A Strategy For Distributed Information Management In A Public Institution

By R. William Maule

This article examines a statewide videotex service which decentralized information management and operational responsibilities to help enlist information providers, save money and increase overall system usage. Automated processing routines enable research scientists to continuously collect, filter and disseminate their data to clients, industry and the general public.

The study is of an online information service begun in 1988 to deliver in electronic formats information previously available only in print. The goal of the service is to reduce overall operating costs by shifting from print to electronic delivery, and to increase the efficiency of information transfer by shifting from postal mail to online resources. Startup operations are centralized since the print and information editing functions being upgraded were a centralized, staff operation.

The service is in support of the graduate and undergraduate teaching programs, and the statewide research and cooperative extension programs, of the University of Florida, Institute of Food and Agricultural Sciences. The service is for approximately 1,000 faculty, 2,500 support staff, and 900 students. Private business users and the general public access the service through extension offices in each county in the state. The primary information providers are scientists working in such fields as agriculture, environmental science, biotechnology, genetics, microbiology, and public health. The service steadily expanded in information content and utilization from 1988 to 1991, and as of August, 1991, contained about 35,000 pages of information and received about 2000 accesses per month.

Networks and Computing Resources

The information processing operations are based on VAX minicomputers. The VAX 6320 cluster on the main campus is used by students, faculty and staff in county offices and state research centers in central Florida. Information from the campus machines is downloaded to VAX minicomputers in northern and southern Florida for use by scientists in research centers and county offices in those areas.

The online service operates in VMS using a custom "C" program to produce a videotex application. The service is cross-referenced with an information inventory system resident on an IBM 3090 600J mainframe/supercomputer through a code which identifies the general category of information, the originating source, a topic and subtopic, and a unique identification number. On the VAX minicomputers, the code indicates the directory structure and physical location of information within the service.

The IBM 3090 also hosts the automated library bibliographic record system for the state university system and is the entry point through which university information is transferred into the national libraries, such as OCLC and the National Agriculture Library. Using the reference code, bibliographic information from new research can be derived from the videotex service, tracked to the inventory management system, and entered into state and national libraries.

Users access the service either directly through the campus ethernet LAN, or by using the Florida Information Resource Network, which is an X.25 packet-switched network which provides access to public computing facilities and information centers throughout the state. The videotex service disseminates university, state, and federal research. Information from other states, and information services from other countries, are accessed from a menu which automates network connections to those resources using the Internet.

User Interface and Operations

Central control of software and hardware operations are needed to ensure a uniform system interface and consistent operational functions (see figure 1). To achieve uniformity, guidelines for screen formats and information loading and processing are provided. Information can be in either ASCII format for online viewing or in binary format for formatted data, graphics, or downloadable software. This strategy proved advantageous for the launch of the service since in the past users had been confused by inconsistent information interfaces and retrieval processes.

The interface features option boxes across the bottom of the screen and with a single keystroke, a user can activate such services as "read," "search," "menu," "print," "help," or "exit." The service began with information resources internally processed and organized. Once the internal information had been loaded, the databases were expanded to accommodate specialized needs.

The Need to Diffuse Control

The videotex service was started with general purpose information such as news, events listings, and publications designed for the general public. It soon became apparent that the service might also function as a conduit through which scientists communicate, store, and share their data. Information management and operational functions were decentralized when it became evident that the types of information which would make the system evolve and grow were too specialized, technical, and sensitive for a central staff to administer.

Many of the databases contained information which changed continuously as new county, state or federal guidelines were established or as new research discovered better processes or materials. Legal responsibilities made it imperative that the scientists responsible for the accuracy of the data oversee the videotex information resources.

Information Management for Distributed Control

Data administration responsibilities were simplified to enable management by a scientist with little computer experience, or by the scientist's representative or secretary. As figure 2(next page) illustrates, the process consisted of first establishing an account and copying the system communication files into that account. Next, subdirectories were established to contain the executable files that would drive the database routines, such as the automated menu system.

Also located at this level were the databases containing the information files, the system files containing the path information, and the databases containing the full text search indices. In operation, the executable files would "look" at the information files, extract titles and topics for use in the menus, and index the words in the manuscripts for full text searching. Messages would flash on-screen during the execution to enable the scientists to watch for system errors. The need for debugging was rare, but when errors did occured, the information providers generally required input from the central staff.

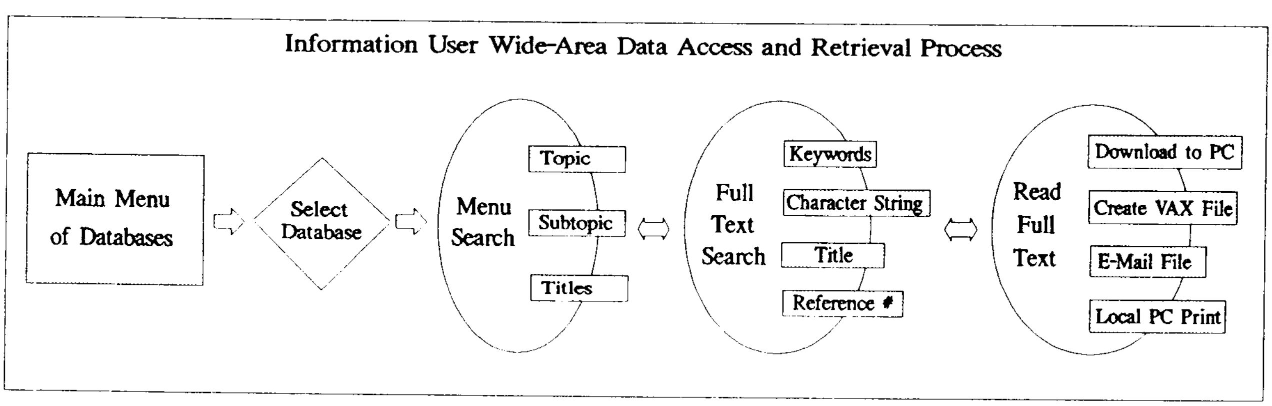

Figure 1. System interface and basic operating functions.

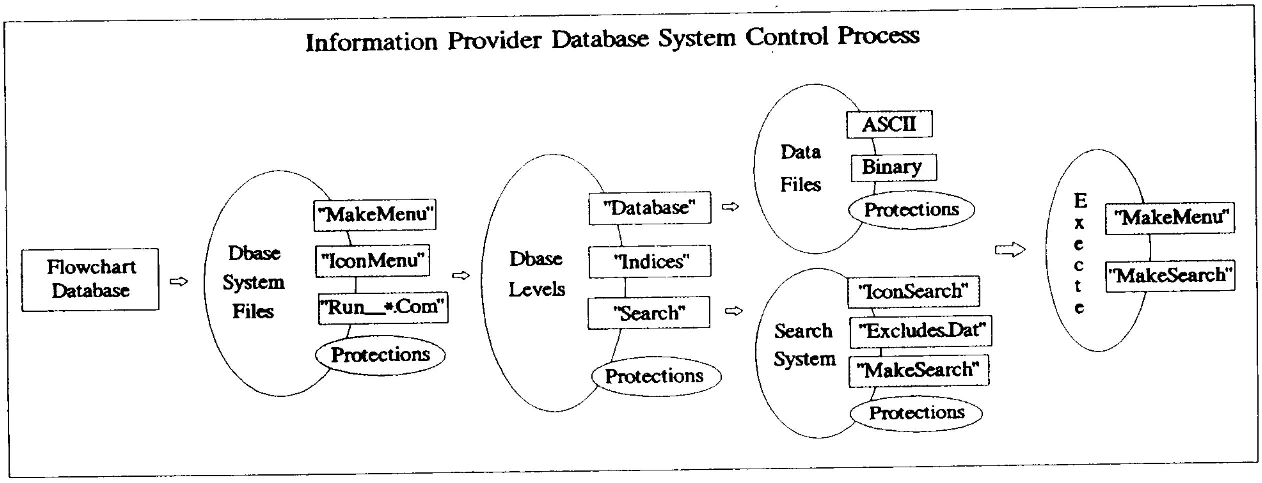

Figure 2. Information administration process for system management.

Training Processes

Information loading and managment processes were automated to ease the shift of information responsibilities from the central staff to the subject matter specialists. A training regimen was established. The training generally required 3-4 hours per week for 6-12 weeks. Those with knowledge of networking in VMS could master the process in 6 weeks, while a relative novice might require as long as 12 weeks.

In the first phase, authors were acclimated to the networking and formatting process. They were instructed to send their data, in the correct format, to the central administrative staff. The staff would build the database infrastructure. Authors were next brought into the teaching laboratory and trained in account management, at which time they learned the database structure and basic operating commands.

In phase 3, the scientists would work with the system files which control the database and provide the advance functions, such as full-text search, automated menu generation, and categorization of topic-subtopic indices. At this stage, authors would be able to fully administer the database on the primary VAX node. Finally, authors were trained to manage their databases on the remote systems and to access and process complimentary information gathered from across the Internet.

As the scientists became more involved with the service, they began to actively promote the system to their peers and to the users of their data. Thus, each new user helped enlist other users and build awareness of the system. Such a promotional effort was not possible with the centralized operation. This "pyramiding" of support and promotion became one of the most important factors stimulating statewide use of the service.

The Synergism of Complementary Resources

The overall strategy in implementing the service was to first build a platform of general resources containing information appropriate for a wide variety of users. Then, to locate a few dynamic resources which contained information sufficient to stimulate new users to learn the networks and applications. A menu containing all of the resources was visible upon initially activating the service to enable users to experiment with other databases. A daily news summary service provided capsule summaries of pertinent news stories from prominent newspapers around the world and was the resource which most successfully aroused interest and active participation. Electronic versions of print materials, mainly manuscripts, newsletters and reports, were added to satisfy organization-wide demands for ready access to 5,000 printed publications. The above resources were implemented by the central staff. Promotion occured through a newsletter, training session, and word-of-mouth. Next, interdepartmental and externally managed resources were added. A database begun in July, 1990, with information on federal research, education, and extension initiatives, was a cooperative project of the central staff and the unit responsible for program evaluation. Federal initiatives were referenced to corresponding state major programs, enabling users to track an initiative from the President or Congress, through the various federal agencies, to state and local implementation. An "instructional resources" database was initiated in August, 1990, with information on network resources for education including BITNET and Internet services, downloadable software, and a section for professional journals distributed electronically. In May, 1990, federally-mandated environmental protection media was added.

Important entomology and nematology databases were added beginning in November, 1990. These resources were the first externally managed databases and also the first cross-departmental applications. Management was by internationally prominent researchers from the main campus, several county offices, two research and education centers, and a state agency. The leading proponent and coordinator was a scientist with more than 10 years international videotex and database management experience who was very skilled in enlisting scientists and users, and in the overall promotion of the service. The databases were highly visible and featured prominently in newspapers, magazines, and trade journals throughout the southeast U.S.

Two complementary, dynamic databases were initiated by scientists in other departments and modeled after the entomology/nematology resources using a similar data structure (topics, subtopics, screen display). These resources were designed to cross-reference complementary databases. A passage in one database might refer uses to another resource for more specialized data, to a newsletter for more general but pertinent information, to a state or national news story, or to a pertinent publication in the reference encyclopedia. Using the reference code, a reader could move into this supplemental resource within 4-5 keystrokes. Cross-referencing was especially helpful since such associations would not generally be established during a full text search, and true hypertext linking would be too time-consuming for technical data which changed on a daily basis.

The Personnel Department began active participation in the service in early 1991 and quickly became influential database managers. The director had previous experience administering information services, and the associate director recognized the potential of online resources and became an avid promoter. Predictably, the "jobs" database, with listings of current state, federal and national openings, became a very popular resource and helped entice new users into the service.

Individual and Institutional Incentives

Pioneers invent new devices or initiate new processes, but it is the innovators who are instrumental in determining projects' future prospects through their practical implementation. In the above example, the innovators supported the system because it could be customized to perform to their specifications. And, because it enabled them to personally manage the resources thereby satisfying their professional responsibilities to perform research and to manage and disseminate data. Thus, a configurable system, capable of being modified, adapted, and administered by the information providers, helped ensure that a reasonable degree of benefit would accrue to the innovator. Needs of the casual users seemed to reside more with interface consistency and with matters related to the ease with which the service might be assimilated into daily work agendas. The reward of having one's information in an information service — even a large system seen by peers and clients throughout a state or region — seemed to carry only a moderate level of prestige. Rather, it was the act of being directly in charge of a major network resource from which the participants in this study exhibited the traits of pride in ownership. Instead of considering information maintenance an additional chore, the scientists (especially those with extensive experience in computer-based information delivery) recognized that distributed information management offered the best opportunity for mass participation. Training requirements, and the overall training program, were very well received. The instruction evolved from laboratory teaching, to online sessions with trainees working interactively with the trainer through the networks.

Conclusion

Information systems managers would agree that information and telecommunications technologies have not achieved their potential. Hardware and software have advanced faster than a user's capability to integrate the resources into the work day. In addition, the technology is advancing at such an accelerating pace that it is simply not economically feasible to continually increase the size of the central information systems staff to stay current with specialized applications, or to monitor exponentially expanding data resources. The premise throughout this paper is that the management of unstructured information may be delegated to those who work most closely with that information. A process was shown which was implemented and successfully distributed information management responsibilities. The system involved procedures to evolve users into information managers using software designed to automate data collection, processing, and dissemination. User training and promotional initiatives were very important.

R. William Maule, Ph.D
is Assistant Professor, University of San Francisco College of Professional Studies. He developed the IFAS service while an Assistant professor at the University of Florida.