is a formal method for classifying information, including agreed-upon terms and relationships. The best known taxonomy for published literature is the Library of Congress Subject Headings (LCSH). General taxonomies like LCSH are usually too broad for specialized domains or applications. As a result, many taxonomies exist that focus on information important to a specific domain. The Library of Medicine Medical Subject Headings (MeSH), is an example that classifies medical information according to common definitions and a controlled vocabulary (National Library of Medicine, 2004).

A good taxonomy can serve as the basis for an information architecture that makes the system usable and effective. A usable system increases user productivity, acceptance, and satisfaction. Architectures based on taxonomies work best when the taxonomy follows the users' mental model of the information contained in the system. That is, a good taxonomy reflects meanings and relationships that users know and agree upon. For example, in the medical field, physicians undergo extensive professional training that standardizes the terminology they use throughout their work. This standardization is reflected in MeSH, which uses terms and relationships that physicians have already been trained to understand. In some professions, courses are offered on the effective use of the domain-specific taxonomy as part of basic professional training.

No taxonomy exists for many applications of information systems. Existing taxonomies are either too broad or do not apply to the domain at hand. This is especially true for systems supporting a broad community where users vary in their interests, responsibilities, priorities, and backgrounds. In these cases, creating the taxonomy typically falls to system engineers or to a development team composed of stakeholder representatives. Unfortunately, members of a development team may not understand the information in the same way as the intended users and the resulting taxonomy may represent the team's mental model not the users'. The task facing the development team, then, is to investigate the users' mental models and base the taxonomy on their findings. In practice, there is not a single mental model shared by every user but the taxonomy can be constructed to match the mental models of as many users as possible, especially for their critical tasks.

Human factors engineers, given their experience collecting and analyzing data on how people do their jobs, can help system developers create taxonomies that are consistent with the mental models of users. They can evaluate taxonomies for usability. Card sorting is a technique commonly in the human factors field to gain insight into the mental models of users (Nielsen & Sano, 1994; Department of Health and Human Services, 2004). Card sorting involves creating sets of cards where each card contains one item that may be included in the system. Participants

sort the cards into groups based on their own ideas about where items should go. Depending on the approach used, the participants may name the groups using their own terminology (known as an "open card sort"), or use predetermined group names ("closed card sort"). Once groups have been completed, the data are analyzed to determine overall categories and organization. This can be accomplished by manually identifying patterns across participants or through statistical techniques. In all cases, careful investigation is required to determine the best organization and labeling of information.

The advantages of using the card sorting technique are that it is simple to administer, inexpensive, and fast. The information gained through a card sort is straightforward to present and easily understood by stakeholders. Some disadvantages of the technique are that data analysis can be time consuming, it can be difficult to resolve the differences between users, and although the users categorize the items, it does not involve the users actually using the system to accomplish a task (i.e., to locate information). In other words, a participant may put an item in category "A" in the card sort, but may look in category "B" when actually searching for the item using the system. Although this technique can provide valuable input to the organization of information, it should not be the only technique used to develop information architecture for a large system.

This technical note describes the card sorting technique. We describe the insights a card sort can provide, how to collect and analyze card sorting data, and discuss lessons learned from our own applications of the method. Throughout the document, we use a recent project at the Federal Aviation Administration (FAA) William J. Hughes Technical Center (WJHTC) as an example.

2. DESCRIPTION OF THE EXAMPLE

In 2004, the WJHTC Office of Knowledge Management (ACK) sponsored a project to consolidate and standardize many information systems used by employees into a web-based portal. Some examples of systems to be included were the existing employee intranet, several human resources systems, the employee directory, various accounting and tracking systems, management information systems, e-mail and collaboration systems, and the WJHTC library card catalog.

The primary rationale for the project was to increase employee productivity by standardizing and simplifying user interfaces, improving access to information, reducing redundant data entry, and allowing users to log on with a single username and password. The project also had broader goals of increasing employee collaboration and improving overall knowledge management at the WJHTC. In addition, the project sought to reduce the hardware and software infrastructure (e.g., servers, databases) and workload needed to maintain so many separate systems.

3. DATA COLLECTION METHOD

As the name implies, the card sorting method involves participants placing cards into groups that are meaningful to them. Each card contains one item of information that may be included in the

system. An item may be a document, a function, a form, or a system. Some examples of items in the portal project were:

- WJHTC Announcements: This item is a document that provides information about the operation of the WJHTC, including information about visitors, human resources, and facility maintenance schedules.

- 411 Directory: This item is a function that allows users to find contact information about other employees at the WJHTC, similar to a white pages phonebook. For example, a user may enter another employee's name and will receive the employee's phone number, organization number, and office location.

- Badge Authorization Form: This item is a document that allows users to print and complete the paperwork needed to obtain a WJHTC security badge.

- FedTrip: This item is a system that allows users to find and reserve arrangements for transportation and lodging for official travel.

The program managers for the portal wanted us to include as many items as possible when we created the taxonomy. Even though the initial deployment of the portal would not include all the items, the program management wanted to allow room for future growth. Because of the large amount of information used by WJHTC employees and organizations, we used several methods to identify information that could be incorporated into the portal. We obtained an initial list of items through meetings with the sponsoring office where high-priority items were identified. We expanded this list through examinations of the existing WJHTC intranet and systems used by employees, and through targeted interviews with members of specific user communities. For example, we asked individual employees about tasks specific to their managerial or administrative responsibilities. While they sat in front of their computer, we asked them to identify and describe frequently used systems, functions, documents, and forms. Almost all of the common items identified by the managers or administrators were bookmarked in their browser or contained in a special folder on their computer desktop. Our final list included 95 items, 10 of which were generally used only by managers and their assistants. The other 85 items were applicable to all WJHTC employees though employees used only an individual subset of these in their day to day work.

3.1 Participants

The participants in a card sorting task should reflect the breadth in abilities, jobs, and environments of the targeted user community. Having a broad sample helps developers be more confident they have not missed important perspectives and that the system will work reasonably well for the entire user community. This is different than trying to achieve a truly representative sample. Some jobs in an organization are very important, but are accomplished by so few people or so infrequently that a truly representative sample would neglect them in favor of frequent tasks accomplished by many people. For this reason, we recommend using a sample with several "regular" users along with selected important "outliers" who are still members of the community but whose jobs are markedly different from the regular users.

In the portal project, we tested nine participants drawn from different job categories at the WJHTC. These categories included research scientists, administrative assistants, financial analysts, facilities personnel, and group managers.

3.2 Materials

We used standard 3 x 5 inch white index cards with handwritten or printed labels for each of the card sort items. We provided each participant with writing implements and blank 4 x 6 inch index cards, and lined paper to record category names. We built three individual packs of index cards, so it was possible to handle three participants simultaneously.

Although it is possible for researchers to create complex computer applications for administering a card-sorting task on a computer, we recommend using physical cards. Using physical cards requires almost no training and is very inexpensive. More importantly, it is very easy for participants to reorganize their stacks of physical cards when they change their minds. Even a well-designed computerized card sorting application would have a very hard time matching this simplicity and speed. Forcing users to follow slower, more precise sorting actions may discourage them from reorganizing their stacks and may encourage participants to commit to a particular organization "too early" because reorganizing the cards on the screen is too much hassle.

The names that we used for each item were the official names of the system, document, or form, or the name as it exists in the current employee intranet. We did not make an effort to change these names for fear of introducing our own biases to the nomenclature. Each card contained only the name of the item, but no description. The rationale for this approach is that the labels should be understandable to the participants but the less written about the item on the card the better. It may seem helpful to include a descriptive phrase about each item on the card. We find that this slows down the process and, worse, may influence the participant's judgment. Instead, we recommend using just the name of the item or its acronym when the acronym is better known than the actual name.

3.3 Procedure

A card sorting task can be accomplished nearly anywhere. It is helpful to have a large flat table or desk so that participants can place cards into stacks and easily move them around. For the portal project, we conducted the task in empty conference rooms with large tables so the participant could work undisturbed. Participants worked in different rooms to prevent them from influencing each other's groupings. Although card sorts are sometimes performed by groups of individuals working together, we decided to have the participants work individually. If the participants worked as a group, deciding the categories through a consensus process, individual approaches to the information organization might be lost. As users would likely use the portal individually, we wanted to capture the diversity of approaches to the information. The drawback to this approach is that during the analysis phase, we had to resolve the differences between the user categories.

Before each run, the participants should receive instruction sheets that describe the purpose of the study, the tasks to be accomplished, and their rights as participants. These sheets must

conform to FAA rules regarding informed consent and confidentiality. The sheet we used in the portal project is provided in Appendix A. Before beginning the task, participants should be informed if the group names are fixed or open, how many groups they may use, and what to do with cards that they do not understand or cannot group. In the portal project, we asked the participants to choose their own group names, allowed them to use as many groups as they wanted, and told them that should create an "I Don't Know" group if necessary. Before each run, we thoroughly shuffled the cards so that previous participants did not influence the current one.

Researchers should allow participants to work at their own pace and alone. Researchers should check on the participant occasionally but should try to intrude as infrequently as possible. In the portal project, participants took about 60 minutes to categorize 95 cards. The conference rooms that we used for the portal project were equipped with telephones. We gave the participants a number where they could reach us in case they had any questions as they were going through the exercise.

When the participant has completed sorting the cards, researchers should examine the stacks and make sure that it is clear which group names correspond to which group. We found that using rubber bands to secure the groups and the group names was a simple way to keep the data together properly.

4. DATA ANALYSIS AND REPORTING METHODS

Card sorting data can be analyzed in a variety of ways. Depending on the needs of the study and the scope of the project (especially the number of participants), the methods vary in how appropriate, interpretable, or helpful they are.

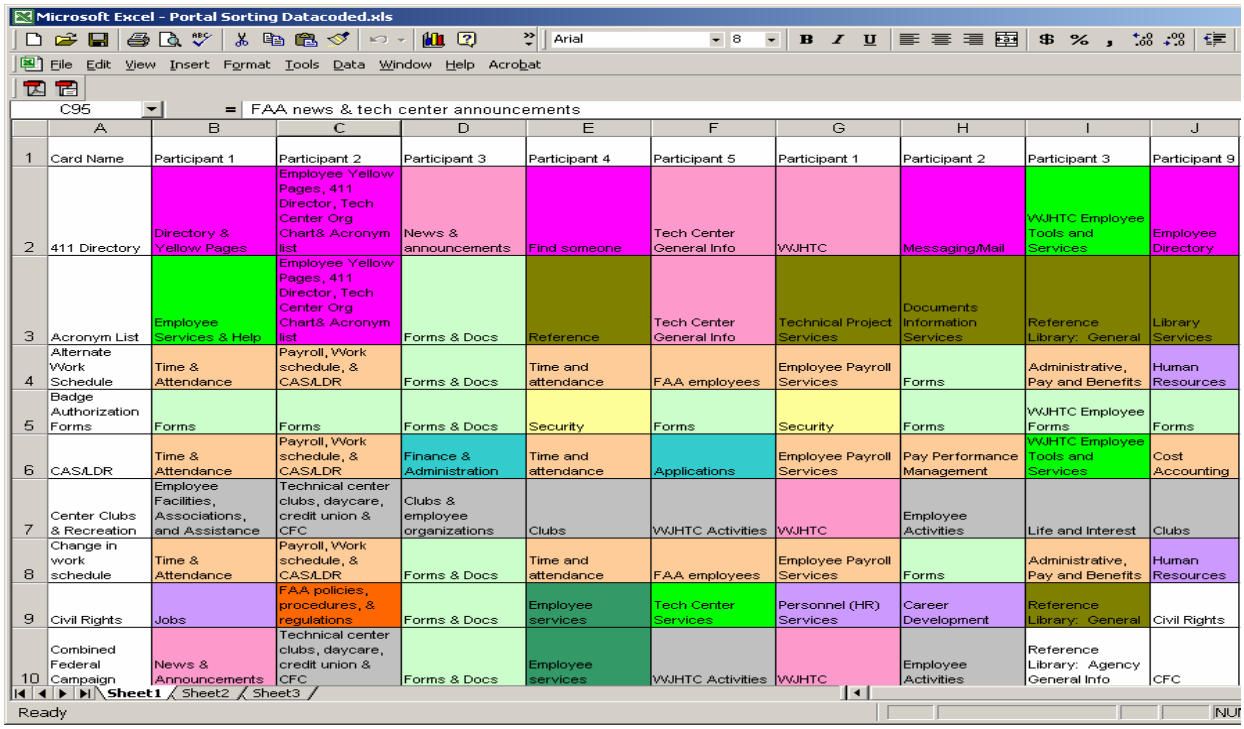
One of the downsides of using physical cards is that they are easy to mix up accidentally. For this reason, we recommend entering each participant's data immediately. A spreadsheet can be built ahead of time for this purpose. For the portal project, we built a spreadsheet with each item represented as a row and each participant represented as a column. We recorded the name of the group assigned by the participant in each cell.

4.1 Deriving Categories By Hand

After all data are entered, it is straightforward to examine the group names used and look for patterns. Patterns are indicated by different participants using the same words or ideas. These patterns are used to derive categories. For example, one participant created a group called "Directory," another created "Find someone," and another created "Employee Directory." This pattern resulted in a derived category that dealt with locating and communicating with fellow employees.

We assigned a color code to each derived category and colored each cell in the spreadsheet with the chosen code (see Figure 1). We used colors instead of category names so we could focus on the categorization of the data before optimizing the name of the category. The goal of the color coding was to collate all of the responses and present the data in a more meaningful way. The advantage of using color-coding was that it allowed us to present a clear picture to the project

sponsors showing how the participants organized the items into groups. This method also allowed us to rapidly identify the categories that were most and least used for particular items.



	A	B	C	D	E	F	G	H	I	J
1	Card Name	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Participant 1	Participant 2	Participant 3	Participant 9
2	411 Directory	Directory & Yellow Pages	Employee Yellow Pages, 411 Director, Tech Center Org Chart& Acronym list	News & announcements	Find someone	Tech Center General Info	WJHTC	Messaging/Mail	WJHTC Employee Tools and Services	Employee Directory
3	Acronym List	Employee Services & Help	Employee Yellow Pages, 411 Director, Tech Center Org Chart& Acronym list	Forms & Docs	Reference	Tech Center General Info	Technical Project Services	Documents Information Services	Reference Library: General	Library Services
4	Alternate Work Schedule	Time & Attendance	Payroll, Work schedule, & CAS/LDR	Forms & Docs	Time and attendance	FAA employees	Employee Payroll Services	Forms	Administrative, Pay and Benefits	Human Resources
5	Badge Authorization Forms	Forms	Forms	Forms & Docs	Security	Forms	Security	Forms	WJHTC Employee Forms	Forms
6	CAS/LDR	Time & Attendance	Payroll, Work schedule, & CAS/LDR	Finance & Administration	Time and attendance	Applications	Employee Payroll Services	Pay Performance Management	WJHTC Employee Tools and Services	Cost Accounting
7	Center Clubs & Recreation	Employee Facilities, Associations, and Assistance	Technical center clubs, daycare, credit union & CFC	Clubs & employee organizations	Clubs	WJHTC Activities	WJHTC	Employee Activities	Life and Interest	Clubs
8	Change in work schedule	Time & Attendance	Payroll, Work schedule, & CAS/LDR	Forms & Docs	Time and attendance	FAA employees	Employee Payroll Services	Forms	Administrative, Pay and Benefits	Human Resources
9	Civil Rights	Jobs	FAA policies, procedures, & regulations	Forms & Docs	Employee services	Tech Center Services	Personnel (HR) Services	Career Development	Reference Library: General	Civil Rights
10	Combined Federal Campaign	News & Announcements	Technical center clubs, daycare, credit union & CFC	Forms & Docs	Employee services	WJHTC Activities	WJHTC	Employee Activities	Reference Library: Agency General Info	CFC

Figure 1. Example of color-coding the card items by category.

Each item received a score by category based on the number of participants who placed the item in that category. For example, in the first row of Figure 1, the *411 directory* item received a score of five for the magenta category, three for the pink category, and one for the bright green category. We developed an initial taxonomy using a majority rules approach where we would locate the information on the card in the category with the highest score. In the example, we would locate the *411 directory* item in the magenta category. Clearly, however, a number of people would look in the pink category for this system instead. Therefore, this would be an instance where we suggest putting links to the system in more than one location to accommodate the widest range of users.

This manual method has a number of benefits. First, it is straightforward to execute and does not require sophisticated analysis tools. Second, unlike many statistical techniques, small sample sizes do not restrict it. Third, results from this method are easy to present to audiences who are not experienced interpreting multivariate statistics. The method provides quick answers that are easy to understand and act upon.

This method also has several drawbacks. First, there is a level of subjectivity required to derive the categories. Researchers are required to make "judgment calls" as to when a group name is similar enough to another that they count as a one category. Second, the method becomes time-consuming and extremely tedious when the number of items or participants is large. Third, the method examines the relationship of items to categories rather than items to other items. It does

not provide information as to how items could be arranged *within* a category. Thus, we supplemented the manual coding with the statistical techniques discussed in the next section.

4.2 Cluster Analysis

Cluster analysis is a data exploration technique that algorithmically groups items based on pairwise similarities (Aldenderfer & Blashfield, 1984). This means that the method examines each pair of items individually and determines how close or far apart they are. To do this, an association matrix is constructed. An association matrix consists of all items listed in both the rows and columns. Each combination of items receives a score. The diagonal of the matrix will be each item paired with itself and will represent the highest similarity (or least dissimilarity) possible in the matrix.

The cluster analysis method uses the association matrix to find the arrangement of items that best accounts for the distances between all the pairs. What results is a diagram containing all the items grouped into one or more clusters. Clusters are often depicted on a diagram known as a dendritic tree diagram or dendrogram, which represents the overall similarities among items as branches on a tree. Pairs that are more similar are located closer on the tree and have short branches (line lengths) between them. Items that are less similar are located farther apart and have long branches. The dendrogram can then be used to determine categories by selecting the major branches of the tree.

For the portal project, we constructed a 95 x 95 association matrix for each participant. Each cell of the matrix recorded whether or not the participant placed the items into the same group, regardless of what that group was called. If the participant grouped the items together, we recorded 1. If the participant did not group the items together, we recorded 0. For example, a participant placed the items *411 Directory* and *Yellow Pages* into one group that he called "Find Someone." We recorded 1 in the cell for pair [*411 Directory* / *Yellow Pages*]. However, the participant did not include *CAS-LDR* in the "Find Someone" group so we recorded 0 for the [*CAS-LDR* / *411 Directory*] and [*CAS-LDR* / *Yellow Pages*] cells.

We calculated an overall association matrix by summing the values in each cell across all participants. This provided a similarity score between 0 (no participants grouped the items together) and 9 (all participants grouped the items together) for each pair. Because some statistical packages require a particular input method, we inverted the similarity scores to provide a measure of dissimilarity or "distance" between the items (dissimilarity = 9 minus similarity). For example, items *Alternate Work Schedule* and *CAS/LDR* have a dissimilarity score of 5, indicating that they are moderately close together. This makes sense because both are related to time and attendance. However, *Badges Authorization Forms* and *Center Clubs & Recreation* have a dissimilarity score of 9, indicating that these are maximally distant concepts in the participants' mental models. Figure 2 shows an example from the overall association matrix.

Card Name	*411 Directory*	*Acronym list*	*Alternate Work Schedule*	*Badge Authorization Forms*	*CAS/LDR*	*Center Clubs & Recreation*
411 Directory	0	7	9	9	8	8
Acronym List	7	0	8	8	9	9
Alternate Work Schedule	9	8	0	7	5	9
Badge Authorization Forms	9	8	7	0	9	9
CAS/LDR	8	9	5	9	0	9
Center Clubs & Recreation	8	9	9	9	9	0

Figure 2. Example from the overall association matrix.

We analyzed the overall association matrix using the Cluster Analysis-Joining/Tree Clustering procedure in Statistica (Statsoft, 2004). Figure 3 shows a sample of the dendrogram for the portal project. The large branches of the tree related well to the categories derived by hand, even when the precise members of the category differed somewhat.

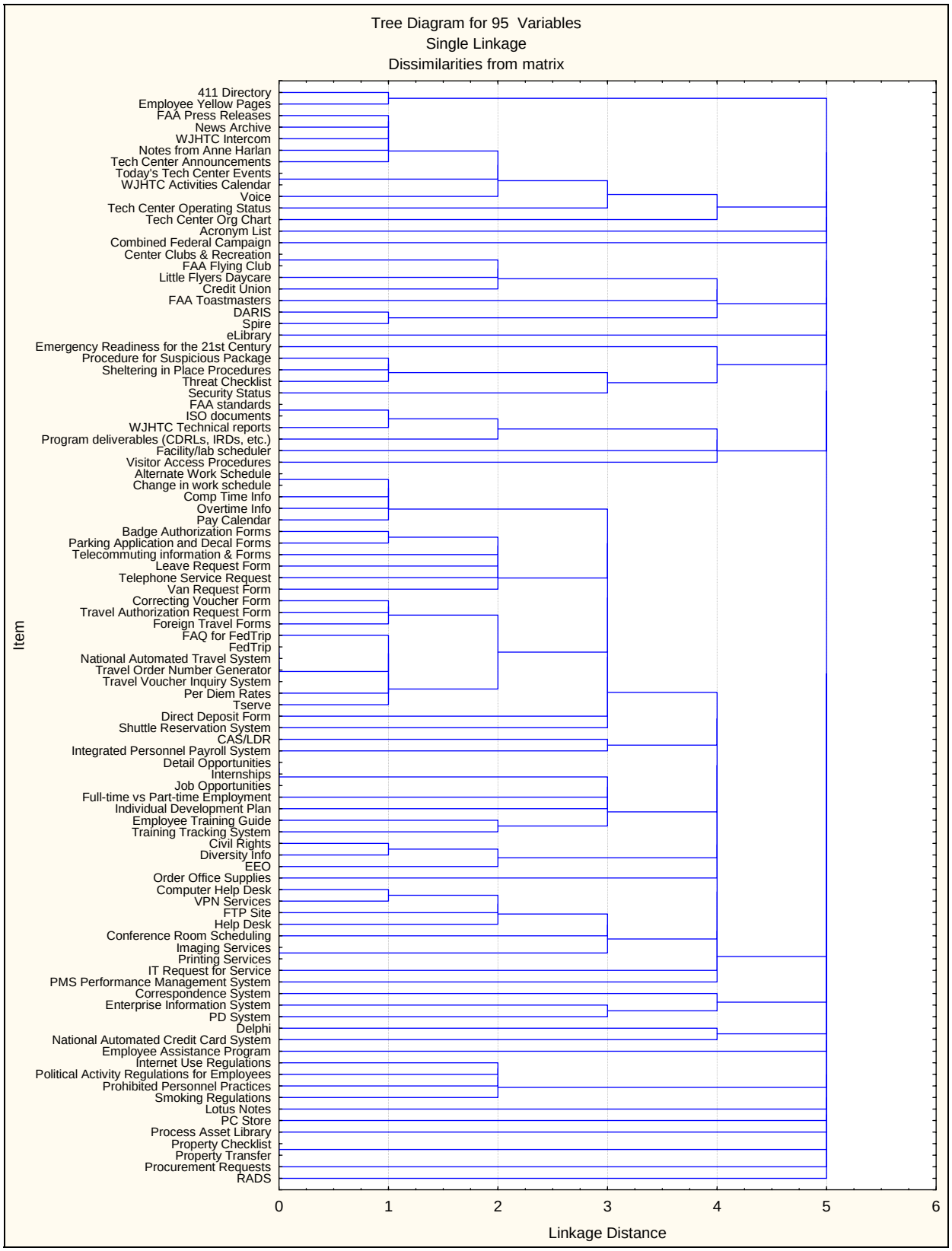


Figure 3. Sample of the dendrogram from the portal project.

For example, the participants frequently grouped items *FAA Press Release*, *News Archive*, *WJHTC Intercom*, *Notes from Anne Harlan*, and *Tech Center Announcements* together and these have low dissimilarity scores with each other. This is reflected in the dendrogram by showing all the items grouped next to each other in the upper left corner, connected by short branches. Participants grouped items *Today's Tech Center Events*, *WJHTC Activities Calendar*, and *Voice* nearby also but with longer branches. This indicates that the second three items are similar to the first four but not as similar as the first four are to each other. By moving along the branches of the dendrogram, we can see that these seven items fall into a larger categories relating to news and announcements.

4.3 Factor Analysis

Factor analysis is another statistical technique that determines underlying structures of data. The analysis looks for patterns in the data to identify one or more factors. Conceptually, a factor in factor analysis is very similar to a cluster or a category in the other methods. Items "load on" or relate to each factor by varying amounts. If an item loads highly on a factor, it is closely related to the factor. If the item does not load well, it is not closely related (or there is too little information to determine). Metrics known as eigenvalues measure how well each factor explains the data. A researcher typically sets an eigenvalue criterion before running the analysis to specify what is a worthwhile factor. Traditionally, an eigenvalue greater than 1.00 is considered worthy of further consideration as a factor.

The statistics underlying factor analysis require large sample sizes especially with large numbers of items. The factor analysis we conducted for the portal project, with only nine participants and 95 items, does not meet the sample size requirements for proper factor analysis. However, to continue discussion of the example and to demonstrate the output of this method, we conducted a principal components factor analysis with Varimax rotation nonetheless. The analysis identified 15 factors with eigenvalues above 1.00. We assigned general names to each of the identified factors (see Table 1). The categories identified in the factor analysis correspond well to the large branches of the dendrogram and the categories identified manually.

Table 1. Sample of Factors Identified from Principal Components Analysis

Factor	Provisional Name	Top Loaded Items	Eigenvalue
1.	"News"	*WJHTC Intercom, Notes from Anne Harlan, FAA Press Releases*	15.3
2.	"Services"	*Help Desk, Computer Help Desk, Virtual Private Network (VPN) Services*	13.8
3.	"Travel"	*Travel Voucher Inquiry System, Travel Order Number Generator, FedTrip*	11.2
4.	"Forms"	*Parking Application and Decal Forms, Badge Authorization Forms, Telephone Service Request*	9.7

4.4 Naming Categories

Cluster analysis and factor analysis can be used to determine which items belong together. However, there is no statistical method to determine what those groups should be called. Developers must use their expertise and established human factors guidelines for the development of labels and titles. See Chapters 6 and 8 of the Human Factors Design Standard (Ahlstrom & Longo, 2003) for guidelines on developing effective labels and menus.

For example, in the portal project, the factor analysis identified 15 factors but this is too many to use as major sections of a portal because of user abilities to remember and distinguish categories and technical considerations like screen real estate. Because we asked participants to provide category names, we already had a good corpus of names. By combining the results of the three methods, we developed the following categories. Each category contains 2-12 items. In addition, some of these names are too long to fit on a regular sized push button or menu. The labels can be further optimized through iterative design and testing. Sample category names derived using the combined techniques discussed here and then refined through iterative reviews and testing:

- Clubs, activities, and employee organizations,
- Collaboration,
- Directory and yellow pages,
- Human Resources/Training,
- News & Announcements,
- Library / Reference Services,
- Policy, Procedures & Regulations,
- Management systems,
- Security,

- Service & Support,
- Pay and Benefits, and
- Travel.

This number of categories is still large for use as major sections of a portal and could be combined further if necessary. The lengths of branches in the cluster analysis yield clues as to the distance between categories. For example, the "Pay and Benefits" and "Human Resources/Training" categories are located on different but adjacent branches in the dendrogram. If space is needed, these categories could be merged into a larger category called "Employee Services" or "Personnel."

In retrospect, we could have limited the number of categories that the users were allowed to create, forcing the users to create a smaller number of broader categories. Another alternative is to bring the users back together to come up with agreed-upon names for the groups. We used the cluster analysis to help combine categories with the understanding that if additional refinement is needed, it would be identified during the usability tests.

We identified locations for several items, such as *Civil Rights*, that did not fit reliably into one category. For these items, we recommended that links be established in several likely locations to increase the chances that users will locate the item. In the case of *Civil Rights*, we established it as part of the "Human Resources/Training" category but also cross referenced it in "Policy, Procedures, & Regulations."

By having the users perform the card sort individually, we were able to look at the areas where categories differed between individuals. Examining these differences showed which content items were not understood, where content may need to be located in more than one area to accommodate different users and alternative paths to the information, and how different users see the information items.

5. CONCLUSION

The card sort method and the associated analyses can help build taxonomies for the organization of the information in a system. By basing the taxonomy on data collected from the users rather than developing it without user input, it is more likely to match the mental model of the users. As a result of matching the organization of data to the users' mental models, the user will be able to find information more quickly and efficiently.

We found a high level of similarity between the statistical and more subjective methods that we used to create the taxonomy in the portal project. Although the different methods that we used provided information on what items to put in the different categories, we based the titles of the categories on our best judgment and the input from our participants, with the understanding that this is part of an iterative process. The card sort itself is one input to information design but cannot create the final structure by itself. The benefit of the card sort is that it can identify general trends so that the organization will better fit the user mental models. The card sort must be part of an iterative process that includes usability testing.

One of the issues that arise with the organization of information based on the card sort is how to handle differences between individuals. For example, one of the users categorized travel forms

and benefits forms into a category called “Forms,” whereas many others put the travel form in a “Travel” category and the benefits form in a “Human Resources/Training” category. These individual differences may reflect differences in the mental models of the users or differences in the way they use the information. Our recommendation for items that are categorized inconsistently is to provide links to these items in multiple places and to verify the location of the information based on results of usability testing.

References

Ahlstrom, V., & Longo, K. (2003). *Human factors design standard: For acquisition of commercial-off-the-shelf subsystems, non-developmental items, and developmental systems*. (DOT/FAA/CT 03/05). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Aldenderfer, M. & Blashfield, R. (1984). *Cluster analysis*. Thousand Oaks, CA: Sage publications.

Department of Health and Human Services. (2004). *Methods for designing usable websites*. Retrieved April 30, 2004, from http://usability.gov/methods/data_collection.html

National Library of Medicine. (2004). *Fact Sheet: Medical Subject Headings* (MeSH). Retrieved September 22, 2004, from http://www.nlm.nih.gov/pubs/factsheets/mesh.html

Nielsen, J., & Sano, D. (1994). SunWeb: User interface design for Sun Sicrosystem's internal web. Proceedings of the Second WWW conference. Available on the internet at http://archive.ncsa.uiuc.edu/sdg/it94/proceedings/hci/nielsen/sunweb.html

Statsoft. (2004). Statistica (Version 6.1) [Computer software]. Tulsa, OK: Statsoft Inc.

Appendix A
The Organization Of Information Items For An Employee Portal - A Card Sorting Task

Background

The William J. Hughes Technical Center is developing an Enterprise Information Portal (EIP) to provide the employees with improved access to information and services. The information portal is proposed as a means of providing a single source for frequently used resources and a means of knowledge sharing.

Purpose

This goal of this study is to investigate how users classify the information into taxonomies. This information will be used to develop a taxonomy for the information that approximates the taxonomies developed by the users.

Participants

As different users across the WJHTC have different tasks and functional requirements, it is expected that the taxonomies may vary. Therefore, participants represent a cross section of the intended user group.

Procedure

The time requirement for the task is approximately 60 minutes. You will be given approximately 90 cards. Your task is to sort them into categorical groups and develop a heading for each group. There may be some items that you are unfamiliar with. You can put these items into a heading called "I don't know". Although there may be more than one category that a particular item may fit into, we are asking you to choose a single category for each item. When you have developed a category for a group, please write down the category on the paper provided and keep it with the index cards that you have designated as going into that category.

Rights of Participants

Participation in this study is strictly voluntary and the privacy of participants will be protected. No individual names or identities will be recorded or released in any reports. Strict adherence to all Federal, Union, and ethical guidelines will be maintained throughout the study.

Point of Contact

Your support is important to the success of this project, and your cooperation will be greatly appreciated. If you have any additional questions, then please do not hesitate to contact:

Deborah Germak, Office of Knowledge Management

(609) 485-9862