THE EVOLUTION OF DOMESTIC INFORMATION DISPLAY SYSTEMS:

TOWARD A GOVERNMENT PUBLIC INFORMATION NETWORK

Edward K. Zimmerman
National Telecommunications &
Information Administration
Washington, D.C.

BACKGROUND: INFORMATION FOR GOVERNMENT DECISION-MAKING

Borne of a union of NASA and Census Bureau technology, the experimental Domestic Information Display System --DIDS--was designed to deliver small-area data in the form of statistical maps to policy makers in government using an interactive, full-color, high-resolution display system. Census tract, county, Congressional District and state statistics from Federal agencies are portrayed against digitized geographical boundaries from Census, using a computer-driven display system originally developed by NASA for study of satellite data. DIDS initial demonstrations were conducted in June 1978 in the Capitol and at the White House.

Following the successful introduction, DIDS was subjected to a one-year test period of evaluation, experimentation and research by an ad hoc group of 15 Federal agencies operating under combined White House and Congressional leadership, with management and services provided by the Department of Commerce and NASA respectively.

At one level, DIDS is an experiment in using an interactive geo-statistical display as a component of a decision support system. At another level, it is an experimental means of exchanging data among executive branch agencies to satisfy each other's needs, as well as to satisfy Congressional and White House requirements. At a higher level, it is a forcing agent for long-needed improvements in the Federal statistical system--changes which, while mandated by law[1], still have not been pursued as actively as many government leaders would have preferred.

The DIDS project seeks to distinguish itself from previous efforts by creating an institutional means of dissemination and exchange of statistical information, which will serve both legislative and executive branches of government, and which will survive changes of administration. It is felt that this can be done through careful matching of concepts, institution, and technology.

Concepts: Users Are Already Networked

The developing concepts of distributed data bases and computing power, combined with telecommunications for data networking, make it technically possible to leave the data bases where they are originated and maintained, and to communicate selected portions of them to the places where they are to be used. Indeed, that is how the decentralized Federal statistical system operates today, but the "network" is composed of heterogenous and inefficient media. Most statistics used for decision-making in the Federal government today move by hand through the mail and messenger network in paper form; or, in verbal form, face-to-face or via the telephone network. The question then, is not whether we use "networks", but rather which network technology we choose.

Public Information Network

Without changing the basic structure and behavior of the Federal statistical system, the provision of a homogenous, modern communication network to move these data in their native digitized form would have an enormous positive effect on the performance of the system. Such a government public information network would make the

[1] For instance, see the Congressional Budget and Impoundment Act of 1974 (PL93-344). Title VIII amends Title II of the Legislative Reorganization Act of 1970, with Sections 201-203 being relevant.

outputs of the Federal Statistical System more available, constructively, to client agencies in Federal, state and local government than they are now.

Common media would facilitate--in fact, demand--various levels of de facto standardization among statistical data providers and consumers. Such standardization is achievable, as the DIDS project has demonstrated. The initial data exchange formats have evolved over the past year, and will continue to evolve. Perhaps it is best not to think in terms of enforced official standards in the near future. After all, some of our more useful standards (e.g. computer programming languages) were de facto for several years before formal adoption. And, networked decentralized computer systems have inherently efficient data conversion capabilities: given a public information network to support DIDS data exchange, it would not be necessary to mandate onerous file format standards to all data contributors.

Minimal Common Services

What must exist centrally, however, is a central indexing capability. Just as our telephone network would be of little use without phone books and information operators, a data network for statistics must have a continuously updated catalog of its resources and their vital statistics, which include bibliographic data, pedigree, and caveats regarding usage and reliability.

Another concept must be considered to respond to the need for frequently requested standard reference data. Main libraries have larger collections of reference books than branch libraries, as it is easier to fill multiple and frequent requests for the same information from one central point. This is analogous to the standard time and weather numbers in telephone central exchanges, and to the store-and-forward messages switches in the TWX/Telex system. The initial DIDS system is an example, with its ground rule that only publicly-available data could be catalogued and maintained on the system.

Applying Common Sense to the Cost Problem

In all such situations, the reason for keeping the information in a place other than its origin (e.g. at the National Weather Service) is to reduce the cost of communicating it to its frequent users. On the one hand, one must consider the frequency of updating the information, and the differences in cost in updating it once (either at the point of origin or at a central point), or at every point maintaining a copy for usage.

On the other hand, one must consider the cost of communicating information among all these points. It is in this area that we see the key advantage of networking: the significantly lower communication costs of contemporary data networks suggests that, once the majority of information sources and sinks are served by the net, it will be cheaper to leave all information at its point of origin, regardless of frequency of access, or usage at other points.

Interpreting this in the context of the DIDS project, it seems that it will be necessary initially to maintain the most popular and frequently-used data in one host computer, equally available to all participants. The importance of this virtual network should decline as more and more providers and users join the actual network serving DIDS[2].

Decentralization Allows User Responsibility

The strongest argument for a distributed system relates to the needs for privacy- confidentiality or security requirements for information. With each network participant possessing and being responsible for his own terminal facility, including a switch to absolutely disconnect it from the network while handling sensitive information locally, the game can be played by the present rules. This is not possible when users are forced, by centralized systems, to communicate sensitive information outside the confines of their own controllable (and thus presumably benign) environments.

Much information handling for decision-making purposes would require access to adjunctive data to supplement or qualify publicly-available information. In several

2 Other parallel technical developments will tend to support the trend toward decentralization. Among them are the datacomputer with virtual associative processing capabilities, large and cheap mass storage (e.g.$150 for 10^{11}bits), and the concept of remotely executable search algorithms.

of his speeches and draft papers, Richard Harden[3] has discussed the concept of three levels of information used in decision-making: the first, which might be thought of as raw data, is the basic statistical imput to the policy process; it is (or could be made) publicly-available, and should be equally available to all stakeholders in the policy process. The second, which might be thought of as policy information, is the work product of policy analysis,and typically relates to the options being prepared for the top decision makers. The third, which might be thought of as political determinations, are the work product of the political decision makers during the usually-reiterative (and often agonizing) final phase of the policy process. Prior to announcement of the decision, the political determinations are the most sensitive and thus least able to be shared.

Conceptually then, the proposed system must be able to support the above behavior. Functioning as a terminal on a network, the policy analyst's information handling facility must be able to reach out for raw data, reviewing and selecting from a rich variety of remote public resources. Once the raw data is gathered, the analyst may wish to disconnect from the network on order to try out various private assumptions with the public data, perhaps in common with relatively small files of policy-related information maintained locally. The product of this effort would be the policy options or recommendations.

In large organizations, the policy analyst's information handling facility might be one of several on a local area network, in which case the policy information might be forwarded to a separate internal facility reserved for use by the top decision makers (1,2). In smaller units, paper outputs would be produced at this point, and conventional handling would suffice.

How does the Private Sector Help?

One more concept remains to be dealt with: private sector participation. There are at least three good reasons why private-sector organizations should be

[3] Special Assistant to the President for Information Management, and chief White House sponsor for the DIDS project.

involved in the next phase of DIDS. First, the private sector is a rich source of both information and information technology, both useful to DIDS and its clientele. Second, serving the needs of state and local government and other interested non-federal users might best be done outside Federal Government. This would be consistent with the Administration's general policy of encouraging a marketplace economy for information. Third, and also consistent with current policy, the provision of the data communication portions of the services by a commercial common carrier, at arm's length from both the government data providers and users, might be more politically palatable than past approaches such as FEDNET.

The appropriate level and means of private sector participation in the project is a very important consideration, and one which thus far has had little discussion. This discussion has begun, and is an early and high-priority activity of the DIDS FY1980 program.

INSTITUTION: WHO SHOULD DO IT?

At the present time, it seems to be logical and consistent with recent policy decisions to operate the DIDS project from the Department of Commerce. The Office of the Chief Economist and its subsidiary Office of Federal Statistical Policy and Standards (OFSPS) have been full partners in DIDS since the beginning of the project. DIDS goals are for the most part goals supported by OFSPS (3), or are also espoused by a current policy document on the subject of improvements to the Federal statistical system (4)--the thrust of which is to establish a central statistical organization (CSO). While some observers may not agree with the Bonnen study's recommendation for institutional arrangment for the CSO, far fewer critics argue about the substantive duties and responsibilities it should have.

The most reasonable assumptions to make at this time, pending action on the Bonnen study recommendations, are that OFSPS is currently the best vehicle for DIDS; that any CSO establishment would include the OFSPS function; and that whatever is spawned from the DIDS initiative would be the proper business of the CSO.

TECHNOLOGY: HOW CAN IT BE DONE?

The capability to support interactive full-color displays of geographically-based information is the most distinguishing feature of the DIDS system. One problem with DIDS is that the technology required to support the mapping is, at present, too expensive to justify its wide-spread use.

Fortunately, two developments will help with this problem. First, the cost of the necessary computer, storage and display hardware components is dropping dramatically. Second,DIDS experimenters are discovering that the computer system required for interactive statistical map displays is suitable for a host of other data display, sharing and manipulation services useful in a decision support system.

The key technical goals of the project, then, must be 1)to reduce the cost of the mapping capability to a sufficiently low level as to make practical its use in a multiple-location, decentralized fashion; 2)to integrate the mapping capability with other complementary tools to make a more generally useful decision support system;3)to identify means of sharing data resources among the participants; 4) to provide an appropriate host facility for common functions and data,evolutional standard development, message switching, indexing and bibliographic facilities, auditing and accounting functions; and 5)to provide for external connections in order to receive data from non-participating sources and to provide data for further distribution--say, via private-sector organizations. In short, the technical capabilities designed, developed and used by the project must support the concepts outlined earlier in the Concepts section of this paper.

What follows is a discussion of how an inter-agency communications system for DIDS might evolve into a public information network. This is not intended to be a complete technical description. Experts in such areas as data management systems must be called upon for complementary contributions, for instance on the subjects of maintenance of central indices of decentralized data contributions and storage management and migration control techniques.

Evolution History

The DIDS system development has been and should continue to be a phased evolution. The accompanying series of diagrams depict that evolution. All but the most essential detail is removed, in favor of making the major connections obvious. Straight lines represent local hard-wired connections; jagged lines depict telecommunications to remote locations. In the following paragraphs the technical history of the project is reported, with elements of future planning interspersed to show how these plans evolved.

Figure la represents the DIDS system as it existed in June 1978, just prior to the time of the initial formal demonstrations. It was essentially unchanged (but for software enhancements and data) from the NASA Atmosphere and Oceanic Information Processing System, AOIPS. It consisted of a DEC PDP 11/70 computer and associated peripherals, equipment, several specialized devices which weren't used for DIDS, and a tailor-made image analysis terminal made by Hazeltine to NASA specifications, called IAC-2. The initial demonstrations on Capitol Hill and at the White House involved specially-balanced video lines to carry the high-resolution video picture image from the IAC-2 terminal at the Goddard Space Flight Center (GSFC) , in Greenbelt, Maryland, to a downtown slave monitor; a 30-megahertz[6] capacity trunk was used for these June 1978 demonstrations, and a high-quality display resulted. As the telecommunication cost for this demonstration was about $40,000, it seemed unfeasible at the time to consider usage of video frequencies of this order to support a downtown DIDS terminal. So, with both computer and its television-like terminal located at NASA Goddard, initial FY79 DIDS experiments and demonstrations required users to make a 36-mile round trip trek from downtown Washington, D.C.

In order to increase usage, the project sought a solution to the problem. Among ideas considered was hooking up <u>all</u> major downtown Federal buildings with a closed-circuit television system--cable TV for the Federal triangle. While there are many good reasons for doing this, cost and lead time were significant

[6] 39 million cycles per second--a capacity equal to 12,000 tlephone lines.

DIDS Hardware Development - Phase I

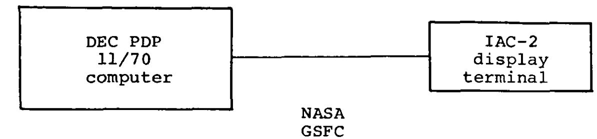

Fig.1a: Original DIS system, as of June 1978.

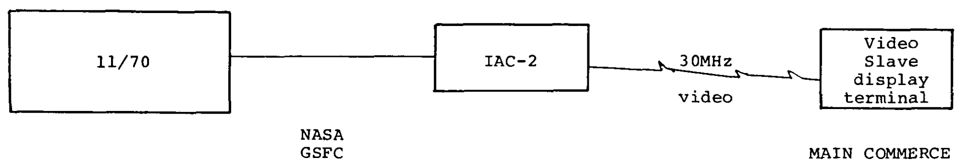

Fig.1b: DIDS at June 1979, with a video-based interim downtown terminal. A similar arrangement was used for the original demonstrations in 1978.

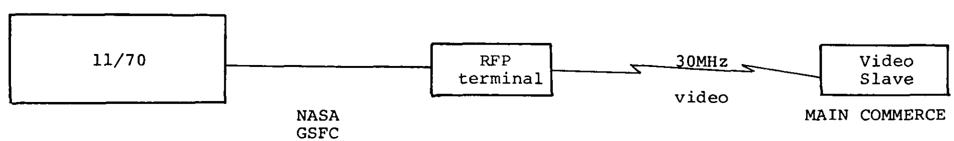

Fig.1c: Due to be operational by January 1980, this evolutionary step includes an inexpensive replacement for the specially-built Hazeltine IAC-2.

deterrents. But while our thoughts focused on a video-based solution, evidence of possible alternatives based on different approaches began to appear in the marketplace. For a time, it seemed that a 'quick and dirty' free-standing system, roughly equivalent to the DIDS system then in use, could be constructed for a cost and in a time frame that would be less on both counts than ordering up a 39-MHz line for a six-month experiment.

As time passed, thinking progressed, and it was ultimately decided to proceed with a video-based solution as an interim measure, while a more carefully considered development of a free-standing system proceeded. Figure 1b represents the result, as of June 1979. Three 10-MHz lines were used to convey a tolerable quality of live video from the IAC-2 at Goddard to a slave terminal located in the Main Commerce building, at 15th and E Streets, N.W. in Washington. As in the initial downtown demonstration, another telephone line was used to support communications between a regular alphanumeric terminal used for system control and the host PDP 11/70 computer. As its small capacity pales by comparison to the 30MHz trunk, this line is not depicted here.

The first step in designing a free-standing system was to reduce the cost. The jump of upgrading from the 11/70 computer to a more modern and suitable off-the-shelf design was felt to be far easier than the transition from IAC-2 to a state-of-the-art color graphics terminal of equivalent capacity. This thinking led to an initial procurement program for a logical replacement for the IAC-2, depicted in Figure 1c as "RFP terminal," due to become operational by January 1980.

Phase Two: A Separate System

At this point in the project, several different forces were driving the technology. Inexpensive solutions to the design problems in a free-standing terminal didn't come as quickly as burgeoning demands for more AOIPS time. NASA was committed to use the PDP 11/70 system on which DIDS was based to support severe storm research in FY1980. It was this motivation which supported the Request for Procurement for the IAC-2 terminal replacement, and to the substitution of a second host computer for the original 11/70.

This reasoning, plus a series of developments in the availability of micro-computer hardware and software led to definition of phase two of the DIDS project hardware program, as depicted in Figure 2a. A DEC VAX 11/780 computer would replace the 11/70, the RFP terminal would replace the IAC-2, and for the time being, the video slave terminal would continue at Main Commerce. These steps are targeted for accomplishment by March 1980.

The micro-computer developments led to a hypothetical design for an economical stand-alone system (5), depicted in detail in Fig. 2b. Presumably starting with a second copy of the RFP terminal, a suitable microcomputer such as the DEC LSI 11/23 would be used to provide a slimmed-down version of the original 11/70 capabilities. This seemed much more feasible after a series of announcements regarding available software for the LSI 11/23.

The next component, floppy disk storage units, are to provide the principal means of emancipation from the centralized concept. Consider the analogue of a classroom in which each student sits at his desk with a map outline and a box of crayons. Students color the maps according to the instructions from the teacher.

In our model, the floppy disk replaces the map outline. To replace teacher's instructions, we depend on statistics available through communications from their original sources. To avoid carrying live high-resolution video, we decentralized the maps on floppy discs, and communicate only the attribute information, socio-economic and demographic statistics. These change with regularity,whereas the geography does not. Thus a stand-alone system for DIDS applications would possess two kinds of storage: one for the geographic base file currently under consideration, and one for the attribute (i.e. statistical) data relevant to that geography. Work on this design has commenced, and should result in the first working prototype of a stand-alone system by May of 1980.

Such a system is depicted in the lower part of Fig.2c at the instant in time, planned for June 1980, during which the downtown video slave terminal is being replaced by the first copy of the stand-alone system. This will be a very significant event in the DIDS project for two reasons. The first, and least important, is that

DIDS Hardware Development - Phase II

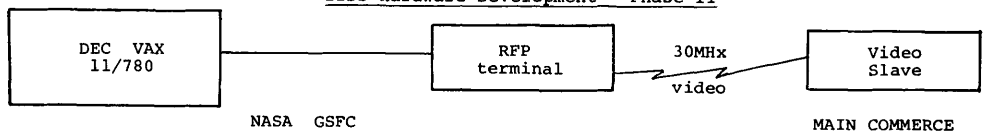

Fig.2a: A more modern computer is to replace the original DIDS machine. Conversion from the PDP 11/70 and operation with the RFP terminal should occur by March 1980.

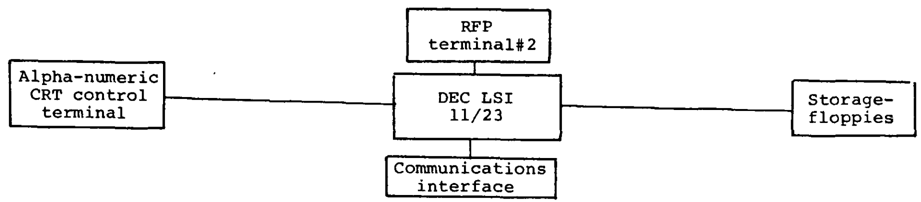

Fig.2b: Reproduction of original sketch of stand-alone DIDS system. Such a configuration could privide the basic capabilities of the DIDS system, at a greatly-reduced price. Prototype should be available by May 1980.

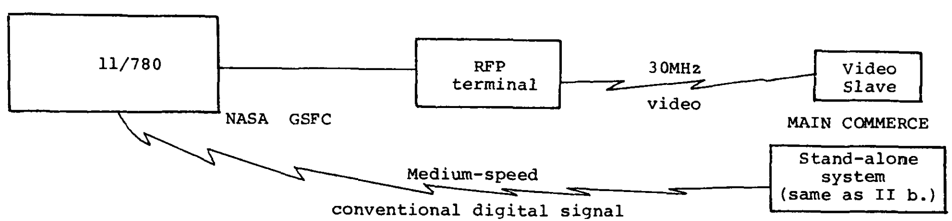

Fig.2c: June 1980 is the planned date for side-by-side tests of the original video-based terminal and the new stand-alone system. When these tests are successful, the video slave terminal connection will be discontinued.

when success is achieved, the project can discontinue the 30MHz video line and its $13,000 per month price. The second will be the evidence of a successful sybiosis between an independently-operating DIDS stand-alone remote system and its occasionally-supportive host system, the VAX 11/780.

Future Plans-Phase Three

Depending on funds availability, perhaps at that point other arrangements might be made to foster the progress toward full networking. For instance, TELENET or ARPANET connections on the network could facilitate data exchange and electronic mail outside the immediate DIDS community. Another idea is to build electronic bridges to alternate system approaches--for example, a communication facility between the DIDS VAX 11/780 and a similar system used for geo-based graphics at the George Washington University. Another is to begin an alternative terminal procurement--with different hardware components, but exactly the same imput and output specifications as the first effort. An important benefit of moves such as these would be to give incentive to private industry and academia to support the effort.

All these goals are consistent with the FY1980 DIDS project objectives of providing an operational capability. Figure 3 depicts, in conceptual form, one way in which the basic capabilities developed in 1980 can be used in 1981 and beyond to meet the perceived user needs. While a ring network may seem to be suggested by the diagram, this should be regarded as a perhaps temporary fascination of the writer--other architectures might suffice as well. The nodes labeled Sn depict stand-alone systems who continuously or occasionally connect to the network for exchange of geographic, statistical, or other information relevant to local decision-making needs. Nodes labled Si are independent stand-alone systems (or logically-equivalent decentralized information handling facilities) which, while they are dependent on the community of network correspondents for advice and information, choose not to be directly attached to the network. Perhaps the "network" exists for them by exchange of floppy disc media!

DIDS Hardware Development - Phase III

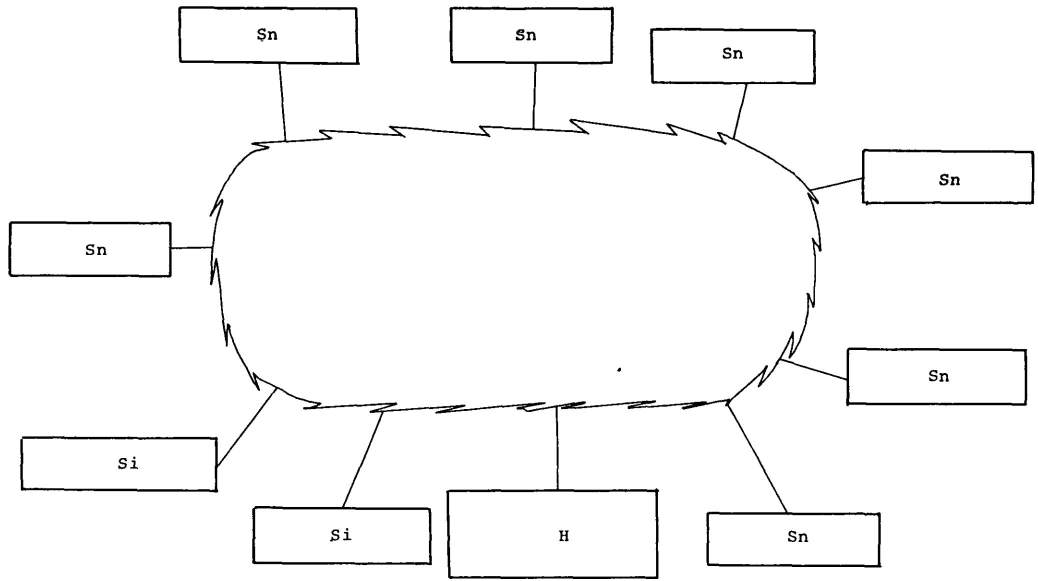

Fig.3: DIDS in 1981 and beyond, in the context of a public information network. Sn are attached network nodes which are DIDS stand-alone systems and Si are independent, more loosely-coupled but logically equivalent systems. H is the best computer used for community-wide functions.

The node labeled H depicts the DIDS community-wide host computer node, or the message-switch, with duties and responsibilities as described earlier under comments on the need for a minimal and decreasing set of centralized support functions.

Little can be said at this point about how such a government public information network is to be realized. Dedicated or dial-up telephone communication ports on the VAX 11/780 might suffice at the beginning. The author's fantasy is an ARPANET-like approach, complete with software exchange and electronic mail facilities. The project participants should not wait for development of Federal policies and standards to try such promising appproaches to the problem of data sharing for decision making. Indeed such trials can provide, by example, motivation for development of appropriate standards.

Summary and Conclusions

In this paper, I have cited the accomplishments achieved so far by the Domestic Information Display System. It is recommended that three components--conceptual, institutional, and technological--be matched so as to attain a higher level of Federal Government DIDS operations. I regard further research on goal definition and planning valuable; I have attempted to provide a rationale for the establishment of a government public information network.

One fundamental goal should not be abandoned: Congress and the White House should share the same basic resources for decision making information, and should continue to cooperate in joint demands for such resources. Such a notion does no violence to separation of powers, and in fact should foster competition for ideas as how to govern our country better. If the Executive-branch agencies and other needed sources of information continue to witness such joint resolves in this area, it will have a salutary effect on the Federal statistical system, and thus on the present and future quality of decision making in the United States.

References

1. Harden, Richard M.
1979 "Management of Information in the Executive Office of the President" (unpublished), May 10.

2. Duda, James L.
1979 "The Future of DIDS in the Department of Transportation - DIDS Future Group Think Piece" (unbpublished), October.

3. Department of Commerce
1979 "A Framework for Planning U.S. Federal Statistics for the 1980's", Office of Federal Statistical Policy and Standards, July.

4. PRP Statistical Reorganization Project
1979 "Improving the Federal Statistical System", 27 pp. (unpublished draft), June. (also known as the Bonnen Report).

5. Zimmerman, E.,
1979 "Free-Standing DIDS", Memorandum, April 24.