

Transforming Data Into Information: The Development and Demonstration of a Model to Support Transportation Planning

by

David P. Racca

**Center for Applied Demography & Survey Research
College of Human Services, Education and Public Policy
University of Delaware**

April 2005

**Delaware Center for Transportation
University of Delaware
355 DuPont Hall
Newark, Delaware 19716
(302) 831-1446**



Transforming Data Into Information:
**The Development and Demonstration of a Model to Support
Transportation Planning**

by

DAVID P. RACCA

**Center for Applied Demography and Survey Research
College of Human Services, Education, and Public Policy
University of Delaware
Newark, Delaware 19716**

**DELAWARE CENTER FOR TRANSPORTATION
University of Delaware
Newark, Delaware 19716**

This work was sponsored by the Delaware Center for Transportation and was prepared in cooperation with the Delaware Department of Transportation. The contents of this report reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views of the Delaware Center for Transportation or the Delaware Department of Transportation at the time of publication. This report does not constitute a standard, specification, or regulation.

The Delaware Center for Transportation is a university-wide multi-disciplinary research unit reporting to the Chair of the Department of Civil and Environmental Engineering, and is co-sponsored by the University of Delaware and the Delaware Department of Transportation.

DCT Staff

Ardeshir Faghri
Director

Jerome Lewis
Associate Director

Wanda L. Taylor
Assistant to the Director

Lawrence H. Klepner
T² Program Coordinator

Sandi Wolfe
Secretary

DCT Policy Council

Robert Taylor, Co-Chair
Chief Engineer, Delaware Department of Transportation

Eric Kaler, Co-Chair
Dean, College of Engineering

The Honorable Tony DeLuca
Chair, Delaware Senate Transportation Committee

The Honorable Richard Cathcart
Chair, Delaware House of Representatives Transportation Committee

Timothy K. Barnekov
Dean, College of Human Resources, Education and Public Policy

Michael J. Chajes
Chair, Civil and Environmental Engineering

Ralph A. Reeb
Director of Planning, Delaware Department of Transportation

Stephen Kingsberry
Director, Delaware Transit Corporation

Shannon Marchman
Representative of the Director of the Delaware Development Office

Roger Roy
Representative, Transportation Management Association

Jim Johnson
Executive Director, Delaware River & Bay Authority

*Delaware Center for Transportation
University of Delaware
Newark, DE 19716
(302) 831-1446*

**Transforming Data Into Information:
The Development and Demonstration
Of A Model To Support
Transportation Planning**

prepared for

Delaware Center For Transportation

by

David P. Racca

Center for Applied Demography & Survey Research

College of Human Services, Education, and Public Policy
University of Delaware
Newark, DE 19716

April 2005

TABLE OF CONTENTS

Project Summary	1
DelDOT Division of Planning and Information Systems.....	4
Key Issues To Be Addressed To Improve Planning Information Systems	7
Creation of an Information Utility To Support Planning	10
Overview Of DUROS	17
Organizational Frameworks For Transportation Data.....	28
Further Implementation of the Interface.....	33
Conclusion.....	34

LIST OF FIGURES

Figure		Page
1	Examples of Planning Questions	6
2	Examples of data used by transportation planning groups	7
3	Voice Commands	15
4	Tool Bar	17
5	Server and Place Listings	18
6	Full Documentation Screen	19
7	New File Screen	20
8	File Listing	21
9	Directory Listing By Size	22
10	Organizational Area Subdirectory Listing	23
11	File Listing In An Organizational Area	24
12	Places/Directories In An Organizational Area	25
13	Project View	25
14	Advanced Search Page	26
15	Results of a Search	27
16	Project Pools in the Pipeline Process	28
17	Impact Categories	29
18	Examples of Important Planning Measures	29
19	Transportation Data Organization	32

Project Summary

A previous study* of information systems at the Division of Planning at the Delaware Department of Transportation (DelDOT) was conducted in response to concerns about data management issues and what could be done to improve information systems. Some of the findings of that study were:

- Improvement is needed to meet information needs that are less predictable, that involve bringing many types of information together from many sources.
- There was a need to be able to more easily find information.
- Easier access to information was desired.
- There was a need for better information integration. Often transportation data used by the Division of Planning was combined from many sources and there were some difficulties in relating the information.
- There was an interest in determining an appropriate direction for further development of geographical information systems to meet analysis and presentation needs.

The initial focus of this project was to outline an organization for transportation data that considered planning activities, data processes, data relationships, and integration so that information resources could be structured in a way that was more responsive to the Division of Planning's needs. Such an approach was focused on data modeling and how data could be related. It was perhaps more of a traditional approach focused on how data could be structured in some type of data repository or Oracle enterprise database in a way that information could be better integrated and utilities could be designed to assist planners with their analysis.

As the project proceeded, representatives from the Division of Planning expressed where the primary focus of the research might best be directed. The Division's use of information systems most often starts with a request for information, or a problem or issue that must be addressed. From there, staff must collect and find necessary information, prepare and analyze the data, and then present it. This day after day "Find → Use → Present" process is much of the focus of the Division of Planning's use of information systems, and anything that can assist at any stage of this process can make a significant improvement. The Division of Planning requested a focus on anything that could be done that would make this process easier or more efficient.

While the formulation of a model for transportation data is necessary for a large scale data warehouse development or implementation of an enterprise database, thoughts turned more to what might help the Division in the near term, what would focus on the specific needs and the ways planning uses information systems, and what would have a good chance of being successfully implemented. Through a rethinking of the problem the following issues/facts emerged:

* Racca, David P., "Directions in Transportation Data Management", prepared for the Delaware Transportation Institute (now Delaware Center for Transportation), February 2000

- Many sources, many types of data, and many types of software – the Division of Planning deals with many types of information, in many data formats, that include reports, presentations, data tables, text documents, pictures, maps, and others that actually are difficult and time consuming to integrate. Being bits of varied information in varied formats, much of the information couldn't be input to an enterprise database. The best that can be done with most of it is to be organized and referenced.
- Planners and analysts only have so much time that can be devoted to a potential improvement to information systems. – Staff in the Division of Planning have been very open and cooperative over the years to efforts to improve information systems, but they are pressed on a day to day basis and can't invest much more time than they do now.
- The nature of how planning uses information systems is decentralized. - Planning tasks are distributed among many professionals and analysts, using various information technologies, each addressing various aspects of projects and issues. Data is spread out over the division and managed by various staff on numerous personal computers. Planning staff need independence and diversity in the use of information technology. There is a limitation to how much a centralized agency wide system can help and how much information resources staff can assist on a day to day basis in such diverse, specialized, and unpredictable needs.
- It is hard to predict what information will be needed at any given time and preparing information beforehand may be a waste of time – There is a temptation to imagine that a sufficient body of information could be collected in a suitable organization that would make everything easier in the future. The issues facing the Division of Planning however are very diverse. The planning process must not only have a strong grasp of all relevant transportation issues and resources but it must also be sensitive to the context in which problems exist with respect to natural resources, cultural resources, regulations, political initiatives, and many other factors. Much of the analysis that is conducted could be considered a “one-off”, a specialized, quick look at a specific problem existing at only one time.
- New information is constantly being generated – Any information resource must be dynamic and maintained and growing on a day-to-day basis.
- Planners and analysts must be intimately involved in the organization and upkeep of their information resources – There are some sensible ways to organize transportation data but there are many ways that planners can look at problems. Planners must drive how their information is managed and must have freedom to view information in a variety of ways. The maintenance of any information resource must be part of the day to day operations and done by the users.

These points above would argue against large scale efforts to collect, integrate, and/or reorganize data into a large centralized system. Certainly agency wide standards are necessary. It has been shown that some difficulties in integrating information can be attributed to various methods used within the agency to identify transportation features and their location. There must be some standards and guidelines for use and acquisition of computer hardware and networks.

In light of the above points the decision was made to focus on and demonstrate a facility that would

- improve data organization
- offer tools to find information
- preserve the massive amounts of diverse information being generated on a yearly basis
- provide easy documentation tools
- bring together the necessarily decentralized information systems
- get the users of information more involved in the management of their information

In this project a functional prototype of web based documentation, search, cataloging, and organizational tool was created to demonstrate a potentially powerful aid to the Division of Planning. For purposes of discussion this utility will be termed the DUROS, the Documentation Utility for Referencing, Organization, and Search. The DUROS has the following features:

- Easy documentation tools. If data isn't documented or cataloged it can't be searched effectively. If documentation is not easy, it won't get done.
- Ability to organize data in numerous ways without copying or moving data
- Easily created customized views of files, directories, projects, and organizational areas. The system goes far beyond the limitations of organization by physical data folders.
- Availability of the tool from any internet browser on a network
- Data management utilities
- Fast search utilities that can operate across several data servers or other areas specified by the user. Searches can be assigned as views or organizational areas for late reference.
- Documentation at the time of data creation

While many other features than those demonstrated could be built into the DUROSS it represents a simple but powerful utility that could be developed and implemented in the near term with relatively low cost when compared to large scale data warehouse efforts.

Division of Planning and Information Systems

DelDOT Division of Planning and Information Systems

A good starting point is the Division of Planning's mission, which is as follows:

“The mission of the Division of Planning is to provide comprehensive transportation planning and project development services to address the mobility needs of Delaware residents, as well as visitors to the state, by providing a safe, efficient, multi-modal, and environmentally sensitive transportation system that conforms with the goals and objectives of Livable Delaware”.*

Key objectives of the Division are:

- Promote mobility for people and goods by working with our customers to create plans that will result in a comprehensive system of transportation options in coordination with Livable Delaware goals and objectives, the State Investment Strategies, and the county Comprehensive Plans.
- Define and solve transportation problems through a planning process that is sensitive to the context in which the problem exists and respects natural resources, cultural resources, and the wishes of the affected community while complying with all applicable laws and regulations.
- Provide transportation information and advice to local governments with land use decision-making responsibilities to help coordinate zoning, sub-division, and annexation decisions among state agencies, counties, and municipalities.*

The mission and the objectives themselves speak to a comprehensive, broad view toward the planning of projects and services. Transportation problems are to be identified and solved with consideration of many factors. Mobility, safety, efficiency, and environmental sensitivity, are just a few of the areas that information systems must address.

Results of Previous Studies

A previous study of information issues in the Division of Planning indicated that improvements were needed in four primary areas for: better data integration, the ability to better find information, the ability to more easily access information, and determining direction of further development of geographical information systems (GIS) to meet analysis and presentation needs. Naturally, the focus on projects requiring collection and synthesis of large amounts of varied information from many sources would require the

* from DelDOT 2003 operating budget documents.

best access, integration, and presentation. Improvements to information systems for planning must address these needs.

The Question → Find → Process → Present Process

Discussions with DeIDOT planning personnel indicate that much of their time is spent going through a particular cycle:

- Information request on some topic occurs.
- Find the available information.
- Figure out a strategy for using the available information and tools together to answer the question.
- Process/analyze the information.
- Present the information and satisfy the need/request.

Anything that can be done to make any of these steps easier and/or quicker represents a large gain in productivity over the several hundred information requests and queries each year. This goes on and on, and generally requests are very different and the effort can often be considered a “one off”, an exercise that probably won’t be needed again and addresses a unique need. So information systems have to be flexible and cover a broad base. A need for a particular view of particular information cannot be predicted many times or prepackaged. The knowledge and resourcefulness of planners and analysts to make use of available information and employ a variety of information tools in real time is still required.

Use of Information In Planning, Informational versus Operational

Traditional approaches to information systems at DeIDOT have focused primarily on the operational aspects of data processing, and on specific computer programs and applications. Needs that are met by systems such as those for accounting, database maintenance, accidents, the road inventory, project tracking, or the preparation of standard reports for federal compliance, are generally well met by existing systems.

Improvement is needed to meet information needs that are less predictable, that involve bringing information together from many sources, such as the analysis and presentation of transportation system enhancements, the effects of future growth and distributions of populations, and collection and analysis of information for project development. Planning applications have a greater demand for more informational rather than operational systems. Impacts of projects to the transportation system, the environment, historical resources, land use patterns, and the public in general, must be continually evaluated and presented, requiring a wide range of information and various perspectives and views at the available information. The types of questions that planners must consider are given in figure 1 as an example.

Figure 1. Examples of Planning Questions

What is transit service in this area? Who is using transit?
What are traffic volumes in this area? Where are congested areas?
Where are people going here?
Where are the major attractors? Who's coming into this area?
What are projects in this area?
What is the population, housing, and employment in this area?
What is the future picture of this area? Is it growing?
What are markets for transit? Where will it be most successful?
What is the volume to capacity ratio of a road at a particular point in time?
Where are the congested areas?
What projects are in a particular area and what is the status?
What is the current condition of facilities?
What will be the distribution and character of populations in the future?
What will be traffic volumes and characteristics of the network in the future?
Where is congestion?
What are the data needed for environmental review?
What are the trends in air quality?
What information does the project development group need?
What is the best data strategy for project development?
What are the primary data needed for public presentations?
In what situations and where are carpooling efforts most effective?
What will be the expected impact of a proposed regional shopping center?
Where can detailed information about a particular corridor be found?
What policies have the greatest support from the public?
Where are the portions of the transportation system that have elicited the most complaints?
When and why did the Department accept maintenance on a road?
What is the percentage of "through" vs. "local" trips on a given roadway?
If emissions are decreased in new cars what will be the effect on air quality?
How congested will roads get in the future?
How many roads of a particular type and place are there?
Where are the major trip attractors in New Castle County?

Key Issues To Be Addressed To Improve Planning Information Systems

To address how information systems may be improved for a transportation planning group, it is useful to examine some key issues that need to be addressed

Issue 1) Many sources, many types of data, and many types of software

The Division of Planning must address a wide range of topics and sources of information. Data is available in numerous locations in and outside the agency, and the data is of many different types, and produced by a wide range of software. Integration of information is often very difficult, if possible at all, because the data types and sources can be so varied. A list of common data types used by the Division of Planning is shown in the figure below.

Figure 2. Examples of data used by transportation planning groups

- * DBMS data sources
Examples: DBMS, Oracle tables, MS Access, Vax data, Road Inventory
- * Tabular – non DBMS, usually requiring software
Examples: GIS tables, Exel spreadsheets, vax dump, fixed format or flat text tables, census data.
- * Text data
Examples: Any viewable text document, word processing documents, Adobe PDF, html, e-mail,
- * Graphics
Examples: All photography, images, and scanned data, maps, posters, Flow diagrams, Power Point Presentations, Web images
- * Specialized software formats
Examples: Travel Demand Forecasting Data, SPSS and other statistical packages, transportation network modeling software, federal reporting software
- * HTML and Web resources
Examples: data and images from websites
- * Spatial data
Examples: Geographical Information System Data

There is a wide range of sources for information, including other transportation agencies, federal, state, and local agencies, the media, correspondence, universities, and the private sector.

Issue 2) Centralizing Data For Planning is Practically Impossible

In DelDOT Division of Planning, data exists across several users and personal computers. Given the many “projects”, studies, presentations, inquiries, etc that Planning must address on a day to day basis, there is a tendency towards decentralization and information system capabilities that must address a very wide variety of data that is difficult to structure, predict or prepare for before hand. Data is most often organized by project. Creation of centralized data warehouses or compilations has generally not gotten off the ground over the past years for a few reasons such as:

- Compiling or structuring “all” information that might be used is a massive, expensive, and often wasteful and fruitless task given that it’s hard to predict what information will be needed.
- Planners and analysts are extremely busy, facing day to day demands and deadlines, and can only be involved to a small extent. Information systems personnel, while often highly effective with the development of operational systems and information challenges that can be specified in detail, have understandable difficulty in expending resources to the unpredictable moving target of most informational uses.
- If an effective data repository of some kind is somehow initiated and created through large amounts of effort, it usually will require more effort than planners and analysts are able to provide to maintain and update it.
- If there is any difficulty in the use of a repository or data warehouse which there generally will be, planners and analysts will turn back to their own independent or local means to satisfy the pressing cycle of “find→process→present”. This is how it has always been, answering the question has always come first, and planners by necessity have had to develop systems that work for them in the near term and there has generally been a low interaction with corporate information resources, and not much incentive for using Oracle and corporate DBMS

Issue 3) Planning personnel need a level of independence in the information systems they employ.

Planners are heavy information users. They generally have expertise in an area of planning or transportation, and are primarily not from an information science background. However, many are analysts who are expert with information systems, particularly those that assist in their area of expertise. Expertise with information systems is a very desirable and necessary part of a planner’s or analyst’s work. A central information resources department often would neither have the immediate knowledge nor the resources in general to deal with the very specifics of an analysis or specialty application. Often information challenges are more of an organizational or production nature rather than a strictly technical issue. Pressures and mission has required an independence from central IS group, though certainly the Division of Planning would need to be involved with and adhere to basic standards

Issue 4) The environment of decentralization

The result of a necessarily distributed approach is a very common situation where there is a collection of professionals each with their own expertise, maintaining data in folders (directories) on separate personal computers, often using different software and data analysis tools, with no standard naming or organizational scheme. Where and what the information is, is typically only understood by individuals who created the information, and there is very little if any documentation available. When someone needs information they ask around or they are familiar with someone in their immediate or extended vicinity that they know may have what they are looking for. When the project is over no one is generally going to take or have the time to wrap it up and document it neatly.

There is no librarian available who will take the information and integrate it into a centralized system and even if there were they would have to go to each user on a continual basis to extract all of the old and new data. Some agencies are blessed with an unusually organized person who structures and documents everything they do (or at least one can imagine such a person). More often the best hope is that if a person leaves the organization, somehow the knowledge of the information they managed can be preserved.

Its not pretty, but this is basically the way things stand, and it's not like it hasn't worked over the years. Everyone can appreciate the idea of a system where information can be found more easily, and where things were a bit more organized, but no one seems to know how that might practically happen.

Issue 5) Information users must have the primary responsibility for the organization of their information resources

The organization of the information has to be what is most natural for the creator/user and not have difficult requirements to which they must conform. Organization of information could be by any number of themes, by project, by individual, or other topic.

Issue 6) Sufficient Tools are not available for Planning staff to document, organize, or find information

To effectively find information it must be documented and organized by the users. Nobody has time to document or catalog. While there are now many types of search technologies available, they all depend on having something to search, and ultimately depend on some form of documentation and transparent organization. At present there appears to be no powerful tools that the Planning staff can employ to organize or document their data in an efficient way. Any such tools would have to be the easiest to use and require the least attention of the users, or they would not be used.

Creation Of An Information Utility to Support Planning

The development of data warehouses and repositories is a typical approach toward making data easier to find and more accessible. Thoughts are that if we can just collect “all the data”, and organize it some way that makes sense, then various interfaces can be created to allow users to interact with it. Organizational frameworks for transportation data were studied in this project* and relationships between the various data were examined toward coming up with a method of handling information that would make finding data and integrating it easier. In an operational system, this approach with its entity relationship diagrams and neat categories might work well. In the case of planning data though, the primary feature is the great diversity of information and its form. There in general is no way to predict what information will be needed on a given day. Appropriate organizational topics can be suggested but not reliably predicted in general. If a massive organizational or data collection effort is initiated it is likely that only a fraction of the collection will be useful and more likely that it will quickly get stale, unattended, and forgotten.

The primary need is to be able to reference and search data as easily as possible, and as part of that be able to document data. To organize, but to be able to easily reorganize. What is needed is a catalog and organizational system that is user driven.

Features and capabilities of and effective information system for planning

Keeping in mind the Division of Planning’s mission and goals, and summarize previous discussion and move toward the next step, this section identifies the features and capabilities of an effective information system utility for planning.

- The system must work well with a range of software that planners use.
- Documentation is vital but the documentation process must be as easy as possible.
- Files must be referenced and accessed across a network of servers. Generally planners work in a personal computer based environment.
- A physical data file could be associated with more than one project, theme, or organizational area. The system should be able to create views and organizational areas without copying physical files. The standard process of everyone organizing data in a physical folder is a limitation.
- A utility should be able to be accessed from any computer in a network. It was decided to make the catalog web based and accessible for anyone who had a browser.
- The system can define users, groups, and administrators. The catalog will look different depending on who is logged into the system.
- Search utilities must be incorporated into the catalog. The results of searches should be easily updated, organized or viewed at other times.

* See Appendix B

- Cataloging must be an easy process that users can employ at the very start of creating data.
- Users should have the ability to open files directly from the catalog.

A search was done for other programs that had been already written to serve the purpose. Software was found that catalogued file systems but did not have easy tools for documentation of data, or the ability to define other organizations of data (views), or sufficient search tools. Many references for data management were in the area of paperless office applications.

Documentation Utility for Referencing, Organization, and Search (DUROS)

In this project a functional prototype of web based documentation, search, cataloging, and organizational tool was created to demonstrate a potentially powerful aid to the Division of Planning. For purposes of discussion, this utility will be termed the DUROS, the Documentation Utility for Referencing, Organization, and Search. The DUROS has the following features:

- Easy documentation tools. If data isn't documented or cataloged it can't be searched effectively. If documentation is not easy, it won't get done.
- Ability to organize data in numerous ways without copying or moving data
- Easily created customized views of files, directories, projects, and organizational areas. The system goes far beyond the limitations of organization by physical data folders.
- Availability of the tool from any internet browser on a network
- Data management utilities
- Fast search utilities that can operate across several data servers or other areas specified by the user. Searches can be assigned as views or organizational areas for late reference.
- Documentation at the time of data creation

Software chosen for the Development and Deployment of the Catalog

The software components used in the demonstration catalog are:

Web server and database:

- Microsoft Web - Web page development and access
- Coldfusion Server – web site database management and query
- Oracle Version 8 – database system

File system reading and import to databases

- ARC Catalog – import of tables into Oracle
- SPSS Professional – processing file system information
- Advanced File Organizer (Softprime Development, www.softprime.com)

Users can access all features of the catalog through a web browser. This provides a very convenient, inexpensive method to allow access to the catalog from many locations. The application was written in HTML and Cold Fusion was used for database operations.

Implementation of Organizational Areas and Projects

Typically data for an organization is spread across one or more network servers. A full path can be specified for a data file in terms of the server name or share name rather than any drive letter that an individual user will map to a server location. So if data files are stored off the home directory of the “d” drive of a server called “SERVER1”, one way to reference the file would be “ [\\SERVER1\d\\$\home\afile.txt](#) “. Or perhaps a Share on the network (MicroSoft) could be defined as [\\SERVER1\Homefiles](#) “ and the fullpath to the file could be “ [\\SERVER1\Homefiles\afile.txt](#) “. The catalog that was built allows for filepaths to be specified in terms of drive letters (i.e. s:\home\afile.txt) but each user could use different drive letter mappings so the server path is used instead.

The same type of hierarchal naming method used for folders within physical directories is used for organizational areas or project areas. Rather than the starting point being a server it is the name of an organizational area or project area. Organizational areas and project areas, and any type of “view” are all handled in the exact same way. Because of so much of planning is organized by projects, a decision was made to have the catalog address them differently but otherwise there is no difference. Within the database the path name of an organizational area are preceded by [\\org](#) and then the name of the organizational area. For instance, if users wanted to organize along the lines of the transportation framework discussed earlier in this report, the whole area would be called “TransFramework” for instance and virtual sub areas (analogous to sub-directories) could be created for “Supply”, “Demand”, “Performance”, and “Impacts”, and files within them would have paths like [\\org\TransFramework\Performance\trafficcounst.xls](#). The organizational areas would look just like physical directories but the data is not copied to a new physical location, and files can belong to any number of organizational areas or sub-areas. There can be as many levels of organization (i.e. sub-sub-sub-areas) as exist with physical directories. With this type of implementation, physical directories could actually be dispensed with and replaced with large storage areas.

Documentation

There have been numerous major efforts across the country to establish metadata (information about information, documentation) standards and to encourage information system users to document their data. The argument is that the bulk of all information systems costs are in data development, and that without documentation, without knowing what the data is, the information becomes worthless and lost. Some fairly extensive standards for metadata have been developed, particularly for GIS data, which include descriptions of the source, accuracy, history, applicability, and other information. The reality is though that most users do not take the time to specify any information other than what might be gleaned from a often cryptic physical directory or file name. Usually only the creator of the data knows what a particular data file is and often even then the

location of the data and its content is forgotten. Documentation of any kind is generally time consuming and users lack tools that would make the process easier. There are numerous methods and tools that have been developed for searching data collections. Some utilities can search within the content of specific data types (i.e. MS Office files) and this is useful but sometimes very computer and time intensive.

Utilities were built into the catalog to allow users to fully document a particular data file. In developing this catalog it was decided that at least if users could associate a short descriptive with data files that would be better than the physical file name and nothing else. With each directory and each file, users can provide a short description easily in the listings. This is easier than what is provided by MicroSoft Explorer and any file managers that have been found.

Referencing Files From The Start

The best time to describe a data file or directory is when it is first created. Utilities are included in DUROS that work with personal computer software to make this easier.

As Easy As Possible, Voice Recognition Software

Many extensive and expensive initiatives to organize and document data ultimately fail beyond the initial implementation. It is important to be realistic. If data management and documentation tools are not as easy as possible there is less hope of users employing them, and without users being involved there is little chance of success. Voice recognition software (Dragon Naturally Speaking) was incorporated into the web interface as one way to make the use of the catalog easier, especially in regards to documentation. The approach taken in the development of this project has been one of a pessimist. Most users generally hate taking time to document, reference, or organize. The catalog is shooting for the absolute minimal difficulty of use.

Voice recognition software is often employed in dictation applications. Users train the software to recognize their voice and the system gets more accurate the more a user corrects mistakes that occur. Accuracy is influenced by enunciation, noise conditions in the immediate environment, acoustic settings in the software, and the quality of the microphone and computer hardware components. Accuracy has gradually improved over the years. Systems are smarter in that they are more aware of context as in distinguishing for instance whether “pool” or “pull” is the word that is meant. There are always some errors with dictation but in general dictation can be taken at a normal speaking pace followed by a bit of proof reading and correction.

The application of voice recognition for the DUROS is for the most part in executing commands and filling in web forms. Recognition of commands is much more accurate. Dragon Naturally Speaking recognizes sets of commands. There are standard commands built into the system that users would most often use with popular software. Users can also define their own commands. As commands can be associated with certain verbs (actions) the system does quite well in recognizing when a command has been issued and

determining the correct command from previously defined lists. Examples of commands might be “Start Word”, “Reference a File”, “Do Advanced Search” , and “Open File”. Any voice command can be associated with a few or many keyboard strokes. Voice commands can build on other voice commands. Web pages can be navigated using the “accesskey” option in and HTML Forms Element that is accessible from the keyboard using the “Alt” key with the specified access key. Users can navigate the web forms of the catalog by associating commands with the access keys associated with elements on the web page. Some dictation is of course part of filling in the information about the file particularly for a long description of the data. The thought was that where long or even short descriptions were added, it was better to have a description with an occasional typographical error or missed word than no description at all.

The addition of the voice recognition feature was to add ease and an element of fun to the documentation and referencing process. As full documentation is so difficult to get users to do, it was thought that users might be more likely to say a long description than to type it. It is the basic viewpoint in this project that it is extremely difficult to get users to reference and document their information and that if the process is not as easy as possible there is little hope of it happening without some type of major coercion. DUROS can be operated using a mouse and a keyboard, but voice recognition was employed as an added feature toward as “easy as possible”. For the creation of the voice recognition feature it was necessary to come up with an easily remembered set of commands that would be needed. Often the command was the same as the text that preceded a form entry or a button on the web page. Figure 3 gives some examples of commands that were created.

The Update Process

Advanced File Organizer was used to read all server files and directories. The delimited export text file is then processed in SPSS. The existing file and directory databases are compared with the update. Attributes in the database that include all documentation previously entered, are preserved. Files in the update that are not in the existing catalog are identified. Files and directories that are not in the update that were in the catalog are identified. A completely updated file and directory listing with all documentation is generated and imported into Oracle. At this time ARCGIS is used to import the newly generated listings into Oracle, this could be modified using SQL loader. SPSS and ARCGIS were selected primarily for ease of use during the development process. Many of the update operations could be done using a programming language or by another method.

Figure 3, Voice Commands

```
Applications
ArcView
ArcMap
START SPSS

General
SHOW MORE
NEXT FIELD
GO TO SAVE IT
CLEAR ENTRY
REFRESH

Internet Explorer
REFRESH
GO BACK (previous site, back)
GO FORWARD
CLEAR ENTRY

OPEN FILE (set the file view)
GRAB (numbered entry)
SHOW MORE

CHANGE VIEW
SET CURRENT FILE VIEW
SEARCH BY WORD
SEARCH BY FILENAME
SEARCH BY PLACE
SEARCH BY FILE TYPE
SEARCH BY PERSON
CHANGE AND RETURN

FILE REFERENCE
ENTER FILE NAME
ENTER A PROJECT
ENTER A PLACE
ENTER A NEW PLACE
ENTER TOPICS
ENTER A NEW TOPIC
ADD FILE DESCRIPTION
FILE TYPE
FILE SECURITY
CLEAR FORM
INSERT DATA
```

Search utilities

Users can search on file names, descriptions, or any document field. Collections can be created and indexed to allow search within certain document types. Searches can occur across all servers referenced in the catalog. Results of searches can be associated with organizational areas.

Server Management Benefits

The catalog provides a fast method of viewing server information. In MS Explorer directory size is shown as the sum of all files within a directory but not including data in subdirectories. The size shown in the catalog is the size of all files and subdirectories

underneath the particular directory so that is easy to see how storage is divided on the server.

Other Features

The DUROS utility needs to contain various file maintenance and management operations. If physical files are moved, or if locations of directories or servers are changed, DUROS needs to know about it. Organizational area and project references must be changed if the data they reference is moved. Basic file and directory operations such as renaming, moving, and deleting need to be accomplished using DUROS instead of outside of the system (i.e. Window Explore, My Computer.)

For each user, DUROS will list recent files that have been opened. Recent places visited (directories, projects, organizational areas) can also be listed.

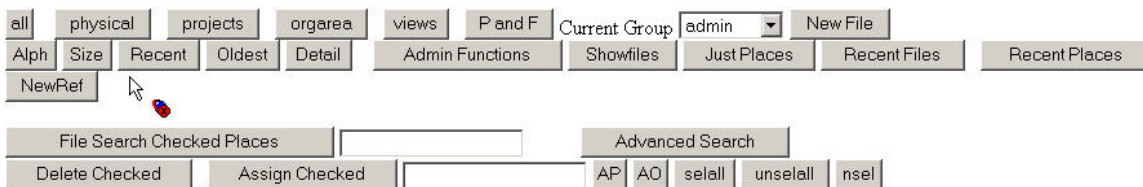
Users, Groups, and access to data has to be defined, and there are administrative tools in DUROS for this.

OVERVIEW OF DUROS

Tool and Button Bar

Each page contains buttons and tools, as shown in the figure below, to perform various functions in the interface. The buttons can alter the type of data that is displayed and the ways listings can be ordered. A checkbox is next to every listing and checked data items can be associated with organizational areas or projects. Checked items can be searched in a variety of ways. Recent places or files visited by the current user can be listed. Administration functions and other utilities can be accessed. The development has focused on the function of the interface, and tool bars will be refined later to provide a better graphic that takes up less space on the page.

Figure 4. Tool Bar



Select a place:	DESCRIPTIVE NAME		
1:dtinterface	Catalog Main Directory	DOC	<input type="checkbox"/>
2:personal	Personal data	DOC	<input type="checkbox"/>
3:dti2002	2002 DTC Proposals, Modeling and Mode C	DOC	<input type="checkbox"/>
4:ocr	Optical Character Recognition Temp Direct	DOC	<input type="checkbox"/>
5:avltransit	Dart AVL Evaluation	DOC	<input type="checkbox"/>
6:filearchive	Data Interface backup and documentation	DOC	<input type="checkbox"/>
7:modechoice	Mode Choice DCT Project	DOC	<input type="checkbox"/>
8:intercon	DPI, O'Donnal, Neighborhood Interconnectiv	DOC	<input type="checkbox"/>
9:Places	Delete, outdated	DOC	<input type="checkbox"/>
10:bikevalues	DCT Bike Path and Land Values Project	DOC	<input type="checkbox"/>
11:dema	DEMA proposals, hazard vulnerability, penii	DOC	<input type="checkbox"/>
12:hazmatrans	DEMA Hazardous Material Transportation f	DOC	<input type="checkbox"/>
13:archive	Archive of numerous older projects	DOC	<input type="checkbox"/>
14:nccgis	New Castle County GIS Projects	DOC	<input type="checkbox"/>

Data Lists

Through a web based form, all data within all intranet servers can be listed, and be ordered alphabetically, by date, or by size. Figure 5, below is a portion of the opening screen for Administrators. It shows all of the servers(disks) in the catalog. It also show organizational areas with paths starting with [\org](#) and project listing with paths starting with [\prj](#) . Organizational areas, project areas, and physical directories on servers are all considered types of “places”. Clicking on any of the places in the first column will always take the user to the next level within the particular place similar to selecting a folder in MS Windows Explorer. Organizational areas and project areas have a tree structure just like physical directory folders, and the interface allows users to search within the “place” trees. The “DOC” Button opens a window to more fully document the place or data file, including the ability to associate key words, security, data type, and long descriptions (see Figure 6). The checkbox allows for operations, such as searches, within selected places.

Figure 5. Server and Place Listings

Select a place:	DESCRIPTIVE NAME	DOC?	
1:\org\modgrids	Modified Grid Area	DOC	<input type="checkbox"/>
2:\org\transdata2	Transportation Data Organization	DOC	<input type="checkbox"/>
3:\oracle\cadsgis	Oracle Database Listing	DOC	<input type="checkbox"/>
4:\lucky1\d\$	Lucky1 D Server, Oracle, power users, SDE	DOC	<input type="checkbox"/>
5:\moses4\d\$	Moses4 D Server, student file spaces and c	DOC	<input type="checkbox"/>
6:\marley1\d\$	Marley1 D Server, CADSR2 primary, user s	DOC	<input type="checkbox"/>
7:\moses5\e\$	Moses5 E Server, CEEP, CHAE, power use	DOC	<input type="checkbox"/>
8:\murphy3\c\$	Murphy 3 C Server, CADSR data	DOC	<input type="checkbox"/>
9:\eddie1\d\$	Eddie1 D Server, CADSR archives, GIS, pri	DOC	<input type="checkbox"/>
10:\moses5\d\$	Moses5 D Server, CADSR, CEEP, CHAE,	DOC	<input type="checkbox"/>
11:\prj\racca	Racca preferred place and file view	DOC	<input type="checkbox"/>

First Plan of Attack, Short Descriptors

One of the easiest, most useful actions users could take to document their data is to supply a brief description of each data file as shown in the column under “Descriptive Name” in Figure 5 above. The DUROS web listing provides a line for a short description of each physical file, directory, project, and organizational area. Users can type in a new description or edit the one shown and the description is immediately input and saved to the interface databases. If users know what their data is, entering short descriptions for listings of places can proceed very quickly. If a particular place is

unfamiliar, users can search within its subdirectories or sublevels to determine appropriate descriptions for contents. This basic brief descriptor represents the minimum documentation of a data place or file. The interface provides an easy tool and entry method to get some information about data in the system. The description can be changed easily at any time but only has to be defined once for a file or directory. Any other organizational area or view will include an information item's description if one has been entered.

Full documentation

Of course it would be nice and very helpful to have more information about data and the "DOC" button next to the listing of each place and file can be selected to associate additional metadata. An example of the documentation page for a file is shown below for a SPSS (statistical software) table that includes six years of travel survey data for Delaware.

Figure 6. Full Documentation Screen

Enter File Name (descriptive):

Physical Location: \\marley1\vd\$\vcadsr2\DEL_DOT\bestsurveydata\combotrips9501f.sav

File Security:

Enter Project Association:

Enter OrgArea Association:

Enter key words (seperated by commas):

Add place description: (search words : followed by description)

Full documentation can be very time consuming but also very important for the preservation of data and for determining suitable uses for it. To make the documentation process as easy as possible, voice recognition software was integrated with the web forms so it is possible for users to dictate documentation into the web form.

Creating New Data Files

The best time to document and reference information is when it is created, rather than as an after thought. Figure 7 below shows the “New File” page that is available to users at the point where they would save their data in any application. The full physical path filename can be cut and pasted into most software.

Figure 7. New File Screen

Enter the descriptive File Name:

Internal file name:

Directory/Place:

Full Physical Path Filename:

File Security: Extension: File Type: File Type Dropdown:

Enter Project Association:

Enter OrgArea Association:

Enter key words (seperated by commnas):

Add place description: (search words : followed by description)

Other examples of listings

Files can be documented just as directories or places. Figure 8 below shows a file listing for a directory containing DeIDOT Household Travel Survey Data. In this example the file size is also displayed. Any file can be opened by clicking on its name.

Figure 8. File Listing

\\marley1\d\$\cadsr2\DEL_DOT\bestsurveydata

Filename	File Descriptive	Doc?		
1:based.aml.txt	Trip purpose category creation procedure	DOC	<input type="checkbox"/>	920
2:combotrips9501f.sav	DelDOT HH Survey Trips data 1995-2001	DOC	<input type="checkbox"/>	17618176
3:grdtz3.dbf	Modified grid to traffic zone lookup table	DOC	<input type="checkbox"/>	268483
4:griddist9501.sav	Grid To Grid distance estimate for o/d in 1995	DOC	<input type="checkbox"/>	1959656
5:gridzip.sav	Lookup table for Modified grids and Zipcode	DOC	<input type="checkbox"/>	44248
6:nostoptrips.sav	Trips involving no stops DelDOT HH Survey	DOC	<input type="checkbox"/>	16878964
7:personmode02.sav	DelDOT HH Survey 2002 Person records	DOC	<input type="checkbox"/>	493048
8:personmode9501c.sav	DelDOT HH Survey 1995-2001 Person Records	DOC	<input type="checkbox"/>	3901244
9:personmode9502a.sav	DelDOT HH Survey 1995-2002 Person Records	DOC	<input type="checkbox"/>	4378224
10:purper.aml.txt	purpose categories for DelDOT HH survey	DOC	<input type="checkbox"/>	900
11:SPSSSrvr.log	Spss log	DOC	<input type="checkbox"/>	418
12:stoptrips.sav	Trips that includes stops DelDOT HH Survey	DOC	<input type="checkbox"/>	558332
13:trips02.sav	DelDOT HH Survey 2002 Trip data	DOC	<input type="checkbox"/>	1127728
14:trips02bak.sav	DelDOT HH Survey 2002 Trip Data backup	DOC	<input type="checkbox"/>	1090888
15:trips02raw.sav	DelDOT HH Survey 2002 Trip Data backup	DOC	<input type="checkbox"/>	799108
16:wilmapco.doc	Wilmapco request, Trip Mode, purpose, dis	DOC	<input type="checkbox"/>	72192

When file systems are read into Oracle data tables a size for each physical directory is calculated. The size calculated is the size of all files within the directory as well as the total size of all files in all subdirectories. This is unlike Windows Explorer that shows the size of only the files in the immediate directory. This is a feature that is very useful for data management. When listings are sorted by size its possible to know where the most space is being used in a particular directory or server. If more disk space is needed, a user or systems manager can more easily determine what areas might be archived or removed. Figure 9 shows this below for one users primary workspace. Size in bytes is in the last column.

Figure 9. Directory Listing by Size

1:personal	Personal data	DOC	<input type="checkbox"/>	07-JUL-04	2071750405
2:dtitraveldemand	Examination and Presentation of Travel Der	DOC	<input type="checkbox"/>	13-OCT-03	1863323554
3:filearchive	Data Interface backup and documentation	DOC	<input type="checkbox"/>	24-JUN-04	1652443228
4:hazmatrans	DEMA Hazardous Material Transportation f	DOC	<input type="checkbox"/>	03-JUN-04	1005430524
5:archive	Archive of numerous older projects	DOC	<input type="checkbox"/>	28-MAY-04	821354052
6:surveywork	DelDOT Travel Survey processing workspe	DOC	<input type="checkbox"/>	05-MAR-03	728883598
7:transitnet2	Transit network models	DOC	<input type="checkbox"/>	20-MAY-03	613007748
8:Exchange	E-Mail Location	DOC	<input type="checkbox"/>	17-JUN-00	597108844
9:hazard	DEMA Hazard Vulnerability GIS data CD	DOC	<input type="checkbox"/>	17-JUN-00	557896606
10:eback	Backup of e-mail	DOC	<input type="checkbox"/>	01-DEC-03	513091692
11:natsurvey	National Travel Survey	DOC	<input type="checkbox"/>	09-JAN-04	456440116
12:filearchive5	Old Catalog processing space	DOC	<input type="checkbox"/>	26-MAR-04	411828399
13:pms	DelDOT Pavement Management System P	DOC	<input type="checkbox"/>	12-NOV-02	323054782
14:interfacebak	Backup of Catalog Program Files	DOC	<input type="checkbox"/>	10-MAY-04	272667410
15:doted	DelDOT data, rd inventory, traffic counts, mi	DOC	<input type="checkbox"/>	21-MAR-02	260881011
16:sra	DNREC State Resource areas	DOC	<input type="checkbox"/>	20-SEP-01	238998303
17:kent	911 maps of kent in autocad	DOC	<input type="checkbox"/>	20-SEP-01	232923762
18:fhwa	City of Newark aerial photo and landuse	DOC	<input type="checkbox"/>	17-JUN-00	212688736
19:CTPP	1990 Census Transportation Planning Pack	DOC	<input type="checkbox"/>	20-SEP-01	204521903
20:jurisdiction	DTI Jurisdictional Maintenance Project	DOC	<input type="checkbox"/>	08-JAN-01	189946704

Organizational Areas

Users of personal computers mostly organize their information by placing their data in particular folders and subfolders on physical data storage disks. In practice using folders as the main way data is organized is very limited particularly for planners or analysts who often can use the same information for a variety of applications. For example, survey data or population projections may be organized and processed in one place but used by several projects that are organized in other places. Without copying data out of one folder and putting it in another, it can be difficult to view at the same time and easily access all of the pieces for a particular topic. Also folders tend to fill with bits of information that are only useful for a particular application. Scanning through 50 files when there is only one of interest is confusing.

When there are groups of individuals that work together and supply information to each other, as would happen in the Division of Planning, there are numerous places where information is kept under various organizational schemes. Sometimes an effort is made by a group to “get organized” and people decide together how the information will be

organized into a new folder scheme, and data is sifted through and copied and collected and put into new collections. This usually requires users to copy their data into some type of repository location on an ongoing basis, introducing another maintenance step, and a different location than where the source usually is created. The new organizational scheme that results is usually an improvement but, is not a sufficient organization for all types of uses. In general the effort is only focused on what can be determined as the most important or widely used information and actually, only addresses a fraction of the information needed, is a step removed from users, is not easy to reorganize, and usually the repository becomes stale and unused.

A method for users to create their own organizations or views of data is needed that:

- can be altered and built upon easily
- can allow for many ways of organizing information
- that is not dependent on physical locations of data or copying data
- get users more continuously involved in the organization of their information

Organizational areas in the web based catalog are implemented like virtual directories and referenced in a similar manner using the slash “\” to separate sub-levels (like sub directories). Within an organizational topic, there can be sub categories/levels, sub-sub categories, etc. Directories or files or even other organizational areas can be associated with any category or level of an organizational area.

As an example, the DeIDOT Household Survey is a travel survey that is conducted each year in Delaware. On the servers, data for the survey is organized and distributed in a number of folders usually based on the year of the survey.

A new organizational area was added and called “DeIDOT HH Survey”. Within the 1st (root) level there are six sub-categories as shown in Figure 10.

Figure 10. Organizational Area Subdirectory Listing

\\org\DELDOT HH Survey			
1:data	Deldot HH survey data files	DOC	<input type="checkbox"/>
2:directories	deldot hh survey directories	DOC	<input type="checkbox"/>
3:doc	survey and documents for deldot hh survey	DOC	<input type="checkbox"/>
4:other surveys	other local and national survey data	DOC	<input type="checkbox"/>
5:spss	spss syntax files for deldot hh survey	DOC	<input type="checkbox"/>
6:tables	tables and figures from deldot hh Survey	DOC	<input type="checkbox"/>

Any sub-category can be viewed by clicking on it name, for instance Figure 11 shows the area where documents are kept. The files shown are actually stored in several different physical directories.

Figure 11. File Listing In An Organizational Area

\\org\DELDOT HH Survey\doc

PROJECT FILE LISTING

File Name	File Descriptive	Doc?	
1:surveyvariables.doc	Explanation of DelDOT HH Survey variable	DOC	<input type="checkbox"/>
2:Trips.doc	2002 trip survey document	DOC	<input type="checkbox"/>
3:survey.doc	2002 DelDOT HH Survey	DOC	<input type="checkbox"/>
4:Trips.doc	2001 Trips chart	DOC	<input type="checkbox"/>
5:survey.doc	2001 DelDOT HH Survey	DOC	<input type="checkbox"/>
6:Survey99.doc	1999 deldot household survey	DOC	<input type="checkbox"/>
7:Trips99.doc	1999 deldot household survey trip chart	DOC	<input type="checkbox"/>
8:Survey98.doc	1998 deldot hh survey	DOC	<input type="checkbox"/>
9:TRIPS98.DOC	1998 deldot hh survey trip chart	DOC	<input type="checkbox"/>
10:Survey.doc	2000 DelDOT HH Survey	DOC	<input type="checkbox"/>
11:Trips.doc	2000 DelDOT HH Trip Chart	DOC	<input type="checkbox"/>
12:SURVEY97.DOC	1997 deldot hh survey	DOC	<input type="checkbox"/>
13:TRIPS97.DOC	1997 deldot hh surve trip chart	DOC	<input type="checkbox"/>

To further group relevant information about the Household Survey, there is a category named “directories”. This is an area that includes all directories that have relevant survey information stored in them. So directories from any server on the network at any level can be joined together in a view. This is useful when a user would like to personalize a view with the places and files that he or she wants to see most often. The user can see a view of directories or files completely independent of where they are located on the various servers.

Figure 12. Places/Directories In An Organizational Area

\\org\DELDOT HH Survey\directories

PROJECT PLACE LISTING

Placename	Place Descriptive	Doc?	
1:\marley1\d\$\cadsr2\DEL_DOT\2000	2000 deldot hh survey directory	DOC	<input type="checkbox"/>
2:\marley1\d\$\cadsr2\DEL_DOT\2001	2001 deldot hh survey directory	DOC	<input type="checkbox"/>
3:\marley1\d\$\cadsr2\DEL_DOT\2002	2002 deldot hh survey directory	DOC	<input type="checkbox"/>
4:\marley1\d\$\cadsr2\DEL_DOT\2003	2003 deldot hh survey directory	DOC	<input type="checkbox"/>
5:\marley1\d\$\cadsr2\DEL_DOT\94	1994 deldot hh survey directory	DOC	<input type="checkbox"/>
6:\marley1\d\$\cadsr2\DEL_DOT\95	1995 deldot hh survey directory	DOC	<input type="checkbox"/>
7:\marley1\d\$\cadsr2\DEL_DOT\96	1996 deldot hh survey directory	DOC	<input type="checkbox"/>
8:\marley1\d\$\cadsr2\DEL_DOT\97	1997 deldot hh survey directory	DOC	<input type="checkbox"/>
9:\marley1\d\$\cadsr2\DEL_DOT\98	1998 deldot hh survey directory	DOC	<input type="checkbox"/>
10:\marley1\d\$\cadsr2\DEL_DOT\99	1999 deldot hh survey directory	DOC	<input type="checkbox"/>
11:\marley1\d\$\cadsr2\DEL_DOT\bestsurveydata	Corrected, best, data files for deldot hh surv	DOC	<input type="checkbox"/>
12:\marley1\d\$\cadsr2\DEL_DOT\whitman	export of deldot hh survey data to whitman r	DOC	<input type="checkbox"/>
13:\moses5\e\$\home\FAE03291\surveywork	DeLDOT Travel Survey processing workspe	DOC	<input type="checkbox"/>

Project Areas

Because much of the DeLDOT Division of Planning’s work is focused around projects, Project areas were also included in the catalog. There is no difference however in how a Project Area and an Organizational Area are handled in the catalog. A project could just as well be viewed as another organizational area. Project data is often handled well by storing all data in a particular folder but as projects get more complex and as the number of projects increases, it is convenient to have one area to view all active projects, and have the ability to reference a variety of data in various locations. It was anticipated that other features that would be helpful for project management might at a later time be built into the catalog.

Figure 13. Project View

Select a place:

	DESCRIPTIVE NAME	DOC?	
1:\prj\Delaware Cadastral Standards	Delaware Cadastral Standards project	DOC	<input type="checkbox"/>
2:\prj\avltransit	Automated Vehicle Locating Cost/benefit ev	DOC	<input type="checkbox"/>
3:\prj\bike paths and housing values	Bike paths and housing value analysis proje	DOC	<input type="checkbox"/>
4:\prj\data management project	Data management project Deldot Planning	DOC	<input type="checkbox"/>
5:\prj\modechoice	DTI Mode Choice analysis	DOC	<input type="checkbox"/>
6:\prj\racca	Racca preferred place and file view	DOC	<input type="checkbox"/>

Search Capabilities

Any “place” including servers, physical directories, or organizational areas can be searched. Several servers or directories can be searched at a time. The check boxes next to the place listings are how the user indicates what areas will be searched. A simple phrase search is available at the top of every listing but more advanced searches are possible. Figure 14 shows the Advanced Search Page. The results of a search are listed and at that point it is possible to check some or all of the entries returned and assign them to an organizational area for later viewing. For instance, all servers could be searched for the extension “.ppt” to find all Power Point presentations that have been created. An organizational area could be defined called “Power Point” that would list all presentations from all users.

Figure 14. Advanced Search Page

ADVANCED SEARCH PAGE

Search for: Files Places

Case Sensitive ? No Yes

Enter search phrase: (Use of wildcards "*" and "?" permitted)

Exclude string: Another exclude string:

Search Word: And Or

Search Physical File Name ? No Yes

Search Full Path Name ? No Yes

Search Short Descrip Name ? No Yes

Search Key Words ? No Yes

Search Long Description ? No Yes

Advanced Search

A sample result from a search is shown in the figure below. By checking the box at the right of any file listing, and then entering a project or organization in the text box to the right of “Assign Checked” any search result can be made part of an organization or captured for later viewing.

Figure 15. Results of a Search

SEARCH RESULTS FOR "*.doc"

Filename	File Descriptive	Doc?	
1:\moses5\home\FAE03291\cadastral\cadastralmeeting1.doc	Meeting 1 Agenda, DGDC Cadastral SubCo	DOC	<input type="checkbox"/>
2:\moses5\home\FAE03291\cadastral\cadastralstandardv1.doc	Meeting 1 minutes, DGDC Cadastral SubCo	DOC	<input type="checkbox"/>
3:\moses5\home\FAE03291\cadastral\cadastralstandardv2.doc	Meeting 2 minutes, DGDC Cadastral SubCc	DOC	<input type="checkbox"/>
4:\moses5\home\FAE03291\cadastral\cadastralstandardv3.doc	Meeting 3 minutes, DGDC Cadastral SubCc	DOC	<input type="checkbox"/>
5:\moses5\home\FAE03291\cadastral\cadastralstandardv4.doc	Meeting 4 minutes, DGDC Cadastral SubCc	DOC	<input type="checkbox"/>
6:\moses5\home\FAE03291\cadastral\cadastralstandardv5.doc	Meeting 5 minutes, DGDC Cadastral SubCc	DOC	<input type="checkbox"/>
7:\moses5\home\FAE03291\hazmatrans\DNREC.doc	HazMat Trans Project, DNREC Meeting	DOC	<input type="checkbox"/>
8:\moses5\home\FAE03291\hazmatrans\HAZMAT RESEARCH NOTES.doc	HazMat Trans Project Research Notes	DOC	<input type="checkbox"/>
9:\moses5\home\FAE03291\hazmatrans\HAZMATRANSfinal.DOC	HazMat Trans Project final report	DOC	<input type="checkbox"/>
10:\moses5\home\FAE03291\hazmatrans\Ideas for Phase II.doc	HazMat Trans Project, initial ideas for phas	DOC	<input type="checkbox"/>

Performance

It is very important that users aren't waiting more than even a second or so for listing of directories, organizations, and projects. DUROSS now has a database that tracks close to 100,000 directories, 7 servers, and about 1.5 million data files. This amount of data is at least comparable or probably much more than managed by Division of Planning. Databases are indexed in a number of ways so that records can be retrieved quickly. The Web Server machine is a single processor Pentium 3, and the Oracle Database is served on a dual processor Pentium 3. The network speed is about 10 megabits at the user stations. On a dual processor, user machine listings were generated in one second to a second and a half, which is considered fast enough and the target performance. On a single processor, user machine listings were generated in 2 seconds to 2 seconds and a half. The user system seems to make the most difference. Going from 1 million files to 1.5 million files made no difference in performance. Search routines were much faster than those in Windows Explorer.

Organizational Frameworks For Transportation Data

The DUROS utility can then organize information in a variety of ways, and is unlimited in the number of views that can be generated and by the physical location of data. This section explores possibilities for how planning data may be organized for better access.

Projects, Impacts, and Performance

The Division of Planning is primarily about projects, and often the organization of information at DelDOT is around projects and programs. The various activities that are considered can be organized in Project Pools as they are in the Division of Planning Pipeline process. The Pipeline is a process where projects are evaluated and prioritized in the context of a variety of needs and impacts.

Figure 16.
Project Pools in the Pipeline Process

- Intersections
- Paving/Resurfacing
- Pedestrians
- Project Development
- Bicycles
- Bridge Management and Painting
- Corridor Capacity Preservation
- Drainage Improvements
- Environmental Improvements
- Rail Crossing Program
- Roadway Reconstruction
- Safety Improvements
- System Expansion
- Traffic Calming
- Transit Access

The success of projects is related to goals that are set and evaluating the effectiveness of activities involves the development of measures for the impact and performance of activities. Impacts can be in safety, mobility, accessibility to travel options, condition of facilities, community impact, and environmental impact.

Figure 17.
Impact Categories

Safety
Mobility
Accessibility to travel options
Condition of facilities
Community impact
Environmental impact

What is important to planners are measures available to specify the impacts. Figure 18 provides examples of the measures needed in the data.

Figure 18. Examples of Important Planning Measures

Safety
Accidents (totals, deaths, by intersection, by roadway, by mode)
Condition data (slickness of road services, visibility)
Transit facility incidents

Mobility
Delay data (average delay)
Failed Intersections
Time saved in system
Travel time between various destinations
Volume to capacity ratio

Accessibility
Availability and level of service of transit
Availability and level of service of other modes
Affordable transportation for low income

Condition of facilities
Condition ratings – for facilities bridges, roadways, signals, etc.
Public opinion
Maintenance costs

Community Impact
Traffic volumes – current, past, projected
Volume to capacity ratios
Open space
Costs for transportation facilities
Measures of economic impact
Energy usage
Induced travel

Environmental Impact
Air quality measurements
VMT
Noise measurements
Drainage problems

Dimensions of planning information

Information views and queries common to planning involve looking at data with respect to a feature or attribute of the information. In particular, information exists and can be summarized for particular geographic areas, time periods, and other variables of interest. Effective information systems would have to allow for these different views of the information, here referred to as the dimensions of the data.

The spatial dimension

Transportation data is predominantly spatial data. Travel demand, supply, and impact information generally contain a spatial dimension. Many questions faced by the Division of Planning are about a particular area or location. Data is most often summarized by a particular area or geographic unit such as a neighborhood, a portion of roadway, a traffic zone, a census tract, a county, or state level. Effective information query tools for planning include the ability to look at data spatially, which is one of the reasons that geographic information system tools are so useful for transportation applications.

The time dimension

Time is another important dimension of transportation data. Planning involves predicting and planning for future conditions. Analyzing progress involves being able to examine past trends. Comparing information often involves a view for a particular time period.

Viewing data with respect to a particular variable

Often information is viewed with respect to another variable (or variables) and these represent additional dimensions. Examples for the spatial and time dimensions mentioned would include for instance population “by” year or “by” area but often other variables are involved such as travel “by” age group or transportation mode preferences by socio-economic status.

Origins and destinations

Often planning data is viewed in terms of trip origins and destinations. The data of interest usually is the travel time between places, the number of trips generated between places, or transit services or other travel mode available between. Data is viewed in these cases for an origin / destination pair.

Supply, Demand, Impact, and Performance

An approach to information systems that address planning's needs best is one that addresses higher level concepts and goals at the start. In this work an organizational framework for transportation data as suggested by the National Cooperative Highway Research Program* is used to work down to the level of data and software and hardware systems rather than the other way around. In this organization the principal categories are Supply, Demand, Impacts, and Performance of transportation systems. Figure 19 on the next page, shows the categories of supply, demand, impacts, and performance of transportation systems that could serve as the data topics and measures necessary to fully describe the current status of the system. These categories can be viewed as the areas of focus of planning activities, and any information request can be viewed as addressing one or more of these areas. Transportation planning requires an understanding of the supply of transportation that includes the features of all facilities and services available to the public. Demand is the need for transportation as produced by populations, employment centers, general land use, and commerce. Planning agencies must continually address impacts of new facilities and projects, and the growth and changes that occur. The public is particularly interested in the performance of transportation facilities, and how well facilities and services (supply) meet the need for travel (demand) "Are facilities adequate?", "Do they effectively serve the public?", "Is there a need that must be addressed", "Are there safety concerns?".

The Need For Tools To Organize Data

As shown in this section there are ways of effectively organizing planning data. But there are many ways that planners may want to organize information at the same time. It is suggested that there is no one method of organization that will meet all needs. If organization is based on physical locations where data is stored, then data must be moved or copied to conform to varying organizations. What planners need is the flexibility to view data in a number of ways and to do this in a decentralized environment. DUROS is the type of tool that will meet this need.

* National Cooperative Highway Research Program (NCHRP) Report 401, Guidance Manual for Managing Transportation Planning Data, National Academy Press, Washington, D. C. 1997, pgs3 and 6

Figure 19. Transportation Data Organization**Supply**

System – network, routes, right of ways, inventory, functional class, capacity
Service – modal access and connections,
Facilities – Inventory of facilities
Condition – pavement data, all other condition, age
Project data – All projects, project evaluations, history, maintenance?`

Demand

Economic – Income, employment, vehicle ownership, commodity, commerce
Demographic – Population, labor force, household characteristics
Land Use – Zoning data, access data, housing,
Commodity - modal factors, modal split, commodity
Travel data – Trip generation, distribution, generator, traffic volume, VMT
Traveler Behavior – Mode choice, route choice, user preferences, historic,
current,proj

Impacts

Air Quality - registration/fleet, VMT, speed, trip data, impact assessment
Other Environmental – visual aesthetic, noise, vibration, parks, stormwater, archeology
Land use – socio economic impact, neighborhood impact
Energy – Energy consumption, efficiency, price impacts
Economic Growth – employment impacts, access to markets

Performance

Safety Data - Incident, accident data, security, medical, emergency services
Performance Measures –
Intermodal system performance
Condition, level of service
Congestion
Delivery times
User cost data

Further Implementation of the Interface

Work will in all likelihood continue developing DUROS at the Center For Applied Demography and Survey Research (CADSR). There is a real need for this type of tool. CADSR has a similar situation as the Division of Planning in that there is a range of analysis and data generation continually going on in a variety of topic areas, among several professionals, in a decentralized environment. Demonstrations of DUROS will be provided to DELDOT and perhaps implementation of DUROS can occur in the Division of Planning.

Other Features Planned For DUROS

DUROS is now a functional interface for documentation, referencing, organization, and search in an intranet environment. DUROS is already in use by analysts and information system managers at the Center for Applied Demography and Survey Research. Its first uses were for file management on the servers, to determine more clearly how disk space was being used. There is room for improvement and additional features to be added beyond this first demonstration version.

There are some additional Administrative functions that are needed involved with the movement and deletion of data, and security.

A utility to take the organizations and create physical directories and file folders would be useful when preparing exports of the data or if system administrators wanted to structure the physical disk along the lines of particular organizations that was defined.

Though DUROSS can be updated to reflect any changes in the servers, an efficient tool to quickly import a data collection would be useful.

Reading file systems and preparing them in Oracle databases each night uses SPSS, File Organizer, and ARCGIS. It is most likely that this software could be replaced using some programming language.

Tool bars and menus could be improved to make them smaller and graphically more appealing.

The ability to handle e-mail files for search and cataloging should be added.

Additional work is needed to complete the voice recognition features.

Many search utilities can search within common file types if they can be specified as part of some collection. This would allow search beyond file names and paths, and beyond documentation added by the user. For instance all "*.DOC" Word files could be searched for any containing the word "transportation" within the text of the document, for instance.

Conclusion

In recent years the DelDOT Division of Planning has been very interested in how data can be better managed and how information systems can be improved. There is a growing need to be able to continually find, analyze, and present information. In interviews with Planning personnel, there was a general sense that information systems were not integrated and that benefits could be realized if information was more widely available and usable outside of the primary function it was created. In studying the Division of Planning's resources, it was obvious that there was a large quantity of data that was available within various parts of the group, but there was no sense that the data was collected and managed as part of a coordinated information library or system that preserved and facilitated access and use of the data.

The Division of Planning's use of information is necessarily decentralized. DelDOT wide resources like the Road Inventory are served well by centralized approaches, but for the most part information is dynamically spread across many people, initiatives, systems, and countless topics. More information and analysis is generated every day. It is not a situation where an outside information group can come in and make sense of it all. Division of Planning personnel must be involved in the organization and referencing of their information. If the Planning group cannot organize their information, it is very doubtful anyone else will.

The problem is, that in the heat of the day no points are awarded for good data management. Personnel are constantly pressed to find, prepare, and present information and then move on to the next topic. Managing information cannot be an after thought or something that will possibly happen when there is a little time for it. Personnel need some type of tool that can allow them to contribute to a Division of Planning information resource in the process of their work. Something that would allow flexibility in how they organize and view data. Something that is as easy as possible. And something that can be built amidst the decentralized chaos.

This project took a realistic look at what might be done to assist the Division. Preparations of data collections, development of repositories and agency wide enterprise databases, reorganizations, and development of standards, are initiatives that have their place. Planning staff can appreciate the benefits of such activities, but they are not a solution, and certainly not a near term solution, to the day-to-day crunch felt by planners and analysts.

The Documentation Utility For Referencing, Organization, and Search is an example of the kind of tool that is needed by the Division of Planning. First, the massive amounts of data produced in the Division must be preserved, referenced, and dynamically organized with day to day involvement of staff. The system could be used not only to reference information inside the Division but also information from other groups.

Delaware Center for Transportation University of Delaware Newark, Delaware 19716

AN EQUAL OPPORTUNITY/AFFIRMATIVE ACTION EMPLOYER

The University of Delaware is committed to assuring equal opportunity to all persons and does not discriminate on the basis of race, color, gender, religion, ancestry, national origin, sexual orientation, veteran statutes, age, or disability in its educational programs, activities, admissions, or employment practices as required by the Title IX of the Education Amendments of 1972, Title VI of the Civil Rights Act of 1964, the Rehabilitation Act of 1973, the Americans with Disabilities Act, other applicable statutes and University policy. Inquiries concerning these statutes and information regarding campus accessibility should be referred to the Affirmative Action Officer, 305 Hullihen Hall, (302) 831-2835 (voice), (302) 831-4563 (TTD)

