REPORT TO THE CONGRESS

Emphasis Needed On Government's Efforts To Standardize Data Elements And Codes For Computer Systems *B-115369*

BY THE COMPTROLLER GENERAL OF THE UNITED STATES

MAY 16, 1974

B-115369

To the President of the Senate and the
Speaker of the House of Representatives

This is our report on the emphasis needed on the
Government's efforts to standardize data elements and codes
for computer systems.

Our review was made pursuant to the Budget and Account-
ing Act, 1921 (31 U.S.C. 53), and the Accounting and Audit-
ing Act of 1950 (31 U.S.C. 67).

Copies of this report are being sent to the Director,
Office of Management and Budget; the Director, National
Bureau of Standards; and the heads of all other Federal
departments and agencies.

Comptroller General
of the United States

Contents

APPENDIX

ABBREVIATIONS

ADP	automatic data processing
ANSI	American National Standards Institute, Inc.
CSC	Civil Service Commission
DOD	Department of Defense
FIPS PUBs	Federal Information Processing Standards Publications
GAO	General Accounting Office
GSA	General Services Administration
HEW	Department of Health, Education, and Welfare
NBS	National Bureau of Standards
OMB	Office of Management and Budget

*COMPTROLLER GENERAL'S
REPORT TO THE CONGRESS*

EMPHASIS NEEDED ON GOVERNMENT'S
EFFORTS TO STANDARDIZE DATA ELEMENTS
AND CODES FOR COMPUTER SYSTEMS
B-115369

D I G E S T

WHY THE REVIEW WAS MADE

GAO made this review to assess progress being made under the Federal program to standardize data elements and codes used in computer operations.

Such standardization could help reduce high costs of Federal computer operations by eliminating unnecessary duplication and incompatabilities in collecting, processing, and disseminating data.

FINDINGS AND CONCLUSIONS

The Government is spending billions of dollars yearly to manage, support, and maintain 6,200 computer systems. Despite advances in computer technology, most data is exchanged between Federal agencies in hard-copy form (paper). (See p. 7.)

Data collected for computer-based systems can be used in many ways. Once collected and deposited in a Federal computer system, it can be transferred from one agency to another. Rarely is data transferred in total to another Federal computer system. However, it can be transferred if originally collected and recorded in a standardized, agreed-upon fashion. (See p. 5.)

After the collector has converted data to machine-readable form, it can be exchanged automatically by magnetic tapes, punched cards, storage disks, computer and communication networks, and so on.

However, before machine-readable forms can be used in the exchange process, the data must be given a standard description, identification, and code.

Data transferred between Federal agencies should be handled in machine-readable form, whenever possible, to take advantage of this method's speed and accuracy, to reduce costly paperwork, and to reduce duplicate and unnecessary data collection.

A Government-wide data standardization program was begun in 1965 by the then Bureau of the Budget to achieve the greatest practicable degree of uniformity of information used among and within data systems. (See p. 10.)

Since that date, Federal efforts to standardize data elements and codes have been slow and not very successful. The Office of Management and Budget (OMB) has given low priority and limited resources to the standardization program and, consequently, has formulated and approved few data standards for Government-wide or programwide use. (See p. 14.)

As a result, the initiative for developing, implementing, and using data standards has been left to individual Federal departments.

Before agencies will commit resources to the standardization program,

Tear Sheet. Upon removal, the report
cover date should be noted hereon.

1

greater central guidance is needed so that their products can be made comparable and applicable Government-wide.

The President transferred responsibility for this program to the Secretary of Commerce in May 1973.

In November 1973, after GAO completed its review, the Secretary of Commerce issued a policy directive which clarified and changed the management and operation of a standards program for data elements and codes. (See p. 17.)

Significant policies established under the new directive include:

--Clarification of terms and new terms to describe data standards.

--Two additional types of standards.

--Priorities for ranking types of standards.

--Responsibilities of the Department of Commerce and other departments and agencies.

Issuance of this new policy directive is a step in the right direction and should help the data standardization program. However, GAO does not foresee a significant amount of automated exchange until Commerce takes additional measures to see that this policy is effectively implemented.

RECOMMENDATIONS

To accelerate development and use of standard data elements and codes, the Secretary of Commerce should:

--Determine where standards would be most beneficial and establish standardization priorities.

--Issue policy delineating accepted theory and terminology and provide for preparation of guidelines, methodology, and criteria to be followed by agencies in their standardization efforts.

--Assign to specific agencies responsibilities for developing standard data elements and codes in specified areas.

--Monitor implementation of data standards to insure their uniform adoption and use. (See p. 33.)

AGENCY ACTION AND UNRESOLVED ISSUES

The Secretary of Commerce in February 1974 said the Department considers standards for data elements and representations to be one of the highest priorities within the Federal processing standards programs.

The Secretary pointed out that its recently approved regulation establishing a Government-wide program for data element standardization embodied recommendations contained in the GAO report. (See p. 33.)

Commerce noted that early responses to the new directive are favorable, but success of the standards program is dependent on support and cooperation from other Federal departments and agencies. GAO agrees that Commerce needs agency help and support and is sending a copy of this report to each Federal agency head along with specific suggestions which should help the Department of Commerce in its data standardization program. (See p. 34.)

MATTERS FOR CONSIDERATION BY THE CONGRESS

GAO is sending this report to the Congress because of:

--Continuing congressional interest in the management and use of computers.

--The necessity for developing standard classifications for information to meet the needs of all branches of Government.

--The specific interest of the Joint Economic Committee and the House Committee on Government Operations.

CHAPTER 1

INTRODUCTION

Although Federal agencies need current and accurate information, collecting raw data and converting it into machine-readable form is expensive and time consuming when done manually. The National Archives and Records Service, General Services Administration (GSA), has estimated the cost of Federal data collection activities to exceed $5 billion annually.

As more Federal information systems are automated, it becomes apparent that computer-based systems offer their users a great opportunity to make many uses of data collected.

Rarely is data transferred in total to another Federal computer system. However, data can be transferred if originally collected and recorded in a standardized, agreed-upon fashion.

When more than one agency needs the same data, that which is already collected and deposited in one Federal computer system can be transferred to another, eliminating the need to duplicate the collection and conversion process. Once the data has been converted to machine-readable form by the original collector, this data can be exchanged automatically in the form of magnetic tapes, punched cards, storage disks and so on. Such exchanges are especially desirable when data can be put directly into another computer-based data system.

We recognize that other standardization efforts (such as hardware, software, and application) may be needed before a completely automated exchange can take place in all circumstances. Pending the completion of those efforts, however, the data standardization program can make significant advances. The direct input of data into another system can be effectively accomplished only if both systems use the same data codes, which is unlikely so long as agencies design and operate information systems independently. Therefore, data standards must be developed for intersystem use.

DEFINITIONS

A data element is a group of informational units which has a unique meaning based on a natural or assigned relationship and subcategories (data items) of distinct units of value. A data code may be the number, letter, symbol, or any combination of these used to represent a data element or item.[1]

Data standards or standardized data elements are descriptions, identifications, and rules, established by authority, for using data elements. For example, standard codes have been established for the States, as shown in the following example.

Data element	Standardized codes	
	Alphabetical code	Numerical code
State of residence:		
Alabama	AL	01
California	CA	06
New York	NY	36
Wyoming	WY	56

EXAMPLE OF PROBLEM WITH NONSTANDARDIZED CODES

Communication barriers resulting from differing agency codes make it difficult and often impossible to consolidate data from different information systems. For example, the Civil Service Commission (CSC) found that agencies could not economically and accurately comply with its seemingly simple request for the total number of Government employees of each sex.

Agencies were asked to code males "1" and females "2" and to provide the data on magnetic tape. That data was readily available, but the agencies defined and coded the

---

[1]The definitions for data elements and codes as contained in an Office of Management and Budget (OMB) circular were revised by The National Bureau of Standards (NBS) in November 1973.

data differently. For example, agency A combined sex data with education data, agency B combined sex data with marital status data. Agency C simply recorded Mr., Miss, Mrs., and Dr. It arrived at sex statistics by assuming that all or most doctors were male.

CSC could not obtain the data it needed without going through a costly manual operation to convert the nonstandard data. As a result, CSC initiated a program to standardize data in the Federal personnel systems.

The example on page 8 shows the problems encountered, while the example on page 9 shows how standard data could have removed the barrier and provided a greater capacity for exchanging data between systems.

FEDERAL AGENCIES EXCHANGE DATA PRIMARILY IN HARD COPY

In fiscal year 1971 GSA completed a survey of 14 civil departments and agencies to measure existing and potential needs for data exchange as a basis for designing nationwide communications networks and systems. The information collected identified how reports were exchanged (manually or in machine-readable form).

Federal agencies have made little progress in exchanging data in machine-readable form for direct input to computers. The vast majority of the survey reports--52,200 of 54,000, or 97 percent--were exchanged in hard-copy form (paper). Of the remaining 1,800 reports, 1,250 were sent in an automated form (punched card, magnetic tape, and so on) and 550 were partially automated.

The departments and agencies reported that they planned to automate 1,200, or only 2 percent, of the hard-copy reports identified in the survey. Most of the reports to be automated were intradepartmental.

UNCOORDINATED DATA DEVELOPMENT INHIBITS DATA EXCHANGE

AGENCY A   AGENCY B   AGENCY C

DATA   DATA   DATA

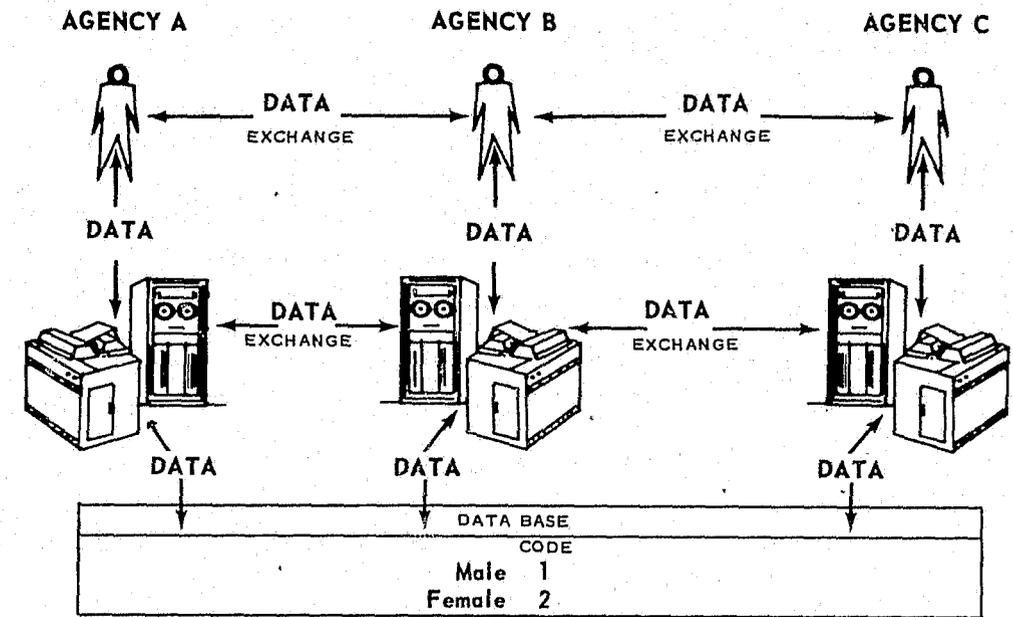COMMUNICATION barrier   COMMUNICATION barrier

DATA   DATA   DATA

DATA BASE	CODE	DATA BASE	CODE	DATA BASE	CODE
Male–college graduate	1	Married–male	A	Male	Mr
Male–other	2	Married – female	B	Female	Mrs Miss
Female–college graduate	3	Unmarried–male	C		
Female – other	4	Unmarried–female	D	Other	Dr
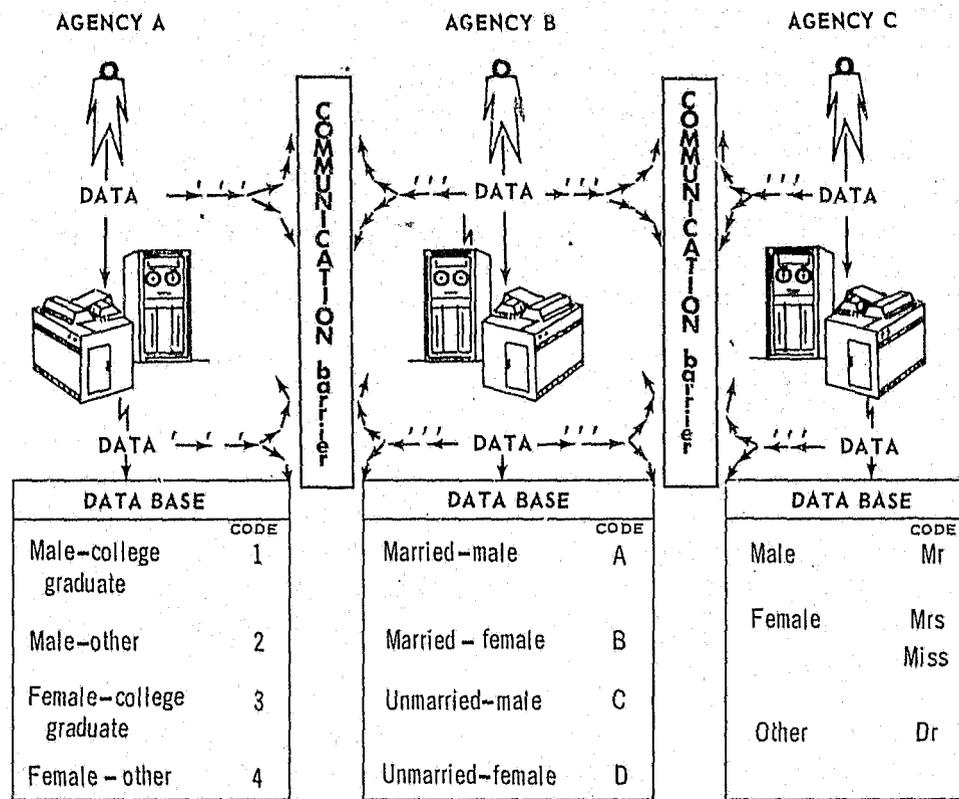
This example shows that the data elements and codes on the sex of government employees was not standardized and prevented CSC from interchanging this data in low-cost, machine-readable form.

DATA STANDARDS FACILITATE DATA EXCHANGE

AGENCY A   AGENCY B   AGENCY C

DATA EXCHANGE   DATA EXCHANGE

DATA   DATA   DATA

DATA EXCHANGE   DATA EXCHANGE

DATA   DATA   DATA

DATA BASE	CODE
Male	1
Female	2

This example shows that, when data elements and codes are standardized, the data can be easily exchanged between computer systems in machine-readable form.

CHAPTER 2

FEDERAL EFFORTS TO

STANDARDIZE DATA ELEMENTS AND CODES

EVOLUTION OF FEDERAL PROGRAM

In March 1965 the President of the United States approved and sent to the Congress a report on the Government's management of automatic data processing (ADP). The comprehensive study by OMB, formerly the Bureau of the Budget, covered various aspects of ADP selection, acquisition, and use, including the need to develop data standards.

The study recommended that OMB assume overall leadership of a program to standardize common Government data elements and their codes. Our report, "Management of Automatic Data Processing Facilities in the Federal Government" (B-115369, Aug. 31, 1965), agreed with that recommendation but pointed out that the standardization efforts involved interplay of data systems in Government, industry, and agency ADP efforts and would require much more central authority than OMB could provide. We felt that a central office in the Office of the President, which would have appropriate authority and responsibility for providing ADP management coordination, was needed to improve the standardization efforts.

In March 1965 OMB issued Circular A-71, delineating executive branch responsibilities for administering and managing ADP activities. OMB retained responsibility for supporting the development and use of standards for data common to the executive departments.

The Brooks bill (Public Law 89-306) was enacted in October 1965 to provide for the economical and efficient acquisition, operation, and use of ADP equipment by Federal agencies. Citing this law as a basic authority, OMB issued Circular A-86 on September 30, 1967, to initiate a formal program for standardizing data elements and codes used in computer-based information systems.

OBJECTIVES OF FEDERAL PROGRAM

The objectives of the program, as set forth in the circular, were to achieve the greatest practicable degree of uniformity of information used among and within data systems and to:

--Help summarize and exchange information.

--Help review and analyze budget processes and programs concerning more than one agency.

--Promote systems integration for communicating information among data systems without interrupting the process for translations or conversions.

--Contribute to improving the products and effectiveness of data systems.

RESPONSIBILITIES FOR
DATA STANDARDIZATION PROGRAM

Circular A-86 allows for the following three types of data standards for the Federal Government.

"--Federal general use

Standards (such as for countries, states, counties, places, organizations, * * * and elements of time) for general use by most agencies in connection with an extensive number and variety of related or unrelated data systems and programs.

"--Federal program use

Standards for use in particular related programs concerning more than one agency. Examples are elements and codes usually limited to use in weather, personnel, supply, and other similarly unique systems. In these cases, the same source data often are used by several agencies and aggregation and exchange of information on a program basis are the rule.

"--Agency standards

Standards limited for use within the programs of a particular agency and either not applicable to or not yet incorporated into a Federal standard."

The responsibilities for, and participation in, developing these three types of standards are listed in the following table.

Types of actions	Federal general use	Federal program use	Agency standards
Provide leadership in program	OMB	OMB	OMB
Arrange for development and maintenance of standards	OMB	OMB or program agency	Agency
Develop standards	Government task force	Interagency effort	Agency
Approve standards	OMB	Program agency	Agency
Maintain standards	Agency named by OMB	Program agency	Agency
Publish standards	NBS	Program agency	Agency
Register standards	NBS	Program agency	Agency
Publish and maintain register	NBS	NBS	NBS and agency
Evaluate and approve requests for exceptions and deferments of standards	OMB	Program agency	Not applicable

Circular A-86 also directed that OMB prepare guidelines and criteria to assist agencies, task forces, and equivalent groups in developing standards.

Federal departments and agencies were instructed to:

--Keep OMB informed on the progress and results of efforts to develop, announce, implement, and maintain program standards.

--Recommend to OMB areas having potential for developing standards.

--Recommend to OMB actions to improve the program.

--Designate a single office as a central contact point on matters pertaining to this circular.

NBS was assigned the responsibility for providing technical advice and assistance to task forces or equivalent groups established to develop Federal general and program standards.

STATUS OF DATA STANDARDS PROGRAM

The Federal program to standardize data elements and codes in information systems has been slow and not very successful. Only six Federal general standards have been issued since 1965, and many agencies do not use all of these standards.

Although some agencies have developed and adopted standards, the overall effort has been hampered by major obstacles in policy direction, approach, and guidelines.

OMB has given the data standards program a low priority and has committed only limited resources to it. Essentially, the initiative for developing, implementing, and using data standards has been left to the Federal departments.

OMB did authorize the development of a handbook for data standardization. A task force completed a draft of it in 1969, but it was never finalized and officially released. Several Federal and non-Federal representatives who reviewed the draft presented a wide range of opinions on the document's adequacy. OMB believed it did not contain the essential ingredients needed to develop data standards. Almost all representatives, however, believed that a document, such as the handbook, was needed to foster a uniform approach to data standards development.

Department of Defense efforts

In the early 1960s the Department of Defense (DOD) instituted a series of standardized procedures, including data standards, to facilitate automated processing of large volumes of transactions. DOD adopted a formal program for data standardization in 1964.

Several directives and instructions pertaining to the DOD data standards program have been issued since 1964. Prime responsibility for administering the program was assigned to the Assistant Secretary of Defense, Comptroller. He was given authority to delegate to selected DOD components responsibility for identifying potential for data standards in areas related to their missions. These component agencies were responsible for reviewing data and establishing proposed standards, coordinating them with appropriate agencies, and maintaining them after they were approved and published by the

Assistant Secretary. The agencies were also required to submit quarterly reports to the Assistant Secretary on their progress.

DOD has developed and promulgated over 1,300 data standards. DOD agencies have been directed to use them in their data systems to the maximum extent practicable.

Civil departments' efforts

Representatives of the Veterans Administration; the Department of Health, Education, and Welfare (HEW); and the Department of Labor advised us that they had not issued policy guidelines and instructions to their component agencies on the Government-wide program. They indicated that OMB must first provide additional guidance and that top management must be made aware of the need for, and the benefits of, data standards to gain their active support for the program. However, these departments have taken the following actions to implement some aspects of a standardization program.

--The Veterans Administration issued instructions for inventorying data in its 19 major information systems.

--HEW drafted a policy and procedural document in October 1970, but it had not been released to the component agencies as of November 15, 1973.

--Labor recently established a new organization responsible for ADP systems and planned to prepare Department-wide instructions that relate primarily to developing agency standards.

As provided by previous instructions contained in Circular A-86, most Federal departments have designated program coordinators for matters pertaining to the circular. These coordinators generally have not been able to convince higher management of the need for data standards, and they appear to be removed from those who need and use the data. Their role is not always formally defined, and operating personnel in some departments are not aware of their existence and/or function.

Several of the program coordinators could not readily give us information on:

--The implementation and use of Federal general standards by subordinate organizations of their respective departments.

--The development and use of program and agency standards within their respective departments.

--Departmental plans for actively pursuing the program and issuing related policy and procedural directives.

NBS efforts

The NBS role with regard to ADP standards was defined by OMB in policy documents first issued in March 1965. In essence, NBS was assigned responsibility for increasing the compatibility of ADP equipment by recommending uniform Federal standards for hardware, techniques, and computer languages. NBS was made program manager for developing standards in these areas, but NBS was not given responsibility for standardizing data elements and codes. Its responsibilities were limited to (1) maintaining registers and reference files of data standards established by other agencies and (2) providing technical advice to activities developing these standards.

NBS developed a series of Federal Information Processing Standards Publications (FIPS PUBs) to announce and describe those ADP standards approved for Government-wide use. In this series NBS has published several hardware, software, and Federal general data standards.

In July 1969 OMB asked NBS to develop a register for program and agency data standards. NBS announced formal guidelines and procedures for registering data codes for program and agency data standards in FIPS PUB 19, issued with an OMB transmittal letter in January 1972. To date, however, NBS has published no program or agency standards due to the lack of success of the registration efforts. (See p. 29.)

RECENT CHANGE IN PROGRAM RESPONSIBILITIES

Executive Order 11717 dated May 9, 1973, transferred to the Secretary of Commerce all OMB functions related to

establishing Government-wide ADP standards, including approving standards on behalf of the President. In effect, this order transfers overall leadership for standardizing data elements and codes to Commerce.

The Director, OMB, rescinded Circular A-86 on August 29, 1973. On November 28, 1973, the Secretary of Commerce issued a policy directive which, in addition to incorporating the general provisions of the rescinded OMB circular, clarified and changed the management and operation of a standards program for data elements and codes. Significant clarifications and policy changes under the new directive include

--a change in the title from "codes" to "representations,"

--transfer of functions as prescribed in Executive Order 11717,

--a clearer glossary of terms,

--two additional types of standards (de facto practices and unit standards),

--priorities for ranking types of standards, and

--responsibilities of the Department of Commerce and other departments and agencies.

CHAPTER 3

ACCOMPLISHMENT OF THE FEDERAL DATA STANDARDS PROGRAM

Federal program and agency standardization efforts have made some progress. Certain agencies acted independently because the costs of paperwork processing and information translation and conversion were becoming prohibitive.

MANPOWER ADMINISTRATION, DEPARTMENT OF LABOR

The Manpower Administration is responsible for collecting information from the States on employment security, jobs, placement, and related matters. Manpower found that the data supplied by the States on its 118 different forms was not uniformly defined and coded. For example, the data element "ethnic group" was coded 18 different ways by 24 States.

The following table shows the differences in codes used by only five selected States to describe "American Indian." The States used different data element names (race, ethnic group, etc.), data item descriptors (American Indian, Indian, Red Race), and codes (3, 6, I, C).

| | | Codes used for items | | |
| | | American Indian | Indian | Red Race |
States	Element name			
California	Race	6	-	-
Kansas	Ethnic	3	-	-
Oklahoma	Ethnic group	-	I	-
Utah	Race	-	-	3
Washington	Ethnic group	C	-	-

According to Manpower representatives, significant resources were spent to analyze the reports and convert the data for input into Manpower's computer-based information systems.

In 1970 Manpower established an agency standardization project to encourage the reporting of data in machine-readable form and to eliminate costly translation and conversion operations. With the cooperation of the States, 59 data standards were established, including a standard for ethnic group.

The standards established by Manpower have been adopted by many States. Officials advised us that when the standards are used little or no conversion is needed to process reports currently being received in machine-readable form. Efforts are underway to expand the number of standards for data being reported.

Direct machine-readable exchange has streamlined Manpower's data collection and reporting operations by eliminating slow and repetitious manual steps that are costly and conducive to error. The following flowchart depicts reporting before standards were used and Manpower's improved reporting capability using data standards. The five operations eliminated are shaded.

Although Manpower's independent standardization efforts are a step in the right direction, many problems still exist. Other Federal agencies have established different codes for the same data elements covered by Manpower's codes. The following example shows these differences for the American Indian.

Federal Agency	Element name	Codes for American Indian
CSC	Minority group	3
DOD	Ethnic group designation	2[a]

[a] Includes persons of Aleut or Eskimo descent.

As a result of these differences in the Federal standardization program, machine-readable exchange between DOD, CSC, and Manpower cannot be made without extensive translation and conversion.

MANPOWER ADMINISTRATION INFORMATION COLLECTION AND REPORTING SYSTEM
BEFORE AND AFTER USE OF DATA STANDARDS

FEDERAL LEVEL

WITHOUT DATA STANDARDS

WITH DATA AND OTHER ADP STANDARDS

STATE LEVEL

20
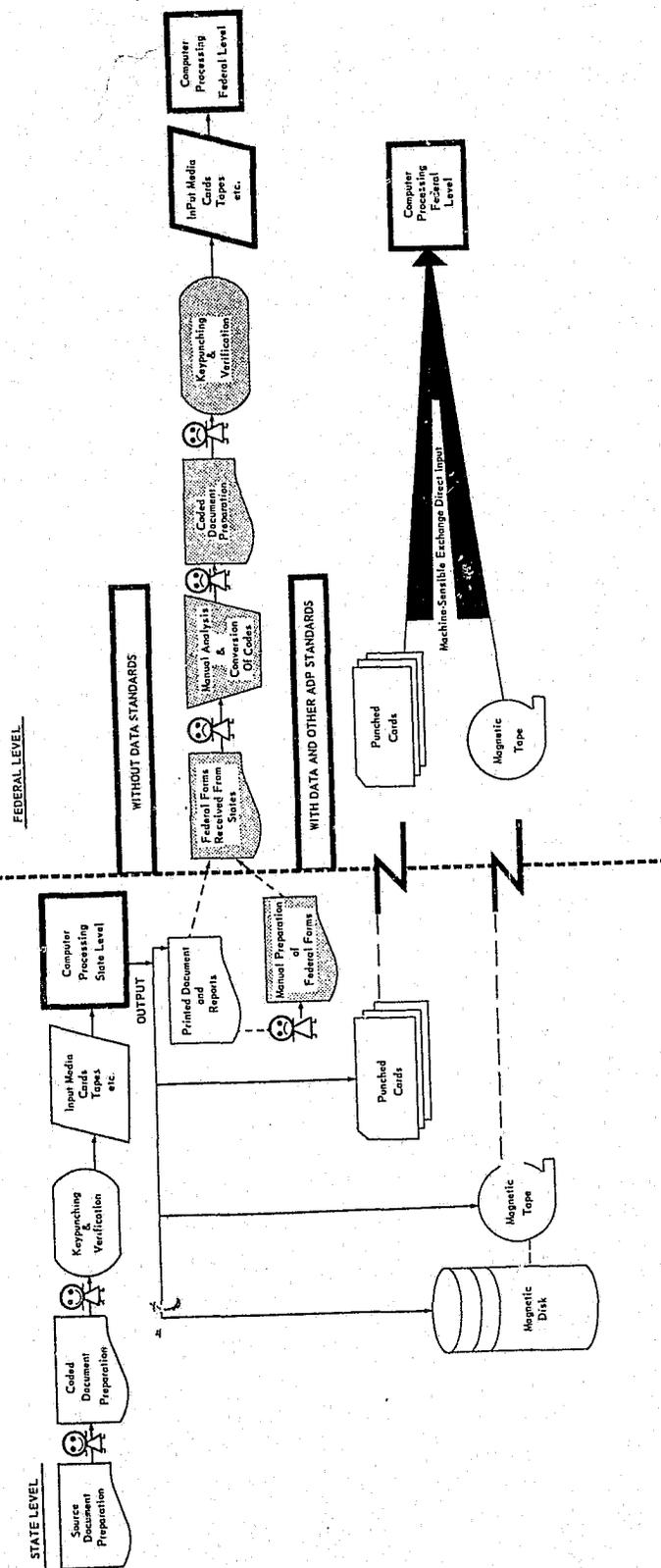
LOGISTICS DATA ELEMENT STANDARDIZATION AND MANAGEMENT OFFICE--DOD

This Office, having been assigned responsibility for standardizing DOD logistics data, has identified and inventoried more than 65,000 data fields (containing one or more data elements) used in more than 39 military logistics information systems. As noted during our review, the Office had analyzed and grouped the data contained in 51,600 of the data fields.

The Office reviewed the data in approximately 13,000 of the fields and established standard names and definitions for homogenous data items. As a consequence, the number was reduced to about 1,600--an 8 to 1 reduction. About 1,000 of these have been adopted for DOD use; the remaining 600 are being reviewed for potential adoption. The Office plans to expand the scope of its review.

The Office advised us that many common data elements with different codes are contained in the DOD logistics systems. For instance, 14 different codes are used in the military to identify items that can be repaired and 16 codes are used to designate the items that cannot. In some cases, the code used by one service to designate a reparable item is the same as the code used by another service to designate a nonreparable item, as illustrated below.

	Codes used by		
	Navy	Air Force	Marine Corps
Reparable items	-	P, U, C	Z
Nonreparable items	P, U, Z	-	C

Each service uses coding to denote several different characteristics in addition to describing whether an item can be repaired. For instance, the Navy's codes identify the

--management attention required to maintain adequate supply,

--frequency of issue, and

--unit price range for the item.

21

The Marine Corps codes are intended to disclose the

--management attention required to maintain adequate supply,

--location where the item can be repaired, and

--unit price range for the item.

OFFICE OF FACILITATION
DEPARTMENT OF TRANSPORTATION

A joint Transportation and industry[1] study showed that participants to a freight shipment had repeatedly to describe the commodities being shipped for the different documents involved in the transaction. Shipping transactions may require the preparation of up to 135 different forms, depending on the type of shipment--export, import, domestic, and so on. Moreover, more than 20 classes of participants, as shown below, may be involved in a shipping transaction.

Possible Participation to Shipping Transactions

Federal:
    Department of Agriculture
    Atomic Energy Commission
    Bureau of the Census
    Bureau of Customs
    Export Control Office
    State Department

Federal and others:
    Buyer
    Consignee
    Supplier

Non-Federal:
    Consul
    Domestic carriers
    Export bank
    Exporter
    Forwarders
    Import bank
    Import broker
    Importer
    Insurance broker
    Insurers
    International carriers
    Local cartage
    Warehouse packers

---

[1] Represented by The National Committee on International Trade Documentation, a nonprofit organization for simplifying and improving international trade documentation and procedures. Membership represents a wide variety of commercial organizations concerned with paperwork costs in international trade.

Transportation and industry[1] are cooperating to develop standard commodity codes for use in freight movement transactions; their efforts have resulted in standards for paper pulp products. Standards for many other commodities are planned. Transportation estimates that savings of $1.2 billion a year will result from the adoption and use of standard commodity codes.

The ultimate goal of this joint effort is to provide one document, containing standard data, to facilitate the automated flow of the necessary information between participants to freight movement transactions, including the Government.

Industry and the Government are expected to realize additional savings through reductions in the amount of paperwork and manual processing required. Government and industry estimate that attaining the project goal of wholly automating the transmission of freight documentation would eliminate 400 million documents and 4 billion copies a year. Annual savings of up to $3 billion are expected.

DATA STANDARDS REGISTERED WITH NBS

The registration program initiated by OMB in January 1972 indicates that other standardization efforts now exist.

--CSC registered 45 personnel standard data elements and codes for use in personnel programs which have been under development for several years. CSC is acknowledged as the program leader for developing these standards because of its responsibilities relating to Federal personnel resources.

--The Library of Congress registered two standard data elements and codes that it developed jointly with the National Library of Medicine and the National Agriculture Library.

---

[1] Represented by the Transportation Data Coordinating Committee, a nonprofit organization for improving coordination of transportation data and information systems through standardization. Membership includes commercial concerns interested in reducing the cost of transportation paperwork.

NBS informed us that, after our survey, registration of data standards had been sporadic. Some data standards, however, have been submitted by the Federal Trade Commission (de facto agency standards), the National Communications System (program standards), and the Atomic Energy Commission (agency standards).

PROBLEMS BEING ENCOUNTERED IN THE

FEDERAL DATA STANDARDIZATION PROGRAM

According to a Government-wide survey by NBS in 1970, most Federal agencies believe that a high priority should be given to standardizing data elements and codes. Despite the indication that data standards are needed as soon as possible, numerous problems related to developing, approving, implementing, and using them have limited the program's progress.

Several problems related to Circular A-86 and its implementation have resulted in limited and independent actions by Federal departments and agencies.

DIFFICULTY IN ASSESSING POTENTIAL BENEFITS

Circular A-86 instructed agencies to consider the technical, operating, economic, or other benefits involved before adopting particular standards. For the most part, no formal or detailed studies have been made to identify and measure the benefits anticipated from specific data standards projects. Interviews with several Government officials involved in information systems indicate that management attention and direction and the necessary resources have not been devoted to the effort because immediate economic benefits cannot be demonstrated.

LACK OF AGREEMENT AMONG AGENCIES
ABOUT APPROACH

Because Circular A-86 terminology was not defined adequately, many points of view on the meanings of terms have developed, precluding a uniform Government-wide approach to the program. For example, such terms as "standard," "agency," and "program," which have various meanings in Government, depending on their application, were not defined. Also, OMB, DOD, and the task force that prepared the draft handbook have discussed the meanings of the various categories of data used in the circular. Differences are still unresolved on the meanings of such a basic term as "data elements" and of such new terms as "representations" and "data use identifiers."

Federal departments disagree on how data standards are to be used in computer-based information systems. At least three views influence the direction of the departmental efforts.

Proponents of one view believe existing information systems represent a significant investment that must be protected. They feel that changes to accommodate any newly established standard data elements and codes would be unnecessarily costly and that it would be less costly to convert and translate data at the time data is exchanged.

Advocates of a second view would use standards only for the data exchanged between systems. But, unlike advocates of the first view, they accept the idea of adopting the standards for internal systems applications. They do not believe the standards should be developed for common data not being exchanged.

Others believe that data common to more than one information system should be standardized even if no data is being exchanged between existing systems. The belief is based on the philosophy that potential future exchanges would be fostered and facilitated and systems integration would be made easier if common data were standardized.

OMB was aware of these different viewpoints and did not take a position on which one should be adopted Government-wide. OMB considered the use of standards voluntary. Departments or agencies may exchange data using any data elements and codes to which they agree.

DIFFICULTY IN IMPLEMENTING STANDARDS
IN SYSTEMS ALREADY DEVELOPED

Most Federal activities have designed, developed, and implemented computer-based information systems with almost exclusive emphasis on meeting specific needs or functions. In this environment, agencies are hesitant and often reluctant to incorporate standards in operating systems because of their preferences for existing codes, the cost of changing existing files, and the need to maintain historical data continuity.

However, changing management requirements, technological advances, and systems growth result in the need to redesign or design information systems periodically, sometimes as often as every 3 to 5 years. Thus, data standards, if incorporated at the point of system design or redesign, would probably cost less than if they were incorporated piecemeal, as approved. Adoption of these standards for internal use at these points would help to insure, in the long run, the ability of departments to achieve the benefits of exchanging pertinent information by automated means as needed.

DIFFICULTY IN ESTABLISHING
FEDERAL GENERAL STANDARDS

Starting in April 1965 OMB, acting under Circular A-71, formed numerous task forces to standardize data elements considered common to almost all Federal information systems. From 1965 to 1970 these groups attempted, part time, to standardize 14 data categories, as follows:

Summary of Federal General Standards Efforts
as of April 1973

Date category	Number of task forces	Date first task force established	FIPS PUB approved	Deferred to ANSI (note a)	In process
Countries	Unknown	Apr. 1965	June 1970		
U.S. States	Unknown	Apr. 1965	b Nov. 1968		
Nonpolitical places	2	Apr. 1965		x	
Federal activities	3	Aug. 1965			x
Calendar date	1	Oct. 1965	Nov. 1968		
Individuals	Unknown	Oct. 1965		x	
Businesses	Unknown	Oct. 1965		x	
Time	Unknown	Authorized in Oct. 1965		x	
Counties and county equivalents	Unknown	After 1965	b Nov. 1968		
Metropolitan statistical areas	Unknown	After 1965	b June 1969		
U.S. congressional districts	Unknown	After 1965	Nov. 1969		
Organizations	Unknown	After 1965		x	
Point locations	1	After 1965		x	
Subdivisions of countries	Unknown	After 1965		x	

[a]The American National Standards Institute, Inc., is a nonprofit national clearinghouse and coordinating agency for voluntary standards in the United States. Its membership consists of trade associations, professional societies, commercial companies, and the Federal Government. ANSI is working to develop standards in these data categories.

[b]This publication has been revised or augmented since this date.

As indicated by the summary, only six data standards have been approved and published. Although more than 4 years have elapsed since their promulgation, many component agencies do not use all the approved standards.

Successful projects took 3 or more years from initiation through publication. Some deferred projects, those which have not produced standards, have been in process longer.

According to agency officials, the following problems were encountered in developing and getting approval for general standards.

--Oversimplification of the task.

--Selection of elements of data considered to be in general use without adequate research, preparation, and study.

--No uniform and comprehensive guidelines to identify the objectives and goals of the standardization efforts and to provide the task forces with the methodology necessary to achieve them.

--No periodic coordination between the task forces.

--Inadequate planning for completing projects promptly.

--Little or no monitoring of the task forces' progress.

--Selection of some task force members who were not aware of information uses or standards development concepts.

--Turnover of key personnel.

--Failure to collect sufficient factual information and evidence to provide a sound basis for, and to support, task force determinations.

--Tendency of departments and their representatives to resist change.

--Lack of knowledge by the Federal departments of the content and use of data in their systems.

DIFFICULTY IN COORDINATING PROGRAM AND AGENCY STANDARDS EFFORTS

Generally, Federal departments did not keep OMB informed of their efforts to develop and use program and agency standards.

In 1970 OMB advised Federal program coordinators of a decision to emphasize the development of program and agency standards. In January 1972 OMB requested departments and agencies to register standard program and agency codes with NBS.

OMB and NBS officials saw registration as an attempt to:

--Identify de facto program standards; that is, data representations currently in wide use that have not been subject to official or formal standardization.

--Provide the basis for broadening the use of existing agency standards.

--Establish an aid to analysts involved in designing new systems.

However, of 62 agencies, 44 did not reply to NBS within the allotted time. Of the 18 replies received, most indicated that there were no data standards to register. Others promised to register standards sometime in the future. Only two departments had registered standard codes at the time of our review.

OMB designated three Federal activities to develop program standards. Except for these, the Government has not clearly delineated who will develop data standards in given program areas.

We have observed numerous program areas where data standards are needed. However, registration results indicate that agencies have taken little action to initiate and develop program standards.

Problems in developing program and agency standards include:

--Absence of external and internal guidance, guidelines, and procedures to provide a uniform approach for (1) identifying the potential for developing standards, (2) developing them, and (3) using them.

--Inability of the program coordinators to obtain the necessary resource commitment.

--Lack of a procedure to monitor the standardization activities of the component agencies and to obtain feedback on development.

PROPOSED CORRECTIVE ACTIONS

In 1970 OMB asked NBS to suggest ways to achieve greater efficiency and economy in Government ADP operations as envisioned in the Brooks bill. With regard to ADP standards, NBS was asked to determine the Government's most pressing needs and how program objectives could be achieved.

NBS, concluding that data standards are an essential element of ADP standardization, asked for overall program responsibility and recommended:

--Appropriate departments to develop and maintain specific standards.

--Preparing guidelines and criteria to assist the departments in developing standards.

--Assessing the program's effectiveness by (1) determining areas that would benefit most from standards, (2) establishing priorities for data standards projects, and (3) monitoring implementation of approved standards.

--Developing training and indoctrination to assist departments in their assigned responsibilities.

CHAPTER 5

POTENTIAL BENEFITS OF DATA STANDARDIZATION

It is impracticable for us to estimate all potential cost savings which would result from standardizing data elements and codes. However, we do know that developing and using such standards would improve Government operations by facilitating exchanges of information in machine-readable form.

Potential benefits from exchanging data in machine-readable form include:

--Reduction in the cost of collecting, processing, and disseminating information by Government agencies.

--Reduction in the amount and cost of paperwork needed to support Government operations.

--Reduction in the effort needed for assigning data elements and codes in Federal systems being designed.

CHAPTER 6

CONCLUSIONS, RECOMMENDATIONS, AND AGENCY ACTIONS

CONCLUSIONS

The increasing need in Government for information, coupled with expanding Federal programs and operations, has highlighted the need for an effective mechanism for exchanging information among data systems.

Automated techniques should be used more to exchange data and information collected by agencies. Data collected and converted to machine-readable form by one agency should be made available to others having a valid need for it. Standard data elements and codes are needed so that data can be exchanged among systems without going through costly re-collection and conversion processes.

We do not foresee a significant amount of automated exchange taking place until the Government standardizes data elements and codes and incorporates them into its information systems as they are designed or redesigned.

Policy determinations are needed on:

--A uniform approach and coordination between departments.

--An across-the-board incorporation of approved Government-wide standards at the most economical time.

--The right of a program leader to initiate standards work with other Federal agencies.

The transfer of responsibility to the Secretary of Commerce for establishing and approving Government-wide ADP standards should be a step in the right direction and could accelerate and improve the data standards program.

The planning, development, and use of data standards must be cooperative and the technical content must represent a consensus of the agencies involved. Consensus imposes a responsibility on the program leadership to form the framework within which standards can be initiated, developed, approved, and used. Because many of the problems hampering present Federal data standards efforts are interdepartmental, we believe it would be more appropriate to coordinate and control the standardization program at a level above the interest of any single department.

The agency exercising program leadership must have appropriate authority, responsibility, and organization to effectively obtain interagency coordination and necessary direction and resources commitments. The program leader must be able to definitively resolve problems that arise and make binding policy decisions regarding them.

We believe that the timely development, adoption, promulgation, and use of data standards in Federal agencies' information systems would provide opportunities for increased savings.

RECOMMENDATIONS

We recommend that, to accelerate timely development and use of standard data elements and codes, the Secretary of Commerce:

--Determine where standards would be most beneficial and establish standardization priorities.

--Issue policy, delineating accepted theory and terminology, and provide for preparation of guidelines, methodology, and criteria to be followed by agencies in their standardization efforts.

--Assign to specific agencies responsibilities for developing standard data elements and codes in specified areas.

--Monitor implementation of data standards to insure their uniform adoption and use.

AGENCY ACTIONS

The Secretary of Commerce advised us in February 1974 that the Department considers standards for data elements and representations to be one of the highest priorities within the Federal Processing Standards Programs. The Secretary pointed out that the recently approved Federal regulation establishing a Government-wide program for data element standardization embodied the recommendations of our report.

--Heads of Federal departments and agencies have been
asked to propose future subjects for standardization.
New areas for standardization will be identified and
priorities will be established on the basis of such
proposals.

--NBS is preparing a procedural manual that contains
accepted theory, terminology and guidelines. The
manual is expected to be completed in fiscal year 1975.

--NBS is considering agreements for specific agencies to
develop and maintain selected exigent standards. Such
agreements are to be effectuated after agency responses
to Commerce's new regulations are received.

--NBS is establishing methods for assessing the impact,
benefits, and problems related to the implementation
of approved standards.

The Secretary pointed out that the success of the stan-
dards program is dependent on the support and cooperation
provided by other Federal departments and agencies. GAO be-
lieves that Commerce needs agency help and support. For this
reason, GAO is sending a copy of the report to each Federal
agency head along with the following suggestions which should
help the Department of Commerce in its data standardization
program. Each agency should

--identify its needs in those areas where data standards
would be most beneficial,

--accept responsibilities for developing data standards
in areas which they have an interest in, and

--support and cooperate with the the Department of
Commerce in the Government-wide program for data
element standardization.

SCOPE OF REVIEW

Our work involved:

--Reviewing the development and use of data standards.

--Identifying the accomplishments of the program and
related problems.

--Analyzing the method of Federal data interchange.

--Making numerous inquiries and interviews with a broad
range of Federal and non-Federal personnel involved
in the data standards and interchange processes.

--Reviewing legislation and supporting circulars, in-
structions, and directives on data standardization.

We visited:

--OMB and NBS because of their responsibilities related
to ADP standards as legislated in the Brooks Bill.

--Some Federal establishments because of their standard-
ization accomplishments.

--Other Federal establishments to obtain information on
what had been done to implement the provisions of the
program and to gain an insight into the problems
encountered.

THE SECRETARY OF COMMERCE
Washington, D.C. 20230

November 28, 1973

MEMORANDUM FOR Heads of Departments and Agencies

SUBJECT: Standardization of Data Elements and Representations

I am forwarding to you the text of a new regulation being issued by the Department of Commerce which will appear in the Federal Register. This regulation provides policy and identifies executive branch responsibilities for a government-wide program for the standardization of data elements and representations used in Federal automated data systems. This regulation supersedes and replaces Office of Management and Budget Circular No. A-86 which was rescinded on August 29, 1973, following the transfer of OMB functions for ADP standards to the Department of Commerce under the provisions of Executive Order 11717 of May 9, 1973.

The principal objective of standardizing data elements and representations is to make maximum utilization of the data resources of the Federal Government and to avoid unnecessary duplications and incompatibilities in the collection, processing and dissemination of data. The program is intended to benefit each department and independent agency through improved utilization of its automatic data processing systems and to benefit the Federal Government and public at large through the effective interchange and sharing of data by government activities and by state and local governments, industry and the public.

In carrying out the objectives of this program, it is necessary that appropriate implementing directives be issued by each department and agency. Accordingly, it is requested that you submit to the Director of the Institute for Computer Sciences and Technology, National Bureau of Standards within 120 days two copies of the internal directives or issuances that will implement this new regulation within your department or agency.

In providing the leadership for this standardization program, I must rely upon the strength, cooperation and support of each department and independent agency. Any suggestions you may have for improving the attached regulation or the standardization program including proposed subjects for future standardization will be welcome.
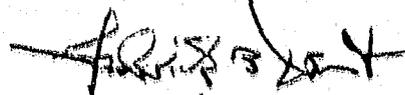
Secretary of Commerce

Enclosure

Title 15 - Commerce and Foreign Trade

Subtitle A - Office of the Secretary of Commerce

PART 6 - STANDARDIZATION OF DATA ELEMENTS AND REPRESENTATIONS

Pursuant to the authority delegated to the Secretary of Commerce by Executive Order 11717 (38 FR 12315, dated May 11, 1973), Subtitle A of Title 15 of the Code of Federal Regulations is hereby amended to add a new Part 6 which implements the provisions of Section 111 (f) (2) of the Federal Property and Administrative Services Act of 1949, as amended (79 Stat. 1127). This new Part 6 supersedes and replaces in its entirety the provisions of Office of Management and Budget Circular A-86 entitled, "Standardization of data elements and codes in data systems", dated September 30, 1967 which was rescinded by the Director of the Office of Management and Budget on August 29, 1973 through a letter to the Heads of Executive Departments and Establishments, same subject.

Part 6 provides policy and identifies responsibilities of executive branch departments and independent agencies for a government-wide program for the standardization of data elements and representations used in Federal automated data systems. The principal features of Part 6 are described in the attached explanatory statement.

Part 6 shall become effective upon publication in the Federal Register.

Issued: November 28, 1973
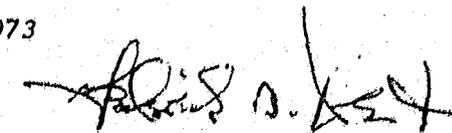
Secretary of Commerce

EXPLANATORY STATEMENT

Public Law 89-306 (79 Stat. 1127) of October 30, 1965, amended Title I of the Federal Property and Administrative Services Act of 1949 (63 Stat. 377) and made provision for the economic and efficient purchase, lease, maintenance, operation and utilization of automatic data processing equipment by Federal departments and agencies.

Pursuant to this statute, the Secretary of Commerce is authorized (1) to provide agencies and the Administrator of General Services with scientific and technological advisory services relating to automatic data processing and related systems, and (2) to make appropriate recommendations to the President relating to the establishment of uniform Federal automatic data processing standards. The Secretary of Commerce is also authorized to undertake necessary research in the sciences and technologies of automatic data processing computer and related systems. The authority conferred upon the Secretary of Commerce is exercised subject to the direction of the President and to fiscal and policy control exercised by the Office of Management and Budget.

By letter to the Secretary of Commerce from the Acting Director of the Office of Management and Budget, dated December 15, 1966, policy guidance was provided concerning the implementation of Public Law 89-306 (79 Stat. 1127). This policy guidance provided

guidelines on the development of voluntary commercial standards and recommendations for uniform Federal standards. The standardization of data elements and codes was excluded in this policy guidance letter.

The Office of Management and Budget on September 30, 1967, issued Circular A-86 entitled "Standardization of data elements and codes in data systems." This Circular (A-86) identified responsibilities and provided policies and guidelines for the management of activities in the executive branch regarding the development and application of standard data elements and their related codes in data systems. The Office of Management and Budget retained the function of providing leadership of the standards program within the executive branch on this subject.

On April 30, 1971, in a memorandum to the Director, Office of Management and Budget, the President authorized and empowered the Director to act finally on behalf of the President, upon recommendations concerning the establishment of automatic data processing standards for use by Federal agencies. Executive Order 11717 of May 9, 1973, transferred to the Secretary of Commerce all functions being performed on that date by the Office of Management and Budget relating to the establishment of government-wide automatic data processing standards, including the function of approving standards on behalf of the President. In this Executive Order (11717) the Director of the Office of Management and Budget retained oversight responsibilities for the transferred functions.

On August 29, 1973, in a letter to the Heads of Executive Departments and Agencies, the Director of the Office of Management and Budget rescinded OMB Circular No. A-86 and stated that revised policies and procedures in regard to the standardization of data elements and codes would be issued by the Department of Commerce. This amendment to Title 15 of the Code of Federal Regulations to add Part 6 is the embodiment of this direction provided by the Director of the Office of Management and Budget.

The new Part 6, in addition to incorporating the general provisions of the rescinded Office of Management and Budget Circular No. A-86, provides clarifications and other changes deemed essential and appropriate for the effective management and operation of a standards program for data elements and codes. The principal changes and additions to Office of Management and Budget Circular No. A-86 provided in the new Part 6 are identified and described below:

Title - Standardization of Data Elements and Representations

The designation of the Part "Standardization of Data Elements and Representations" is a change to the title of Office of Management and Budget Circular No. A-86. The use of the term "representations" is a broader term than "codes" in that it includes codes, abbreviations, names and numeric values. It also avoids confusion with other uses of the term "codes" as in character codes, self-checking codes, redundancy codes and programming codes.

§6.1 Purpose

This section indicates the transfer of functions prescribed in Executive Order 11717 and defines the relationship of the provisions of the Part to Office of Management and Budget Circular No. A-46.

§6.4 Glossary

This section provides clarification of terms and adds new terms needed to describe data standards. The terms used are those defined in Federal Information Processing Standard 20, "Guidelines for Describing Information Interchange Formats" which has been additionally approved by the American National Standards Institute as a voluntary national standard (X10.1-1973). These terms are provided in an appendix as noted in §6.4.

§6.5 Types of Standards

This section provides for two additional types of standards (i.e., De facto practices and Unit standards).

§6.6 Policies

This section provides for a relative ranking of the types of standards; the need to assure that applicable statutes, regulations and Executive Orders are complied with to protect sensitive and classified data files; and the requirement to cooperate with affected activities outside the Federal Government in the development and implementation of data standards.

§6.7 Responsibilities

This section lists the responsibilities of the Department of Commerce and Executive Branch departments and agencies. In addition to the responsibilities for maintaining registers and providing advisory services as in Office of Management and Budget Circular No. A-86, the National Bureau of Standards is assigned the major operational functions for the standardization of data elements and representations. The Secretary of Commerce approves Federal general and Federal program standards.

The responsibilities of Federal departments and independent agencies are for the most part a continuation of responsibilities reflected in the rescinded Office of Management and Budget Circular No. A-86. The need to assess the impact, benefits and problems related to the implementation of approved standards is a new responsibility.

§6.8 Exceptions, Deferments and Revisions of Federal Standards

This section is a change from Office of Management and Budget Circular No. A-86 in that all such requests are coordinated in advance with the National Bureau of Standards rather than with various Government agencies.

§6.9 Effect On Previously Issued Standards

This section continues the standards previously approved under the provisions of Office of Management and Budget Circular No. A-86.

Subtitle A of Title 15 of the Code of Federal Regulations is amended by adding a new Part 6, reading as follows:

PART 6 - STANDARDIZATION OF DATA ELEMENTS AND REPRESENTATIONS

Sec.

6.1 Purpose

6.2 Background

6.3 Objectives

6.4 Glossary

6.5 Types of Standards

6.6 Policies

6.7 Responsibilities

6.8 Exceptions, Deferments and Revisions of Federal Standards

6.9 Effect On Previously Issued Standards

Appendix A - Glossary

AUTHORITY: The provisions of this Part 6 issued under 79 Stat. 1127; Executive Order 11717, dated May 9, 1973 (38 FR 12315, dated May 11, 1973).

§6.1 Purpose

The purpose of this Part is to implement the provisions of Section 111 (f) (2) of the Federal Property and Administrative Services Act of 1949, as amended (79 Stat. 1127) and Executive Order 11717 of May 9, 1973 (38 FR 12315, dated May 11, 1973). It supersedes and

replaces in its entirety Office of Management and Budget Circular A-86 entitled, "Standardization of data elements and codes in data systems", dated September 30, 1967. Office of Management and Budget Circular No. A-86 was rescinded by the Director of Office of Management and Budget on August 29, 1973.

This Part identifies responsibilities and provides policies and guidelines for the management of activities in the Executive Branch relating to the development, implementation and maintenance of standards for data elements and representations used in automated Federal data systems. Its provisions complement the standards and recommendations that have been or may be issued under the statistical procedures prescribed by Office of Management and Budget Circular A-46.

§6.2 Background

Recent advances in computer and communications technologies have made possible the wider use of data and programs that are developed or generated to meet mission requirements of Federal departments, agencies, and activities. While the extended use of these data and programs can contribute to reduced costs in Government operations and improved services, the full advantages of these new technical capabilities cannot be realized until standards are developed and implemented which will provide for the uniform identification, definition and representation of data. These standards for data must also be accompanied by supporting standards for representing graphic characters (alphabets, numbers, and other symbols), communications and device controls. In addition, it is essential to have standards that provide for interchangeable media (e.g., tapes, cassettes and disks) covering both physical and logical specifications.

There is an ever increasing need to interchange data and programs with state, local and other governments, and with industry and the public. This adds further emphasis and dimension to the need for responsive standards that will facilitate interchange.

This Part defines a Federal-wide program for standardizing data elements and representations which are used and interchanged in Government data systems. Other approved standards and guidelines issued by the National Bureau of Standards in the Federal Information Processing Standards series of publications address related ADP subjects and areas.

§6.3 Objectives

The principal goal in standardizing data elements and representations is to make maximum utilization of the data resources of the Federal Government and to avoid unnecessary duplications and incompatibilities in the collection, processing, and dissemination of data.

§6.4 Glossary

Appendix A of this Part provides a glossary of terms as used in this Part and in descriptions of data.

§6.5 Types of Standards

For the purposes of this Part, the following types of practices and standards are identified for data elements and representations:

(a) De facto Practices. Those data elements and representations in current use that have not been subjected to official or formal standardization.

(b) Unit Standards. Those data elements and representations that have been approved by an authorized official for use within that unit. (A unit for purposes of this Part is any Federal organization within the executive branch of the Government, which is at a lower organizational level than an executive department or independent agency).

(c) Agency Standards. Those data elements and representations that have been approved by an authorized official for use within an executive department or independent agency.

(d) Federal Program Standards. Those data elements and representations that have been approved by the Secretary of Commerce for use in a particular program or mission where more than one executive branch department or independent agency is involved with their use. For example, those standards that could be approved and prescribed for use are those which include, but are not limited to, Federal-wide personnel, communications and transportation data systems.

(e) Federal General Standards. Those representations that have been approved by the Secretary of Commerce for Federal-wide use by executive departments and independent agencies in all Federal-wide

programs and for use in all Federal data systems. For example, this includes such representations as calendar dates, state abbreviations and codes, and codes for standard metropolitan statistical areas.

(f) American National Standards. Those data elements and representations that have been approved for voluntary national use by the American National Standards Institute.

(g) International Standards. Those data elements and representations that have been approved by the International Organization for Standardization (ISO), for voluntary use by member nations and international organizations.

§6.6 Policies

The following policies apply to the development, implementation, and maintenance of data element and representation standards:

(a) Data Elements and representations that are prescribed for interchange among more than one executive department or agency or with the private sector including industry, state, local, or other Governments, or with the public at large will be considered for standardization as either Federal general or Federal program standards.

(b) Federal general standards are the highest level standards followed by Federal program standards, agency standards and unit standards in that order. This order establishes a precedence for standards use. For example, a Federal general standard will be used and will supplant a Federal program, agency or unit standard. Likewise a Federal program standard takes precedence over an agency or unit standard.

(c) Approved standards will be implemented by all Federal agencies in all circumstances where technical, operating and economic benefits can be expected to result. These standards will be considered on the basis of their long term benefits and advantages to the Government at large. Local inconveniences or short-term conversion costs need to be recognized, but such factors will not be considered overriding deterrents to the development, implementation, and maintenance of standards that are capable of reducing overall government operating costs or providing improved Government services.

(d) Existing standards will be considered for adoption as Federal general or program standards when these are determined to meet Federal requirements or can readily be adapted to do so.

(e) Approved standards and revisions thereto will be implemented on a time phased basis in order to minimize disruption and conversion costs. Conversion costs will be identified and considered in the submissions of annual budget estimates.

(f) Although data element and representation standards are developed and implemented to provide for the effective interchange and processing of data, Federal departments and agencies must comply with applicable statutes, regulations and executive orders to assure that sensitive or classified data are adequately protected and that only authorized disclosure or release of such data is allowed.

(g) In the formulation of standards for data elements and representations which will have implementation impact on state and local governments, industry or other segments of the private sector,

arrangements will be made to establish necessary liaisons and coordinations with these interests to consider their needs and potential problems in responding to Federally imposed reporting requirements.

§6.7 Responsibilities

Responsibilities for the standardization of data elements and representations are outlined below:

(a) Department of Commerce. The Department of Commerce will provide leadership of an executive branch program for standardizing data elements and representations. Within the Department the following specific responsibilities are assigned:

(1) Secretary of Commerce. The Secretary of Commerce, on behalf of the President, approves all Federal Information Processing Standards. For data elements and representations, this approval will include both Federal general and Federal program standards.

(2) National Bureau of Standards. The National Bureau of Standards will:

(i) Arrange with appropriate executive branch departments and independent agencies to assume leadership and undertake responsibilities for the development and maintenance of specific Federal program and Federal general standards.

(ii) Arrange for the publication and promulgation of approved Federal general and Federal program standards. These will be promulgated by the National Bureau of Standards as Federal Information Processing Standards. The responsibility under this subparagraph includes the authority to modify or supersede these standards whether issued under this regulation or prior to the effective date of this regulation.

(iii)   Maintain and promulgate selected registers of data element and representation standards and practices that are under development or are in current use.

(iv)   Provide procedures, guidelines and criteria to assist Federal departments and agencies in the development, implementation, and maintenance of standards.

(v)   Provide technical assistance, as requested and within the limits of available resources to Federal departments and agencies on matters concerning the utilization of automatic data processing and standardization.

(vi)   Arrange for the assessment of the need, impact, benefits and problems related to the implementation of proposed and approved standards.

(vii)   Coordinate requests for exceptions to and deferments on the implementation of approved Federal standards.

(viii)   Arrange for and coordinate appropriate Federal representation and participation on voluntary industry committees.

(ix)   Arrange for appropriate liaison with state, local and other governments on matters of mutual interest or concern relating to Federal development, implementation, and maintenance of standards.

(b)   Departments and Independent Agencies.   Each of these organizations will:

(1)   Implement approved Federal standards that are announced under the provisions of this Part and assist the National Bureau of Standards in the assessment of the need, impact, benefits and problems related to the implementation of approved standards.

(2)   Assume leadership and support of responsibilities for the development of Federal general and Federal program standards as may be mutually arranged by the National Bureau of Standards.

(3)   Establish within their organizations, mechanisms for the development, implementation and maintenance of agency and unit standards where such efforts will contribute to reduced costs or improved services.

(4)   Establish appropriate procedures and mechanisms within their organizations for the dissemination and implementation of approved Federal standards.

(5)   Review and provide information and comments on proposed standards that are being considered for Federal adoption.   This includes the analyses necessary to assess implementation impact and potential savings or improved services.

(6)   Prepare and submit selected registers of data elements and representations within the data systems of the department or agency as may be arranged by the National Bureau of Standards.   These registers will be used as a source reference to avoid duplication in the design of new data elements and representations and to assist in determining possible subjects for future standardization.

(7) Provide participation on committees and task groups that may be formed to develop and maintain Federal general or Federal program standards.

(8) Provide participation, as requested by the National Bureau of Standards, on committees and task groups that may be formed to develop and maintain voluntary industry standards for use nationally and internationally.

(9) Designate an office or official to act as a single point of contact on matters related to this Part.

§6.8 Exceptions, Deferments, and Revisions of Federal Standards

Requests for exceptions, deferments and revisions of standards will be forwarded to the National Bureau of Standards for consideration and/coordination. These requests will provide detailed justification for the exception, deferment or revision deemed necessary. These should be submitted at least forty-five days in advance of any exception or deferral action.

§6.9 Effect On Previously Issued Standards

All standards that were issued under the provisions of Office of Management and Budget Circular No. A-86 prior to the effective date of this regulation remain in effect unless modified or superseded pursuant to the provisions of this regulation.

GLOSSARY OF TERMS

This Glossary includes definitions of terms used in this Part. Additional terms applicable to data standardization are provided for purposes of clarification. The terms and definitions are either from established vocabularies or have been defined for purposes of this Part.

ATTRIBUTE DATA ELEMENT--A data element that is used to qualify or quantify another data element (e.g., "Date of Birth" and "Mailing Address" would be attribute data elements in a personnel file where the primary element(s) is/are used to identify the person).

CHARACTER TYPE--An indication of the type of characters or bytes to represent a value (i.e., alphabetic, numeric, pure alphabetic, pure numeric, binary, packed numeric, etc.).

ALPHABETIC--A representation which is expressed using only letters and punctuation symbols.

ALPHANUMERIC--A representation which is expressed using letters, numbers, and punctuation symbols.

BINARY--A representation of numbers which is expressed using only the numbers 0 and 1, e.g., 5 is expressed as 101.

NUMERIC--A representation which is expressed using only numbers and selected mathematical punctuation symbols.

PACKED NUMERIC--A representation of numeric values that compresses each character representation in such a way that the original value can be recovered, e.g., in an eight bit byte, two numeric characters can be represented by two.four bit units.

PURE ALPHABETIC--A representation which is expressed using only letters.

PURE ALPHANUMERIC--A representation which is expressed using only letters and numbers.

PURE NUMERIC--A representation which is expressed using only numbers.

COMPOSITE DATA ELEMENT (DATA CHAIN)--A data element that has an ordered string of related data items that can be treated as a group or singly, e.g., a data element named "Date of Birth" could have the data items, "Year", "Month", and "Day of Month".

CONTEXT DEPENDENT DEFINITION--A statement of meaning that relies upon a situation, background, or environment for proper interpretation.

DATE CODE--A coded representation used to identify a data item. Usually codes are designed according to established rules and criteria, and only by chance form a phonetic word or phrase.

DATA ELEMENT--A basic unit of identifiable and definable information. A data element occupies the space provided by fields in a record or

blocks on a form. It has an identifying name and value or values for expressing a specific fact. For example, a data element named "Color of Eyes" could have recorded values of "Blue (a name)", "BL (an abbreviation)" or "06 (a code)." Similarly, a data element named "Age of Employee" could have a recorded value of "28 (a numeric value)."

DATA ELEMENT ABBREVIATION--An abbreviated form of the data element name.

DATA ELEMENT DEFINITION--A statement of the meaning of a data element.

DATA ELEMENT NAME--A name used to identify a data element.

DATA ELEMENT SOURCE--An identification of the source or provider of the particular data element, i.e., individual, organization, sensor, computation, etc.

DATA ELEMENT TAG (DATA ELEMENT CODE)--A symbolic tag used to identify a data element.

DATA ITEM--The expression of a particular fact of a data element e.g., "Blue" may be a data item of the data element named "Color of eyes".

DATA ITEM ABBREVIATION--An abbreviated form of the data item name.

DATA ITEM DEFINITION--A statement of the meaning of a data item.

DATA ITEM NAME--A name used to identify a data item.

DEPENDENT CODE--A code that has segments which are dependent upon other segments in order to provide unique identification of the coded item. Usually, codes having classification significance are dependent codes.

FIELD--In a record, a specific area used for representing a particular category of data, e.g., a group of card columns used to express a wage rate.

FIELD LENGTH--A measure of the length (size) of a field, usually expressed in units of characters, words, or bytes.

FIELD LENGTH TYPE--An indication of whether the field of a record is fixed or variable in length.

FIXED LENGTH FIELD--A field whose length does not vary.

VARIABLE LENGTH FIELD--A field whose length varies. Usually, the boundaries of this type of field are identified by field separators.

FIELD SEPARATOR--A character or byte used to identify the boundary between fields.

FILLER CHARACTER--A specific character or bit combination used to fill the remainder of a field after justification.

FORMATTED INFORMATION--An arrangement of information into discrete units and structures in a manner to facilitate its access and processing. Contrasted with narrative information that is arranged according to rules of grammar.

GENERAL DEFINITION--A statement of meaning that can be interpreted without regard to a specific situation, background, or environment.

INFORMATION INTERCHANGE--The transfer of data representing information between or among two or more points (devices, locations, organizations, or persons) of the same or different (dissimilar) information system or systems.

JUSTIFICATION--To adjust the value representation in a field to either the right or left boundary (margin).

LEFT JUSTIFY--Adjustment of a value representation to the left boundary (high order) of a field.

RIGHT JUSTIFY--Adjustment of a value representation to the right boundary (low order) of a field.

NON-SIGNIFICANT CODE--A code that provides for the identification of a particular fact but does not yield any further information, e.g. random numbers used as codes. Contrasted with significant code.

NUMERIC VALUE--The expression of a data item which denotes a measurement, count, or mathematical concept, usually represented by numerals and a limited number of special characters (i.e., plus (+), minus (-), decimal point (.), comma (,), asterisk (*), and slant (/)).

PADDING--A technique used to fill a field, record, or block with dummy data (usually zeros or spaces).

PRIMARY DATA ELEMENT--A data element or elements that is/are the subject of a record. Usually the other elements, called attribute data elements, qualify or quantify the primary data element (e.g., in a personnel field, the element(s) that is/are used to identify the individual are primary; other elements such as "Date of Birth" and "Mailing Address" are attribute data elements).

RADIX POINT--A character, usually a period, that separates the integer part of a number from the fractional part. In decimal (base 10) notation the radix point is called the decimal point.

RECORD--A collection of related elements of data treated as a unit.

RECORD INDEX--An ordered reference list of the contents of a record together with keys or reference notations for identifying and locating the contents.

RECORD LAYOUT--A description of the arrangement and structure of information in a record, including the sequence and size of each identified component.

RECORD LENGTH--A measure of the length (size) of a record, usually expressed in units of characters, words, or bytes.

RECORD LENGTH TYPE--An indication of whether the records of a file are fixed or variable in length.

FIXED LENGTH RECORD--Pertaining to a file in which the records are uniform in length.

VARIABLE LENGTH RECORD--Pertaining to a file in which the records are not uniform in length.

REPRESENTATION--A number, letter or symbol used to express a particular concept or meaning. It may be in the form of a name, abbreviation, code, or numeric value.

ROUNDING (ROUNDOFF)--To delete the least significant digit or digits of a numeral, and to adjust the part retained in accordance with some rule.

SELF-CHECKING CODE--A code that is appended to another code to provide for validity checking. A self-checking code is derived mathematically from the characteristics of the base code.

SIGNIFICANT CODE--A code which in addition to identifying a particular fact also yields further information, e.g., catalog numbers in addition to identifying a particular item also often indicate the classification of the item. Contrasted with non-significant code.

TRUNCATE--To delete characters from a character string, usually from either end of the string.

TYPE OF CODE SIGNIFICANCE--An indication of the type of significance that a particular code yields.

COLLATING SIGNIFICANCE--A code designed in such a way that it facilitates ordering of the coded item.

MNEMONIC SIGNIFICANCE--A code designed in such a way as to facilitate the human recall of the name of the coded items.

CLASSIFICATION SIGNIFICANCE--A code designed in such a way as to facilitate the classifying of the coded items into classes and sub-classes.

VARIABLE NAME DATA ELEMENT--A data element that identifies a set (array) of similar values (data items). By varying certain identifiers in the name the entire set (array) of values can be identified. For example, a set of values that give population by State and Year could be identified by the data element "Population of (State) in (Year)" where State and Year are variable names. The variable names are used to identify particular values in an array (e.g., "Population of

RECORD LAYOUT--A description of the arrangement and structure of information in a record, including the sequence and size of each identified component.

RECORD LENGTH--A measure of the length (size) of a record, usually expressed in units of characters, words, or bytes.

RECORD LENGTH TYPE--An indication of whether the records of a file are fixed or variable in length.

FIXED LENGTH RECORD--Pertaining to a file in which the records are uniform in length.

VARIABLE LENGTH RECORD--Pertaining to a file in which the records are not uniform in length.

REPRESENTATION--A number, letter or symbol used to express a particular concept or meaning. It may be in the form of a name, abbreviation, code, or numeric value.

ROUNDING (ROUNDOFF)--To delete the least significant digit or digits of a numeral, and to adjust the part retained in accordance with some rule.

SELF-CHECKING CODE--A code that is appended to another code to provide for validity checking. A self-checking code is derived mathematically from the characteristics of the base code.

(New Jersey) in 1970" was 7,168,164.) In this example "New Jersey" and "1970" are variable names used to identify a specific value "7,168,164" in an array.

UNITED STATES DEPARTMENT OF COMMERCE
The Assistant Secretary for Administration
Washington, D.C. 20230
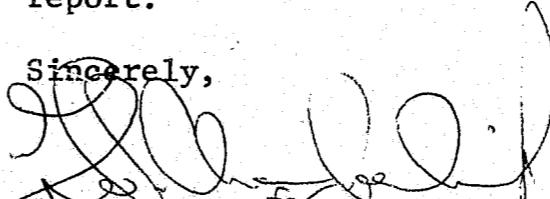
February 15, 1974

Mr. John Landicho
Associate Director
General Government Division
U. S. General Accounting Office
Washington, D. C. 20548

Dear Mr. Landicho:

This is in reply to your letter of December 26, 1973, requesting comments on a draft report entitled "More Emphasis Needed on Government's Efforts to Standardize Data Elements and Codes for Computer Systems".

We have reviewed the comments of the Assistant Secretary for Science and Technology and believe that they are appropriately responsive to the matter discussed in the report.

Sincerely,

Henry B. Turner
Assistant Secretary
  for Administration

Attachment

UNITED STATES DEPARTMENT OF COMMERCE
The Assistant Secretary for Science and Technology
Washington, D.C. 20230

JAN 2 9 1974

Mr. John Landicho
Associate Director
General Government Division
U.S. General Accounting Office
Washington, D. C. 20548

Dear Mr. Landicho:

Reference is made to your letter of December 26, 1973, to
Secretary Dent which forwarded a draft GAO report entitled,
"More Emphasis Needed on Government's Efforts to Standardize
Data Elements and Codes for Computer Systems." We have
reviewed the draft report and find it cogent, timely, factual,
and comprehensive. As you have noted in the report, Secretary
Dent has recently approved a new Federal Regulation that
established a Federal Government-wide program for data element
standardization. This new regulation which embodies the
recommendations that are detailed in the draft GAO report has
been forwarded to the heads of Federal departments and agencies
for Government-wide implementation. Already responses are
being received from agencies recognizing the importance of data
standards and welcoming the Department's leadership in this
requisite aspect of standardization.

The Department considers standards for data elements and
representations to be one of the highest priorities within
the Federal Information Processing Standards Program. I
have requested the Director of the National Bureau of Standards
to initiate his responsibilities as identified in the new
regulation through reprogramming actions. The additional
resources needed to provide the recommended increased support
for data standards will be identified and considered in the
Department's FY 1976 budget submission.

It is further noted that the success of the standards program
is dependent not only upon the leadership provided by the
Secretary of Commerce and the efforts of the National Bureau
of Standards, but, even more importantly, the support and
cooperation provided by other Federal departments and

independent agencies in the development and implementation
of standards. Accordingly, it is suggested that other
recommendations be added to the draft report that address
the responsibilities of other departments and agencies in
an effective Federal Government-wide data standards program.

We appreciate the opportunity to comment on the results of
your review prior to its final release. Other pertinent
comments bearing on the draft report are enclosed.

Sincerely,

Betsy Ancker-Johnson, Ph.D.

Enclosure

COMMENTS

1. The findings on page 2 state that data transferred between Federal agencies should be in machine-sensible form wherever possible. However, the recommendations in the draft report do not address this vital aspect of improving systems performance and data management.

2. In regard to the recommendations on page 5, the following actions are being taken:

    a. In forwarding the new regulation on data element standardization to the heads of Federal departments and agencies, Secretary Dent asked that future subjects for standardization be proposed. Based upon the inputs received and in cooperation with concerned agencies, new areas for standardization will be identified and priorities will be established.

    b. A procedural manual that contains accepted theory, terminology and guidelines is in preparation at the National Bureau of Standards. This manual is expected to be completed in early FY75.

    c. Agreements with specific agencies to develop and maintain selected exigent standards are being considered by the National Bureau of Standards and will be effectuated after agency responses are received implementing the new regulation on data standards. New areas already identified and under consideration include standards for commodity codes, industry codes, organization (Government and non-Government) codes, occupation codes, and curriculum codes.

    d. The National Bureau of Standards is establishing methods for assessing the impact, benefits, and problems related to the implementation of approved standards.

3. The draft report on pages 7 and 12 cites that most data collection is still often a manual operation. The National Bureau of Standards is undertaking new initiatives in the area of automated data collection and related techniques to improve computer utilization through more effective man-machine interfaces. The results of this work

is expected to increase performance and facilitate data interchange. Currently, however, the management of ADP and manual reports control are treated as separate functions in most Federal agencies. The management of data, whether manual or automated, needs to be centralized, if the benefits expected in the draft report are to be realized.

END