

#6152

# A Collaborative Support System for the Review of Building Energy Data in the LoanSTAR Monitoring and Analysis Project

Dean Willis, Jeff Haberl, Ph. D., P. E.

Energy Systems Laboratory, Texas A&M University, College Station, Texas

## ABSTRACT

The LoanSTAR Monitoring and Analysis Project (MAP) collects a large volume of data pertaining to the analysis of building energy usage. Experts within the project review these data for accuracy. This review is collaborative, and is carried out via a traditional paper-driven process. Growth of the project causes an information saturation with insufficient time for the existing system of review. Powerful computing facilities are available to the project, and are currently being used to feed the paper-driven review process. Utilization of these facilities to provide a computer-supported collaborative environment for the review of building energy analysis data is feasible. This paper addresses the requirements such a system must meet, and proposes a design which can be implemented at minimal cost. The basic research questions raised focus on the criteria for user acceptance of and satisfaction with the proposed system.

## BACKGROUND

The LoanSTAR Monitoring and Analysis Project (MAP) is administered by the Governor's Energy Office (GEO) of the state of Texas. It is a statewide project to implement energy conservation measures in public-sector buildings. As the monitoring contractor for this project, Texas A&M University's Energy Systems Lab (ESL) carries out a number of subtasks. One of these tasks is the measurement of energy usage in many of the participating buildings. The purpose of such measurement is to provide data for the analysis of the effectiveness of building system retrofits and operational measures intended to conserve energy.

To achieve this purpose, the buildings are instrumented with automated data recorders. Time series data are collected from these instruments via telephone lines and used for analytical purposes. This collection process is relatively error-prone, as both the instrumentation and its operation are complex. To maintain a high standard of accuracy, the ESL carries out a review of acquired data to detect errors and provide for their timely correction.

Although data acquisition is principally an automatic process, review is not. Review is currently carried out in a labor-intensive manner by a group of reviewers with

minimal automated support. As described by Bottger and Yetton (1), the subject-matter (building systems) experts within the group dominate the process, and their availability is a bottleneck to system performance.

As the number of buildings monitored by LoanSTAR increases, so does the amount of effort required to perform reviews of the acquired data. At some indefinite point, the time required to conduct even the simplest reviews will exceed the time available. As Bush (2) and Englebart (3) pointed out, such an information explosion is occurring in all fields of study.

Automated support for review is required to assure that the saturation point lies outside the bounds of the project and to streamline project operations for maximum efficiency. Efficiently automating the thinking process or expert knowledge used by the reviewers may not be possible, especially since no solid consensus on how errors are spotted has been reached. However, expediting the communications processes involved in the collaborative process is possible by developing what Englebart calls "tools to think with."

## PURPOSE AND SCOPE

This paper discusses the requirements for a support system for the collaborative review of building energy analysis data within the LoanSTAR project, and proposes a design meeting those requirements. This process is carried out in three steps:

- 1) Examining the current manual system.
- 2) Exploring the available computing resources.
- 3) Discussing a feasible model for an automated system.

Additionally, this document discusses research questions which might be addressed during the implementation of such a system.

## CURRENT REVIEW METHODS

The current review method for acquired data centers on circulating two documents: a weekly Inspection Plot

Notebook, or IPN, and a Monthly Energy Consumption Report, or MECR. The MECR is also the principle means of feedback to coordinators within the agencies responsible for the buildings being monitored.

### Inspection Plot Notebook

The ESL data processing staff, prepares the IPN weekly. It contains several pages of time series and scatter plots for each building. Each page contains several plots, and may be unique to a given building. In practice, the IPN is divided into several sections, each of which is circulated separately. Multiple copies of each section may also be circulated. After preparation, the inspection plot notebook is circulated to each of the reviewers. Reviewers make comments on the report by typing their comments into a file on a diskette circulated with the report. Comments by the reviewers are collected by the Site Description and Agency Contacts group (SDAC), and problems detected at this stage are communicated to responsible parties for correction. Figure 1 is a diagram of this process.

### Monthly Energy Consumption Reports

At the end of each month, the accumulated comments and reports on corrections are assembled by the data processing staff into a draft of the Monthly Energy Consumption Report (MECR), which is circulated for comment. Comments are collected, problems resolved, and a revised draft is prepared. The revised draft is examined by a subset of reviewers who coordinate final corrections with the data

processing staff. Eventually, a final MECR is produced, and is released to appropriate agencies. Figure 2 is a diagram of the MECR process.

### Task Description of the Review Process

The overall review process consists of three sub tasks: preparation of the material, analytical review, and compilation of comments. Concurrent with review, savings calculations are performed by the analysis group. Although both preparation of the material and compilation of comments are critical and time consuming, the mechanics involved are straightforward and suitable for automation by traditional means. The conceptually interesting task is the analytical review of the observed data.

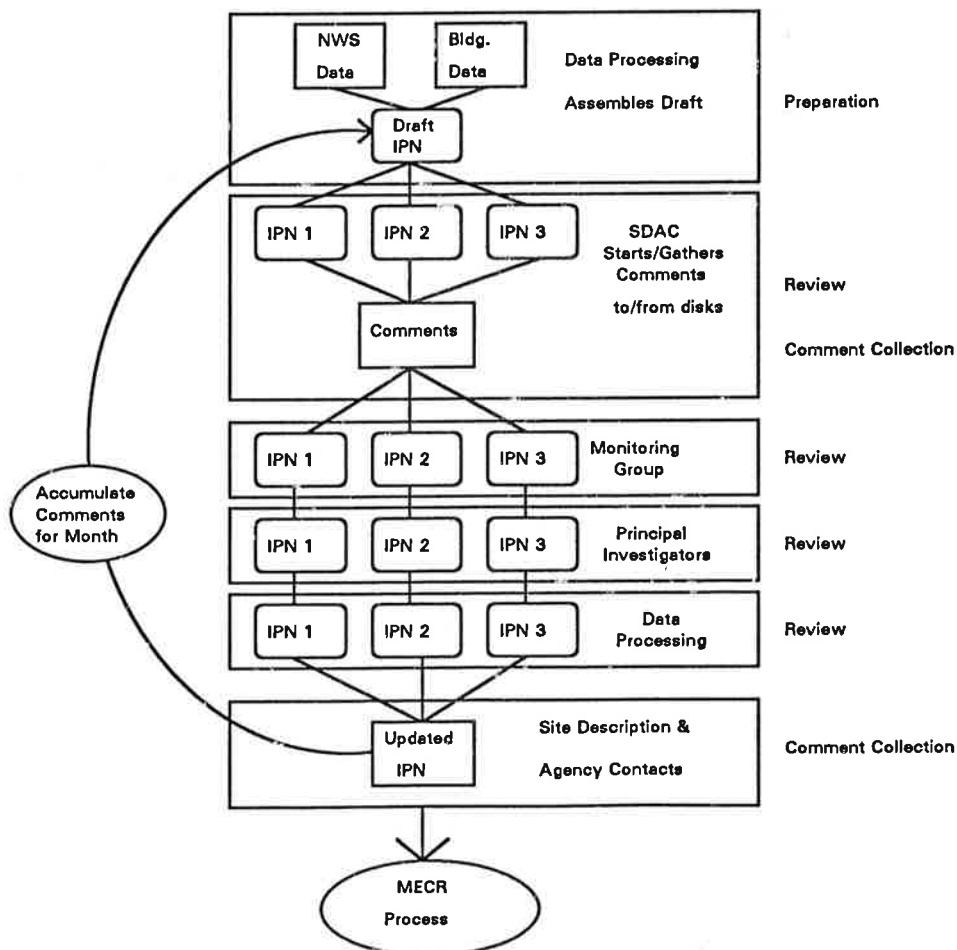


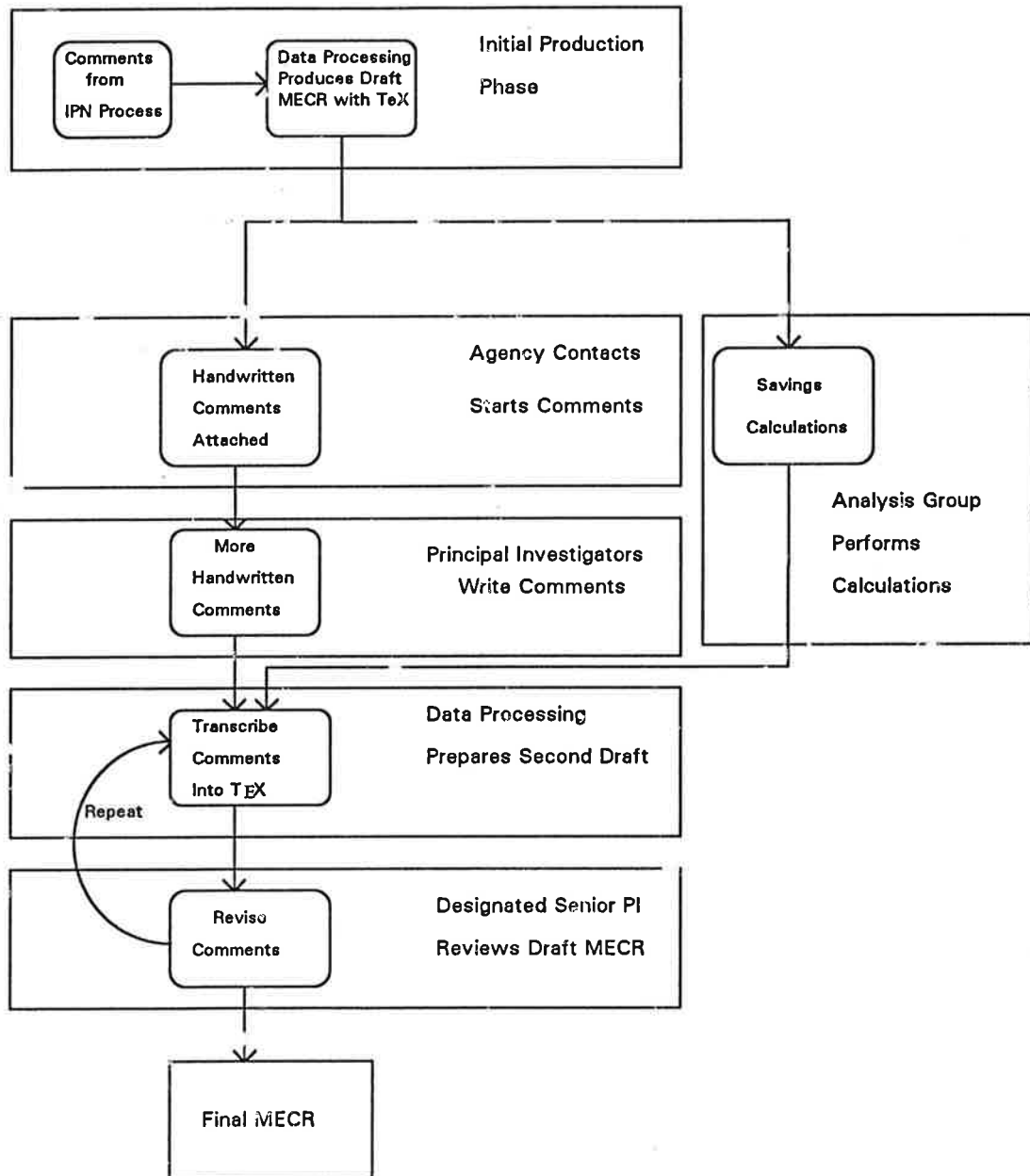
Figure 1: Circulation of the Inspection Plot Notebook

Reviewers visually inspect the plots in the IPN (weekly Inspection Plot Notebook) and MECR (Monthly Energy Consumption Report). In addition to typographical and labeling errors, they are looking for conditions that signify errors in the data acquisition process or in building operations, administration, and maintenance (OA&M). As reviewers detect and/or resolve anomalies, they write comments into the report comment file using a word

the remainder of the organization, thereby reducing dependency on the unique skills of the reviewer (8). Reviewers detect error conditions in two primary ways:

- Discrepancies between observations and expectations.
- Anomalies compared to the building's historical data.

**Discrepancies Between Observations and Expectations.**



**Figure 2: Circulation of the Monthly Energy Consumption Report**

processing program . The additional knowledge generated by the reviewer in this process becomes available through

The first detection method is discrepancy between the observations depicted in the given plots and the reviewers

expectations based on his knowledge of the systems involved or of similar buildings. In the case of a building new to the project, this may be the only readily available method for checking monitoring performance. The reviewer compares the ranges of measurements and the periodic wave forms of observations to general expectations. Further, the reviewer looks for expected interactions between signals, such as the associations between outdoor air temperature and cooling load.

#### Anomalies Compared to Building History.

Sudden changes in the observed behavior of a building system often indicate a measurement equipment failure, a building system failure, or an operational change. Reviewers may detect such changes by comparing new sets of observations with older sets (when available). In practice, reviewers tend to step through old IPNs and MECRs for the building concurrently with the report being reviewed. Reviewers visually compare each new plot to one or more historical plots. In many cases, reviewers actually overlay two pages of plots.

### COMPUTING RESOURCES AVAILABLE

The Energy Systems Laboratory (ESL) maintains a variety of computing systems in conjunction with the LoanSTAR Monitoring and Analysis Project. For practicality, a collaborative support system must operate within the limits defined by this equipment. Systems include advanced RISC based workstations and conventional personal computers. Most of these systems are interconnected, either by the general campus network, or with dial-up telephone lines.

#### UNIX Server and Workstations

The ESL maintains a minicomputer and workstation

environment using RISC processors running a UNIX operating system. Models in use at the ESL are rated at 18 to 36 MIPS (millions of instructions per second).

The UNIX systems handle the bulk of data processing and numerical analysis on the LoanSTAR project. However, there are not enough UNIX workstations capable of graphics for all of the planned users of the collaborative system, which reduces the feasibility of a collaborative support system based solely on the UNIX systems.

#### Conventional Personal Computers

Most of the conventional office-automation work, as well as communication with data recorders, is performed with standard IBM-PC compatible personal computers. Machines in use range from original 4.77 Mhz IBM PC-XT systems with Intel 8088 processors, 640 kb of RAM and 10 Mb disk, to 33 Mhz 80486 systems with 16 Mb of RAM and 200 Mb disk drives.

Each faculty or staff member has a dedicated machine, as do graduate students whose tasks warrant such allocation. Public access machines are available to other graduate and undergraduate students. These machines are for the most part 80386 models, with an average of 4 Mb of RAM and 80 Mb disk. Most machines have VGA (640 by 480 by 16) graphics, with either color or monochrome 14" monitors. Many of the machines have extended VGA graphics, generally to 1024 by 768 by 256 color with 14" or 16" non-interlaced monitors.

Desktop machines provide word processing, spreadsheet, and other functions. Microsoft Windows is the environment of choice. Windows applications are used for word processing and spreadsheet work. Other Windows applications, including graphical/statistical data analysis,

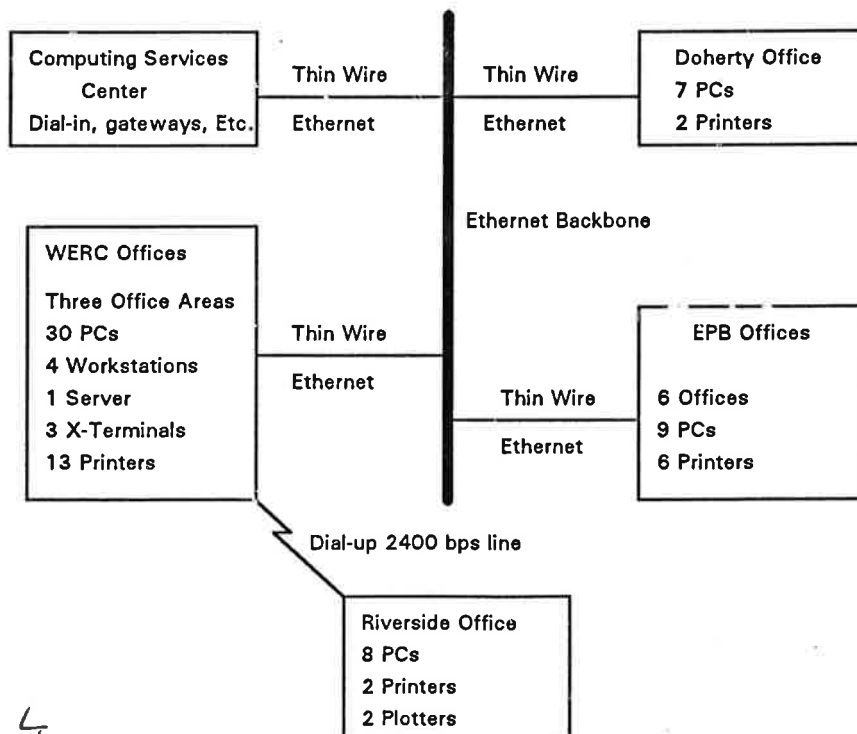


Figure 3: Network Access and Computer Facilities

drawing, and desktop publishing software are available on some machines. Because all users are familiar with Windows, any proposed system running on personal computers should be Windows based.

the data processing staff to translate raw data into the hypermedia (6) structure for review. Of the two, the browsing component is more complex, and is the focus of most of the following discussion, which includes:

## Campus Network

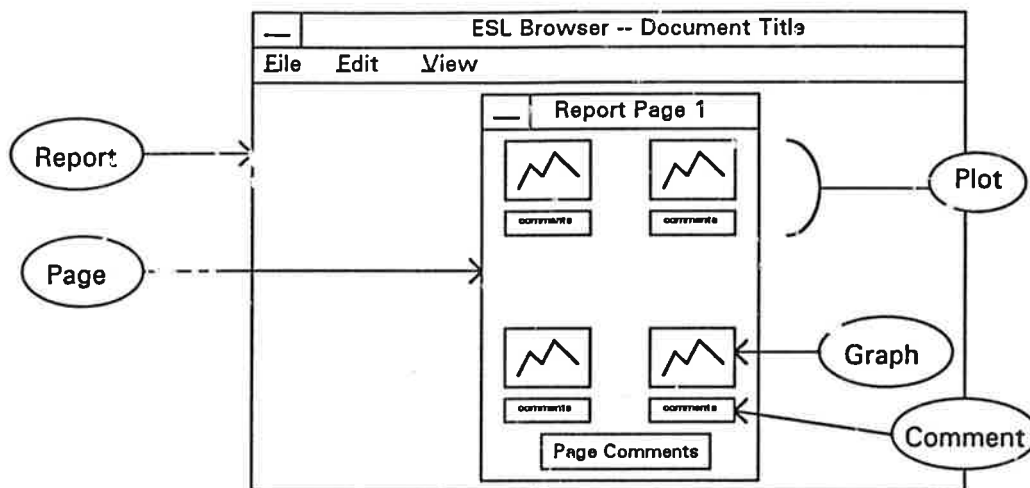


Figure 4 -- Window Hierarchy

With the exception of portable computers and computers used as dedicated data acquisition controllers, all ESL systems on the main TAMU campus are interconnected via the campus Ethernet network. A logical representation of the system is shown in Figure 3. The physical layout is somewhat more complicated, with some redundancy, multiple media types, and different routing strategies for different protocols.

The network allows all of the connected systems to take advantage of shared resources. The main ESL server provides over 2 gigabytes of disk space. Remote systems use networking software to transparently access this disk space. The effect is as if the shared area of the server's disk were physically attached to each remote system at the same time. Files on shared disk can be read and written by the remote systems as if they were local, subject to the security controls of the server.

This transparent file sharing allows the server to be used as a central repository of information that can be accessed by applications running on the remote computers.

## PROPOSED SYSTEM

The proposed system can be divided into two components: A "browsing component" to be used by reviewers to examine and comment on the observed data, and an "automated data set construction" component to be used by

- Functional Requirements of the Browsing Component
- Target Platform for the Browsing Component
- User Interface of the Browsing Component
- Data Model of the System
- Automated Data Set Construction Component
- Development Tools for the System

**Functional Requirements of the Browsing Component**  
On the basis of the task description developed earlier, the browsing component needs to provide facilities for:

- Displaying plots
- Displaying comments made by other reviewers
- Adding and editing comments
- Comparing sets of plots
- Collecting comments into summaries

The system proposed implements on-line analogs of the IPN (inspection plot notebook) and MECR (monthly energy consumption report), while providing extensibility into other structures as needed. The existing structures of IPN and MECR are maintained for two reasons. First, reviewers are familiar with these formats. Second the existing structures were developed to facilitate detection of unusual

interactions between systems (see *Discrepancies Between Observations and Expectations*, previously discussed).

### Target Platform for Browsing Component

The browsing component could be delivered on personal computers, UNIX workstations, or both. All reviewers have access to networked personal computers, and few have access to UNIX workstations, making the personal computer the better choice. Extending the system to include UNIX workstations at some future date is possible but not required at this time.

### User Interface of Browsing Component

The user interface is modeled around the structure of the existing paper reports. The system makes plots and text boxes of the reports into interactive components. "Direct manipulation" (10) techniques are used to help the user work with these components. A windowed display consistent with that of standard Microsoft Windows applications is proposed. The display is organized hierarchically along the same lines as the underlying structure.

The entire application is displayed in a containing or parent window (Figure 4). This window is optionally subdivided under user control as the user moves through the structural hierarchy. Each report, such as an IPN, is composed of pages. Each page is composed of one or more plots and a page-level comment box. Each plot decomposes to the plot graphic and a plot-level comment box.

Moving up or down the hierarchy, any component can be expanded into a subwindow which can be individually manipulated within the parent window. Any subwindow can be resized, expanded to the full size of the parent (zoomed), closed, tiled, or cascaded along with other subwindows. Tiling in this context is the process of evenly dividing parent window space into non-overlapping subwindows. Cascading is a process wherein the subwindows are resized into consistent sizes, and stacked in an overlapping manner along a diagonal. Both processes are familiar to users of the Windows file manager, program manager, Microsoft Word, or Microsoft Excel.

Movement down a hierarchy is accomplished by double-clicking a component, which causes the component to be expanded into a subwindow, which can optionally replace the parent subwindow. Movement in the opposite direction is accomplished by closing the new subwindow, by selecting the parent from a list of available windows, or by clicking on the title bar of a subwindow which is not completely hidden. Although complex to describe, the process is natural and transparent to an experienced Windows user.

For example, consider a user browsing through a single IPN structure. The user begins by activating the browser from the Windows Program Manager, then selecting the particular IPN from the browser's "File Open" menu. Within the browser window, page 1 of the given IPN is displayed -- a set of plots, any previous comments for that page, and any previous comments for each plot. If a signal outage in the third plot on the right is noticed, the user double-clicks that plot. A new window opens up, containing an expanded plot, and a comment box. The user clicks in the comment box, causing a text cursor to appear, then types in a comment. Double-clicking on the bar of this subwindow dismisses the plot, restoring the original view.

As a result of this type of user interface, the user is empowered to arrange the display in whatever manner best suits the needs of the moment while working in a familiar environment. Simple steps can produce complex results, expediting the exploration and comparison process. Multiple reports can be active at one time, facilitating comparison. Overlay of one plot with another might be a useful feature. It is important to remember that there is no "one right way" to use such a system (6). A possible extension of this process is discussed in the later section "Research Questions: Virtual Structure".

### Data Model of System

The proposed system is based on shared files located on the UNIX server. Access to these files across the network can be provided by NFS services and/or Novell Portable NetWare services in a conventional manner. Each report includes a "master" file and a number of "node" files.

The master file describes the overall structure of the report, referencing individual graphic and text items together into commented plots, pages, and reports. The format of the master file is not completely defined, but it will meet several criteria:

- Readable, editable text for debugging
- Suitable for automated construction
- Provide information for structural comparison

Node files can be either text or graphs. Text files will be standard ASCII text. Graphs can be either vector-oriented Windows Metafiles (WMF) or pixel-oriented bitmaps (BMP), both of which are Windows standard formats supported by Visual Basic.

### Automated Data Set Construction Component

The Automated Data Set Construction Component is an administrative tool that assists in the development of reports for the system. It is applied at the start of the IPN and MECR creation process by the Data Processing staff. Using a template, this component creates a working directory and prepares a master file for a report. Individual fields in the

template are filled in with appropriate file names and section titles to produce the report master file.

Over time, the ESL will enhance the function of this component. One interesting possibility is the use of pattern-recognition techniques to flag plots which vary from their archetypes.

### Development Tools for System

Given the choice of Microsoft Windows as a target environment for the browsing component, the available development tools are Borland C++ and Microsoft Visual Basic. Both support structured, object oriented development. Borland C++ may produce faster code, but suffers from a long edit-compile-link-debug development cycle coupled with the large learning curve of traditional Windows development. Microsoft Visual Basic offers an incrementally compiling development environment with rich debugging capabilities and an object-oriented structure for quick development of Windows applications. Both languages offer access to low-level Windows system calls, but Visual Basic does not require the use of such features in most situations.

Development work of the sort proposed is exploratory in nature, resulting in a high probability of design change over a project lifetime. In the university environment, less experienced programmers will be required to adapt and maintain the system after initial development. Microsoft Visual Basic meets these criteria better than Borland C++, and also offers quicker development time for the reasons discussed above. Therefore, initial system development of the browsing component will be done using Microsoft Visual Basic.

Automated data set construction may be performed either on personal computers or on a UNIX server or workstation. The type of programming involved is different from that of the user interface, being oriented towards text stream processing rather than user interaction. Therefore, development work on the automated data set construction components will be done in portable C, using coming PC and/or UNIX C compilers.

## RESEARCH QUESTIONS

A project of this sort poses general research questions in both the computer-human interaction (CHI) and hypermedia areas. Researchable CHI issues include determination of the design characteristics affecting user acceptance and usage, and productivity of the resulting system relative to the current paper system. The primary hypermedia question is whether the use of virtual structure can be exploited to develop a better understanding of the system interactions.

### Acceptance and Usage Constraints

The possibility of developing an on-line system has already drawn criticism from current reviewers who do not expect such a system to be useful. The objections raised include limited on-screen display space, slow performance, and an inability to easily compare sets of data. Determination of which criteria affect user acceptance and performance could assist in setting standards for future development. The variety of computing systems available for use on this project will enable data to be collected on effects of screen size and resolution as well as speed.

Systems available to the project include displays ranging from 12" units with 640 by 350 pixels to 16" models with 1024 by 768 pixels. Careful design of usability experiments could shed light on the effect of display size and resolution within this range.

Performance can have a great impact on usability. According to Robertson, McCracken and Newell, slow response can preclude usability in hypermedia systems (9). Systems in use will range from 8 Mhz 80286 models to 33 Mhz 80486 systems, with a full range of I/O performance and screen update speed as well. A related operating systems project might study the effectiveness of preloading (forward caching) node transitions to minimize the wait time imposed by network activity.

### Productivity of the Automated System

The real metric for the success of a collaborative support system is productivity. Do users of the system produce better results faster than they did with the paper system? If so, or not, why? Pre- and post-implementation studies of effort and satisfaction with results may provide insights into the effectiveness of the system and the determinants of that effectiveness.

### Virtual Structure

Virtual structure issues have been discussed by hypermedia scholars for some time (4) (5). In this context, virtual structure refers to relationships between pieces of information that are apparent due to the arrangement or content of elements in the system. For example, a virtual search operation in a set of inspection plots might translate to "find all page 2s with references to steam traps with comments mentioning bypass valves."

The comparison of structure also falls into this category. One feature of the proposed system is comparing two similarly structured reports in "lock step", advancing through them so that relevant information from each report is displayed collaterally (7). Such a feature might be useful in a review process that focuses on comparing this week's results to last week's. Analysis of the usage of the proposed system will answer that question. Other structural

operations might also be useful -- perhaps use of the system and study of that usage will suggest them.

## CONCLUSION

Review of building energy analysis data in the LoanSTAR project is a critical labor-intensive manual process that must be improved to remain feasible. While automating the thinking process may not be possible, automating the communications and paper-handling is. The proposed system implements a simple shared-file hypermedia system using existing computing systems within the Energy Systems Lab. Analysis of results obtained from the implementation of such a system will answer questions relating to the general usefulness and specific parameters affecting the usefulness of such a system.

## REFERENCES

- (1) Bottger, P. and P. Yetton, "An Integration of Process and Decision Scheme Explanations of Group Problem Solving Performance," *Organizational Behaviour and Human Decision Processes*, vol. 42, 1988, pp. 234-249.
- (2) Bush, V., "As We May Think," *The Atlantic Monthly*, 176:1, Jul. 1945, pp. 101-108.
- (3) Englebart, D., "A conceptual Framework for the Augmentation of Man's Intellect," *Vistas in Information Handling*, Vol. 1, Spartan Books, Washington, DC, 1963, pp. 1-29.
- (4) Halasz, F., T. Moran, and R. Trigg, "Reflections on NoteCards: Seven Issues for the Next Generation of Hypermedia Systems," *Communications of the ACM*, 31(7), 1988, pp. 836-855.
- (5) Halasz, F., "Seven Issues Revisited," *Proceedings of the Hypertext 91 Conference*, San Antonio TX, December 1991.
- (6) Nelson, T., "A File Structure for The Complex, The Changing, and the Indeterminate," *ACM Proceedings of the 20th National Conference*, 1965.
- (7) Nelson, T., *Literary Machines*, Project Xanadu, San Antonio, TX, 1987.
- (8) Pareek, U., "Organizational Learning Diagnostics (OLD): Assessing and Developing Organizational Learning," *The 1988 Annual: Developing Human Resources*, University Associates, San Diego CA, 1988.
- (9) Robertson, G., D. McCracken, and A. Newell, "The ZOG Approach to Man-Machine Communication," *Int. Journal of Man-Machine Studies*, Vol. 14, 1981, pp. 461-488.
- (10) Shneiderman, B., "Direct Manipulation: A Step Beyond Programming Languages," *IEEE Computer*, 16(8), 1983, pp. 57-69.