&EPA

Agency Catalog of Data Policies and Standards

PREFACE

This catalog summarizes EPA-wide and Federal data policies and standards. Each Agency-wide policy and standard is described with the full text version in the appendix. Federal data policies and standards, knownas FIPS, are briefly described as are the efforts currently underway to develop International Data Standards for environmental data.

The policies and standards in this Catalog are the definitive list of data standards that Agency personnel and contractors must meet when developing information systems. As new policies and standards are developed and promulgated, this catalog will be updated. Waivers are possible only upon request to, and approval by, the Director of OIRM.

This catalog is a tribute to the many EPA staff and contractors who have developed and implemented Agency-wide data standards. Their vision and perseverance have laid the foundation for a successful EPA data standards program.

Your comments on how to improve this catalog would be appreciated. Please send them to:

Program Manager for Data Administration
Information Management and Services Division (PM-211D)
Office of Information Resources Management
US EPA
401 M Street, SW
Washington, DC 20460

Printed on Recycled Paper

i

AGENCY CATALOG OF DATA POLICIES AND STANDARDS

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Background

As EPA's information needs have expanded, the Agency has developed numerous computer systems to support its management and regulatory functions. Many of these systems were designed by individual EPA program offices, without the aid of uniform Agency data standards. This pattern of system development has resulted in many non-uniform coding practices and incompatible data structures, leading to inefficient data sharing between systems belonging to different program areas. Moreover, the growth in end-user computing has introduced a trend toward decentralization of systems support and development throughout EPA. This decentralized environment has contributed dramatically to the increased potential for duplication of data and incompatible data formats among EPA information systems and data repositories. This makes it difficult, expensive, and sometimes impossible for data to be meaningfully shared across EPA systems and programs and with state and local governments, industry, and the public.

To take full advantage of EPA's growing technological and data resources, there needs to be an increased emphasis on improving compatibility of data among EPA systems. A data standards program can help establish an environment in which data is managed effectively and is sharable from one program to another. Standards help to improve quality and efficiency, reduce waste and unnecessary duplication in systems, and enhance the portability of data, software and technical skills across systems. The principal goal of a data standards program is to make maximum use of the tremendous investment EPA makes in its data. The increasing need to share EPA data adds further emphasis to the need for responsive standards that facilitate interchange of EPA information resources.

In response to EPA's data management needs, the Agency has undertaken a Data Standards program. Defining basic data standards for national environmental systems is a major initiative for EPA and is critical to progress in data integration efforts and improvements in managing for environmental results. As part of this program EPA-wide data standards are proposed, studied, refined, approved, and promulgated to increase consistency and portability of data within the Agency and between the Agency and outside parties (such as States, industry, and other Federal agencies).

1.2. Purpose

This catalog is among the first products of the EPA Data Standards Program. The purpose of this catalog is to provide a useful source of up-to-date information on the current EPA-wide and Federal data standards that must be followed in developing, implementing, operating and maintaining all EPA information systems and services. This catalog contains all the EPA Agency-wide standards, and lists the Federal standards and guidelines that are issued by the National Institute of Standards and Technology

(NIST) as Federal Information Processing Standards (FIPS), and international standards that are relevant to EPA mission and functions. It will be updated periodically when EPA issues new Agency-wide data standards or when other improvements to the catalog are identified. For example, there are new data standards currently in development at the Agency; many of these are not listed in the catalog due to their draft, developmental status. The next version of the Catalog may describe a new EPA spatial data coverage documentation standard. It may include a standard for consistently representing and encoding EPA's organizational structure. It may present a definitive, agreed-upon standard for identifying biological taxonomy.

As these, and other, important efforts culminate, the catalog will provide one centralized means for disseminating the standards to the broad Agency IRM community. Without aggressive communication of data standards to the relevant community, widespread compliance with the standards is unlikely. The Catalog is an essential communication tool to promote compliance with standards.

1.3. EPA Data Standards Program

The EPA Data Standards Program was established and formalized by the *EPA Information Resources Management Policy Manual*, Chapter 5, Data Standards. Under this policy, the Office of Information Resources Management (OIRM) is responsible for providing leadership in developing standards; coordinating the evaluation and approval process of all data standards with the Assistant Administrators, Regional Administrators, Office Directors and Senior Information Resources Management Officials; exercising final approval authority for the adoption of new data standards and granting waivers, when appropriate, to current Agency data standards; proposing data elements or representations for use by more than one organizational component of EPA; and supporting other EPA data administration efforts. OIRM's Information Management and Services Division (IMSD) is responsible for the Data Standards Program, as one component of its broader efforts in data administration. EPA organizations are responsible for implementing Agency data standards and establishing an organization-wide data standards work group which reviews and provides information and comments on proposed standards. EPA organizations are also in the best position to propose new data standards for the Agency to use within the environmental community.

If you have a data standard to propose for Agency-wide use, please contact the Data Standards Administrator in OIRM/IMSD. IMSD will serve as the focal point for initiating the process to attain Agency-wide consensus on the proposed standard.

Currently, EPA has formally adopted four Agency-wide data standards. These standards were published as EPA Transmittal Orders to be added to the EPA Directives System. They are contained in Section 4 of this manual. In addition to these standards, it is EPA policy and EPA's responsibility to conform with relevant Federal Information Processing Standards (FIPS) in designing and implementing new data processing

systems and in making major modifications to existing data processing systems. A list of applicable FIPS is contained in Section 5.

1.4. Format of the Catalog

The catalog is organized into five remaining sections:

Section 2. Federal Data-Related Laws, Regulations, Policies and Guidelines. This section describes the Federal laws, regulations, policies and guidelines that mandate the need for an effective information resources management program in all Federal agencies.

Section 3. EPA Data Related Policies. This section contains the EPA IRM policies related to the management of data as an essential resource.

Section 4. EPA Agency-wide Data Standards. This section contains all the current EPA-wide data standards. The many excellent data standards used in individual Agency systems are not listed here -- they are not within the scope of this Catalog.

Section 5. Federal Data Standards. This section lists the Federal standards and guidelines issued by NIST as Federal Information Processing Standards.

Section 6. International Data Policies, Guidelines, and Standards. This section describes the ongoing efforts in development of International Data Standards for environmental data.

Bibliography. The bibliography contains a list of Federal and other publications that relate to the management and administration of data.

Appendix A. Appendix A contains full-text copies of the current EPA Agency-wide data policies and standards.

2. FEDERAL DATA-RELATED LAWS, REGULATIONS, POLICIES AND GUIDELINES

2.1. Overview

This section provides an overview of the Federal data-related laws, regulations, policies and guidelines that form the basis for Federal data management activities. All Federal agencies are required to adhere to these laws and regulations. This is accomplished through establishing Federal Information Resource Management (IRM) programs and implementing agency policies and standards. EPA-specific data policies and standards are described in Sections 3 and 4 of this catalog. Since the remainder of this section provides mostly background information, catalog readers may want to go directly to the next two sections for EPA-specific information.

There have been many laws passed that concern, directly or indirectly, information or data management activities. The Federal Property and Administrative Services Act of 1949 was among the first, providing basic procurement and management authority for telecommunication and ADP resources. This act authorized the General Services Administration to prescribe standard purchase specifications and their use in the procurement process. In response to this authority, GSA established the Federal Standardization Program. The Federal Standardization Program is defined in GSA Handbook FIRMR Bulletin C-3.

Many other laws have been passed which deal with various aspects of information activities. The "Federal Records Act of 1968" (44 U.S.C. 29) assigned to GSA a central management role in government wide records management. The "Procurement of ADP Resources by the Federal Government (Brooks) Act of 1965" (40 U.S.C. 759), governs Federal acquisition and management of ADP equipment. The "Freedom of Information Act of 1966" (5 U.S.C. 552) established the right of citizens to request information from Federal agencies. The "Privacy Act of 1974" (5 U.S.C. 552a) established an individual's right to review records maintained about him/her by Federal agencies (among other provisions). Amendments to these laws, and many other laws, regulations, Executive Orders, and Office of Management and Budget (OMB) Circulars have been promulgated which have an impact on Federal information management.

In 1980, the Paperwork Reduction Act (PRA), P.L. 96-511, became law. The PRA formally established the concept of IRM. The PRA was the first major legislation which recognized that information itself is a valuable resource. Among other measures, it gave agencies the responsibility to carry out information management activities in an efficient, effective, and economical manner and to comply with the information policies, principles, standards, and guidelines prescribed by the Director of OMB. It also mandated that each agency designate a senior official, reporting directly to the head of that agency, to carry out responsibilities under the Act. It is the responsibility of the

senior IRM official for EPA to develop and promulgate EPA IRM policies that implement the Federal laws and policies pertaining to information management.

2.2. Applicability

Federal laws, policies, and guidelines apply to all Agency information management activities. Agencies, in response to Public Laws, regulations and directives, have established IRM organizations to bring together the entire range of information-related activities and provide implementing guidance to Agency organizations in managing their daily information activities. OIRM has been delegated the primary responsibility in EPA to provide leadership in developing, promulgating and enforcing the policies derived from the Federal IRM requirements, such as the EPA Data Standards Program.

2.3. Location

All Federal IRM laws, regulations, policies, and guidelines are available in the EPA Headquarters library or can be obtained through OIRM.

2.4. Federal Data-Related Laws, Regulations, Policies, and Guidelines

The following are the key Federal laws and guidelines relating to management and utilization of IRM resources.

2.4.1. P.L. 96-511 - Paperwork Reduction Act of 1980 (44 U.S.C. 35)

This act establishes a broad mandate for agencies to perform their information activities in an efficient, effective, and economical manner. This Act also requires Federal agencies to designate a Senior Official, who reports directly to the agency head, as the responsible party for information; and defines management functions covered under the Federal Property and Administrative Services Act of 1949 (40 U.S.C. 759). The Paperwork Reduction Act specifically recognized GSA's role in the acquisition and management of automated data processing equipment (ADPE).

2.4.2. P.L. 96-591 - Paperwork Reduction Reauthorization Act of 1986 (44 U.S.C. 35)

This act refines the intent of the original act by defining "information resources management" as the "planning, budgeting, organizing, directing, training, promoting, controlling, and management activities associated with the burden, collection, creation, use and dissemination of information by agencies, and includes the management of information and related resources such as automatic data processing equipment."

2.4.3. P.L. 90-620 - The Federal Records Act of 1968 (44 U.S.C. 29)

This act assigned the General Services Administration (GSA) a central management role regarding government-wide records management. Subsequently, P.L. 98-497 (44 U.S.C. 21) established the National Archives & Records Administration as an independent agency, and transferred to it many of the records management responsibilities formerly at GSA.

2.4.4. P.L. 100-235 - Computer Security Act of 1987 (44 U.S.C.)

The Computer Security Act provides further standards and guidance for assuring the cost-effective security and privacy of sensitive information in Federal computer systems. Computer security, as covered in the Act and Federal guidances, implies more than security for "clearance-restricted" files. It includes all "sensitive data: data which is mission-critical to the agency; data which has right-to-privacy implications; and all data which, if lost or changed, would significantly affect the agency's finances, operations, or plans." A key element of the Act is the requirement for each Federal agency to identify its computer systems which contain sensitive information, and to establish plans for minimizing the loss, misuse, unauthorized access to, or modification of information in those systems. Additionally, each agency is required to train in security matters all personnel involved in the management, operation, and use of Federal computer systems.

2.4.5. Computer Matching and Privacy Protection Act of 1988 and Amendments of 1990

These amendments provide additional protection for individuals whose records are used in a computer matching program. Before any action can be taken against an individual as a result of the information produced by a matching program, the agency must independently verify the information as accurate. Each Agency must establish a Data Integrity Board to review all proposed computer matches.

2.4.6. OMB Circular A-130, Management of Federal Information Resources (December 12, 1985)

This Circular provides the general policy framework for management of Federal information resources. The Circular provides that agencies shall establish multi-year strategic planning processes for acquiring and operating information technology that meet program and mission needs, reflect budget constraints, and form the basis for budget requests.

2.4.7. Federal Information Resources Management Regulations (FIRMR)

The General Services Administration is chartered with preparing, issuing and maintaining the FIRMR. This document is the primary regulation governing Federal agencies' management, acquisition, and use of certain ADP and telecommunications resources. The FIRMR is to be used in conjunction with the general procurement and contracting regulations contained in the Federal Acquisition Regulation (FAR). Agency compliance with the FIRMR is the responsibility of the agency head. The FIRMR contains the provisions of public laws and is the primary policy and procedure guideline for IRM in the Federal Government. As such, all policy and procedures issued by Federal agencies are written for the purpose of implementing or supplementing FIRMR direction. Although all parts of the FIRMR are valuable references for all aspects of IRM, Bulletin C-3, which replaces Part 201-9, describes the policies and procedures regarding the Federal Standardization Program.

3. EPA DATA-RELATED POLICIES

3.1. Overview

This section describes the EPA policies on data management. Data policies are the formal statements that implement the EPA's IRM program directions and Federal laws and mandates. These policies are the management framework for EPA's IRM program, of which data management activities are an important component. The majority of EPA's formal policies relating to the management of IRM resources are contained in EPA's *Information Resources Management Policy Manual*.

3.2. Applicability

In general, EPA's data-related policies apply to all EPA organizations and their employees. They also apply to the facilities and personnel of agents (including contractors and grantees) of EPA who design, develop, operate or maintain Agency information and information systems including automated and manual systems developed for program or administrative purposes.

3.3. Location

The majority of EPA's formal policies relating to the management of IRM resources are contained in *EPA's Information Resources Management Policy Manual*. This manual can be obtained from the Information Management and Services Division within OIRM. (The EPA policies discussed in this section are also contained in Appendix A of this manual.)

3.4. EPA Data-Related Policies

The following four policies address the management of EPA data:

3.4.1. Policy on Electronic Reporting, July 30, 1990.

This policy establishes a uniform Agency approach to electronic reporting by the regulated community. The policy pertains to submission of reports via electronic or magnetic media. This policy is intended to promote the adoption of electronic reporting by EPA programs, and ensure that as they implement this technology they do so in a manner that is both consistent across the Agency and compatible with the electronic reporting practices that are current in the regulated community. This policy applies in every case where EPA programs are requesting or requiring reports from members of the regulated community, including those cases where reports are required or requested of Federal facilities. This policy applies whether the recipients of the reports are EPA programs at headquarters or in regional offices, or State or local government agencies acting under EPA-delegated authority.

3.4.2. Chapter 3, "State/EPA Data Management"

This policy establishes the principles that govern the management and sharing of data between EPA and State environmental agencies and the information systems that handle this data. This policy also defines roles and responsibilities for implementing and ensuring adherence to these principles.

3.4.3. Chapter 5, "Data Standards"

This policy establishes the EPA Data Standards Program. The purpose of this program is to provide consistent definition of data and to facilitate cross-media use of data. This policy sets forth Agency principles on data standards and assigns organizational responsibilities for implementing and administering common data standards.

3.4.4. Chapter 13, "Locational Data"

This policy establishes the principles for collecting and documenting latitude/longitude coordinates for facilities, sites and monitoring and observation points regulated or tracked under Federal environmental programs within the jurisdiction of the Agency. The intent of this policy is to extend environmental analyses and allow data to be integrated based upon location, thereby promoting the enhanced use of EPA's extensive data resources for cross-media environmental analyses and management decisions. This policy underscores EPA's commitment to knowing where environmental results occur and to establishing the data infrastructure necessary to enable data sharing and secondary data use.

4. EPA AGENCY-WIDE DATA STANDARDS

4.1. Overview

Data standards are statements of the specific rules guiding the implementation of IRM data policies and day-to-day IRM data management activities. Depending upon the scope or topic of the standard, standards may be implemented EPA-wide or may be specific to one program office. This section contains the EPA-wide data standards. These data standards focus mostly on required data elements, lists of valid values for selected data elements, and issues relating to transmission of data. At this time, they do not cover some types of data standards such as naming conventions, data set description standards, standards for data dictionaries, etc. Agency-wide standards of these types may be developed in the future. The data standards contained in this section were selected because of their unique cross-cutting nature. Currently, all EPA-wide data standards are issued under the EPA directive system as Agency Orders and undergo the Green Border review process, requiring an absolute consensus among all EPA organizational components for approval.

4.2. Applicability

EPA's data standards apply to all EPA organizations and their employees. They also apply to the facilities and personnel of agents (including contractors and grantees) of EPA who design, develop, operate or maintain Agency information and information systems including automated and manual systems developed for program or administrative purposes. Any requests for waivers from these data standards must be sent to OIRM for review and approval.

4.3. Location

Appendix A contains copies of all of the EPA-wide data standards. Because these data standards are issued as EPA directives, additional copies of the data standards are available through the Office of Administration's Management and Organization Division, which manages the EPA directive system.

4.4. EPA Agency-wide Data Standards

The following four EPA-wide data standards are currently in effect:

4.4.1. Chemical Abstract Service Registry Number Data Standard, EPA Order 2180.1, June 26, 1987

The CAS registry number provides consistent and unambiguous identification of chemicals and facilitates sharing chemical information across program offices. The CAS number is a unique identifier assigned by the Chemical Abstract Service to each distinct chemical substance recorded in the

CAS Chemical Registry System. It is represented as a ten digit code with the first nine characters uniquely identifying the chemical and the tenth character acting as a check digit. The Chemical Abstract Service Registry is available online through the EPA Headquarters Library. Information concerning adding a chemical to the registry may be obtained by calling the American Chemical Society at 1-800-848-6538.

4.4.2. Data Standards for Electronic Transmission of Laboratory Measurement Results, EPA Order 2180.2, December 10, 1987

This standard prescribes the formats and electronic representations of data commonly used in transmitting laboratory results. The order covers EPA laboratory transmission standards regarding: media formats describing diskette and tape specifications and standard record lengths; record formats describing each record layout; definition of production runs; record sequences describing the order of records within a file; file and record integrity; date and time formats; and other transmission information. This standard also presents record layouts and field definitions for many commonly used record types. The standard enables consistent formatting of lab sample data and all associated data relating to instrumentation and sample controls.

4.4.3. Facility Identification Data Standard, EPA Order 2180.3, April 9, 1990

This Order establishes a data standard for unique facility identification codes to be maintained in all EPA data collections containing information on facilities regulated by EPA under authority of Federal environmental legislation. Standardization of the format and content of facility identification codes enhances data integration capabilities and increases the utility of all EPA data on facilities. This coding scheme is not based on an external standard, such as the DUNS ID, but rather is a unique EPA identifier comprised of the relevant FIPS standard 2-digit state, followed by a unique random 10-digit identification number. These codes are assigned and tracked by the Facility Index Data System (FINDS). The facility ID standard element must be maintained as a key data element in all EPA systems containing facility-related information.

4.4.4. Minimum Set of Data Elements for Groundwater, EPA Order No. 7500.1, September 11, 1989

This EPA Order establishes the minimum data elements to be collected and managed by the Agency ground-water community, including States, local governments, the regulated community, EPA, and other Federal agencies. It is EPA policy to collect and maintain consistent ground-water data based on a minimum set of data elements. EPA Program Offices and Regions are required to incorporate this policy into their ground-water data collection

activities through appropriate mechanisms such as regulations, policies, directives, orders, guidance, or procedures, when consistent with statutory and other significant policy, administrative, and technical considerations.

5. FEDERAL DATA STANDARDS

5.1. Overview

The National Institute of Standards and Technology (NIST) develops standards and guidelines, provides technical assistance, and conducts research for computers and related telecommunications systems to achieve more effective use of Federal information technology resources. The Federal Information Processing Standards (FIPS) are the Federal data standards for all data interchange among executive departments and independent agencies, and for Federal data interchange with the non-Federal sector including industry, State, local and other governments, and the public at large.

5.2. Applicability

All Federal agencies must comply with the FIPS standards. Nonconformance with FIPS PUBS standards requires a waiver from OIRM. Any requests for waiver, along with the appropriate justifications, should be forwarded to the Director, OIRM, for review.

5.3. Location

A complete set of FIPS is maintained in the EPA Headquarters Library. Automated copies of selected FIPS can be found on EPA's mainframe computer. It is likely, in fact, that multiple copies of certain FIPS exist as parts of separate application systems on the mainframe. Future versions of this catalog may identify the mainframe location of the definitive copy of these FIPS, so that purchase and maintenance of redundant copies can be eliminated.

The following is a list of selected FIPS -- those that deal most closely with specific data element standards. Additional FIPS are listed in the Bibliography. These additional guidelines provide more information on general Federal *data administration* activities.

5.4. Federal Data Standards (FIPS)

The following is a list of 16 FIPS data standards -- primarily those that set formats or provide lists of valid values for specific data elements.

5.4.1.	FIPS PUB 4-1, Representation for Calendar Date and Ordinal Date for Information Interchange

This standard provides a standard means of representing calendar date and ordinal date to facilitate interchange of data among information systems.

5.4.2.	FIPS PUB 5-2, Codes for the Identification of the States, the District of Columbia and the Outlying Areas of the United States, and Associated Areas

Two-character alphabetic abbreviations and two-digit numeric codes are defined for each of the states and U.S. territories. The abbreviations are identical to the Postal Service codes and therefore are useful for identifying establishment locations and for producing mailing addresses.

5.4.3.	FIPS PUB 6-4, Counties and Equivalent Entities of the United States, its Possessions, and Associated Areas

The FIPS representation of this data element consists of three-digit numeric codes for each county or county equivalent in the U.S. The counties are arranged in alphabetical order and the corresponding codes are unique within each state. For example, the first counties in Alaska and Wyoming are each designated '001'. This sequence, when combined with the FIPS numeric state code discussed earlier, unambiguously identifies each county in the United States.

5.4.4.	FIPS PUB 8-5, Metropolitan Statistical Areas (Including CMSAs, PMSAs, and NECMAs)

MSA (formerly SMSA) codes are elements which identify integrated social and economic units. Each contains a unique four-digit numeric identifier plus a one-digit alphabetic population level code. Nine criteria, including population, are used to identify MSAs, which may cross state boundaries. MSAs are used to define large residential centers and other areas of dense population. In October 1984, FIPS PUB 8-5 revised the definitions of MSAs. The updated standard revised the criteria used to identify MSAs and also established three new statistical areas: the Primary Metropolitan Statistical Area (PMSA), the Consolidated Metropolitan Statistical Area (CMSA), and the New England County Metropolitan Area (NECMA). The PMSA is an area analogous to the MSA but is slightly larger. PMSA codes are formatted the same as MSAs. CMSA codes are two-digit numeric groups of PMSAs assembled together into larger integrated areas. NECMAs are separate metropolitan area codes for groups of New England towns and cities. Since more detailed information is available for New England at the town level, additional codes have been developed to allow information to be tracked by more detailed metropolitan areas. The NECMA codes are formatted the same

as MSA codes. They are not intended to replace MSA codes, but rather to be more detailed areas for New England.

5.4.5. FIPS PUB 9, Congressional Districts of the United States

Two-digit numeric codes identify the congressional districts of the U.S. The numbers are unique within each state and unique for each of the sessions of Congress. The districts may be rearranged occasionally, depending upon census results.

5.4.6. FIPS PUB 10-3/104-1, Countries and Principle Administrative Divisions

This original NIST standard, FIPS PUB 10-3, consists of a two-character alphabetic code, one for each of the geo-political entities in the world. Although the NIST standard was developed first, a coding scheme later published by the International Organization for Standardization (ISO) has received wider publicity and greater use. Recognizing this, the NIST has published FIPS PUB 104-1 which allows any agency to adopt either format as its standard.

5.4.7. FIPS PUB 19-1, Catalog of Widely Used Code Sets

This catalog lists and briefly describes code sets that are in wide use in the United States and that might be used in Federal data systems. The purpose of this catalog is to assist Federal Agencies and other organizations in selecting appropriate code sets. There are four commonly used coding schemes included in this catalog that have particular significance to EPA. These include the Hazardous Material ID Number, the Social Security Number, Human Sexes, and Zip Codes. These codes are described below.

The **hazardous material ID number** is a United Nations code for uniquely identifying various harmful substances. Since EPA programs often track information pertaining to hazardous materials, this coding system may provide a common method of identification within the EPA and may increase consistency with international conventions. The applicability to EPA's information needs and the potential usefulness of this code warrant consideration by EPA of its possible adoption as an EPA data standard.

The **Social Security Number** is a nine digit numeric developed by the Social Security Administration to uniquely identify all citizens of the United States. Many organizations across the United States and some systems within the EPA use this code to identify personnel. Although many systems at EPA do not keep records of specific people, there is applicability to EPA for internal staffing-related records where a systematic code would facilitate portability between systems. Due to its wide spread use in all sectors of the U.S., the

Social Security Number is the logical choice for such a standard. Note that all systems employing unique personal identifiers are subject to requirements of the Privacy Act (EPA's Privacy Act Manual is available from OIRM/IMSD upon request).

The **human sexes codes** provide a consistent method for documenting the sex of people. The scheme includes codes for male, female, and codes identifying that gender is not specified or not known. The applicability of human sex codes to EPA is essentially the same as for the Social Security Number. Personnel records within EPA benefit from increased portability with the adoption of a standard code. The scheme outlined in FIPS PUB 19-1 is one that provides flexibility with its additional standard codes for unknown data, and should be widely used by the EPA.

Zip Codes are five and nine digit numeric codes identifying the Postal Service areas in the United States. These codes are crucial to all mailing address information. Currently, many EPA programs have a need for mailing address information and many EPA systems store ZIP Codes for this purpose. However, some Agency systems can only store the five digit ZIP Codes and cannot accommodate the recently introduced nine digit ZIP Codes. Since ZIP Codes are useful throughout the Agency, they should be adopted as an EPA standard and their collection made mandatory for all mailing addresses. Also, EPA should uniformly update all five digit ZIP Codes to the new standard nine digits wherever practical.

5.4.8. FIPS PUB 55-2, Guidelines: Codes for Populated Places, Primarily County Divisions, and Other Locational Entities of the United States and Outlying Areas

This guideline implements ANSI X3.47-1977 and provides a two-character FIPS State Code and a five-character FIPS numeric place code to uniquely identify each listed entity. It contains an exhaustive list of names of incorporated places, census designated places (CDPs), primary county divisions (such as townships, New England towns, and census county divisions), counties, and recognized Indian reservations and Alaska Native villages. The listing also includes names of all populated places in the files of the Geographic Names Information System of the U.S. Geological Survey, as well as names of airports, military bases, national parks, and U.S. Post Offices. A two-character class code distinguishes over fifty entity types.

5.4.9. FIPS PUB 58-1, Representations of Local Time of Day for Information Interchange

This standard provides uniform time representations for both 12- and 24-hour timekeeping systems. It provides a means for representing local time of the day in digital form for the purpose of interchanging information among

data systems. It specifies the time elements and their sequencing, the use of separators between time elements and the representations of the meridiem designator.

5.4.10. FIPS PUB 59, Representations of Universal Time, Local Time Differentials, and United States Time Zone References for Information Interchange

This standard identifies the various time zones of the world. Procedures for expressing universal time (Greenwich Mean Time) and for presenting local time differential factors and time zones are given.

5.4.11. FIPS PUB 66, Standard Industrial Classification (SIC) Codes

This standard provides classifications, short titles, and codes for representing industries and groups of establishments with similar economic activities. The codes may be two to four digits in length and are left-justified. A two-digit number refers to a general industrial classification (such as Agricultural Chemicals) whereas a four-place number identifies a subclassification (such as Nitrogenous Fertilizers) in more specific detail.

5.4.12. FIPS PUB 70-1, Representation of Geographical Point Locations for Information Interchange

This standard specifies three formats for representing geographic point locations: longitude/latitude, Universal Transverse Mercator (UTM), and the State Plane Coordinate System. Of these, the first is the most widely used at the Federal level and is EPA's preferred locational coordinate system, according to EPA's Locational Data Policy (see Section 3 or Appendix A). It employs spherical coordinate representations (degrees, minutes, and seconds) to identify points on the earth's surface. The prime meridians for longitude and latitude are Greenwich, England and the equator, respectively. The UTM method is a rectangular coordinate system which uses linear measurements to specify a location. The State Plane Coordinate system was designed to define the location of points within a restricted geographical area. A standard for altitude data is also included that is represented by the vertical distance between a point and the National Geodetic Vertical Datum - roughly sea level.

5.4.13. FIPS PUB 92, Guideline for Standard Occupational Classification (SOC) Codes

This standard specifies two, three and four digit numeric codes classifying work categories. These categories are based on the actual work performed and not other factors such as skill level, place of work, licensing required, etc. The

two digit numerics are high level categories and the three and four digit numerics are further levels of subclassifications.

5.4.14. FIPS PUB 95, Codes for the Identification of Federal and Federally Assisted Organizations

Four digit numeric codes are defined for each organization funded by the United States Government. The code consists of the two-digit Treasury Agency Symbol (TAS) followed by a two-digit subdivision code. The Department of Treasury maintains this number. The types of organizations covered include the Legislative, Judicial, and Executive Branches, other Independent Federal and Quasi-Federal Organizations, Independent Federal-State and Interstate Organization, and International Organizations.

5.4.15. FIPS PUB 103, Hydrologic Units

The U.S. Geological Survey developed two to eight digit numeric codes identifying each major hydrological area in the U.S. and the Caribbean. The codes are hierarchical and begin with 2-digit codes at the top identifying the 21 major hydrological regions. 4-digit codes subdivide the regions into 222 subregions; 6-digit codes subdivide subregions into 352 accounting units; 8-digit codes further subdivide the accounting units into approximately 2,150 cataloging units. In addition to numeric codes, the standard provides names for each unit which are unique within a branch of the hierarchy.

5.4.16. Proposed FIPS for Spatial Data Transfer Standard (SDTS)

This proposed standard provides specifications for the organization and structure of digital spatial data transfer, definition of spatial features and attributes, and data transfer encoding. The purpose of the standard is to promote and facilitate the transfer of digital spatial data between dissimilar geographic information systems or other computer systems.

6. INTERNATIONAL DATA POLICIES, GUIDELINES, AND STANDARDS

6.1. Overview

There is an increasing need, both nationally and internationally, for data standards to ensure compatible environmental data. Environmental data that are compatible, and can be compared on both local and global scales, are fundamental to a number of important international and multi-disciplinary research programs. Most environmental data bases have been established to assemble or archive data for narrowly specified uses, such as monitoring or enforcing regulations. The techniques and standards used during the collection of the data in these data bases are tailored to their individual purposes, and therefore vary widely.

The increasing need for compilations of environmental information, for a wide range of subjects in regions of overlapping interests, has increased the need for ensuring compatibility among data sets. This will require that careful attention be given to the acceptance and promulgation of uniform techniques of environmental measurements and data management, along with development of a shared understanding of the environmental processes that link the characteristics being measured.

In 1990, a workshop on "Directions for Internationally Compatible Environmental Data" was held to bring together representatives from programs involving environmental data measurement and use. The workshop emphasized major programs and the measurement and data management problems associated with them. The main goals of the workshop were to assess current status and projected needs for:

- measurement standards and methods for air, water, and soil environments,

- data banks and the efficient dissemination of environmental data, and

- background measurements and global monitoring.

The workshop was subdivided into four main categories: air, water, soils, and data management (both water and soils included biota). These four areas were subsequently divided further into small working groups to maximize the opportunity for meaningful discussions. The workshop was intended to encourage further action, not only by CODATA, a subcommittee of the International Council of Scientific Unions, but also by other existing organizations that deal with environmental measurements. The workshop sought to promote better international data compatibility, quality, exchangeability and utilization.

6.2. Applicability

With the Agency's growing concern for internationally compatible data, EPA will most likely participate in the development of and use of international standards to promote environmental data compatibility.

6.3. Location

The results of this formative workshop are documented in a publication entitled, "Directions for Internationally Compatible Environmental Data", by G.C. Carter and B.I. Diamondstone. This publication is available in the Headquarters Library.

BIBLIOGRAPHY

SUPPLEMENTAL STANDARDS AND GUIDELINES FOR DATA ADMINISTRATION

Federal Information Processing Standard Publications (FIPS Pubs)

FIPS PUB 76	Guideline for Planning & Using a Data Dictionary System
FIPS PUB 77	Guideline for Planning & Management of Database Applications
FIPS PUB 88	Guideline on Integrity Assurance and Management of Database Applications
FIPS PUB 110	Guideline for Choosing a Data Management Approach
FIPS PUB 123	Specification for Choosing a Data Management Approach
FIPS PUB 124	Guideline on Functional Specifications for Database Management Systems
FIPS PUB 126	Database Language NDL (Adopts ANSI X3.133 - 1986)
FIPS PUB 127	Database Language SQL (Adopts ANSI X3.135 - 1986)
FIPS PUB 156	Information Resource Dictionary System (IRDS))

Relevant Publications

Durrell. Durell, William. Data Administration: A Practical Guide to Successful Data Management, McGraw-Hill, 1985.

Martin. Martin, James. Information Engineering, Book II: Planning & Analysis, Prentice Hall, 1990.

Date. Date, C.J. An Introduction to Database Systems, Volume I, 5th ed., Addison Wesley, 1990.

Diamondstone. Diamondstone, Barry I. Directions for Internationally Compatible Environmental Data, Hemisphere Publishing Corporation, 1990.

APPENDIX A

Part II

Environmental
Protection Agency

Policy on Electronic Reporting; Notice

ENVIRONMENTAL PROTECTION AGENCY

[FRL-3815-4]

Policy on Electronic Reporting

A. Purpose

This policy establishes a uniform Agency approach to electronic reporting by the regulated community, that is, to their submission of reports in an electronic or magnetic medium. This policy is intended to promote the adoption of electronic reporting by EPA programs, and to ensure that as they implement this technology they do so in a manner that is both consistent across the Agency and compatible with the electronic reporting practices that are current in the regulated community.

B. Applicability

This policy applies in every case where EPA programs are requesting or requiring reports from members of the regulated community, including those cases where reports are required or requested of Federal facilities. This policy applies whether the recipients of the reports are EPA programs at headquarters or in regional offices, or State or local government agencies.

However, this policy is not intended to disrupt existing systems of electronic reporting; therefore, this policy does not apply to systems in place before its effective date, or to any subsequent enhancements of such systems. In addition, this policy does not apply to transfers of data or information that are internal to EPA itself.

C. Force of This Policy

1. For EPA Programs

This policy is intended as guidance to EPA programs, that they will follow to the extent practicable. Program offices initiating nonconforming systems should be aware that the National Institute of Standards and Technology (NIST) has proposed a Federal Information Processing Standard (FIPS) for electronic data interchange (EDI). 54 FR 38424 (September 18, 1989). If promulgated, this FIPS would require all agencies procuring or upgrading EDI systems after the effective date to use either the ANSI X12 or UN/EDIFACT standards, and it is these standards that are adopted in this policy as well.

2. For the Regulated Community

This policy is intended to inform the regulated community of the general approach that EPA will take to electronic reporting. As individual programs adopt electronic reporting within the framework of this approach,

they will specify the obligations of their respondents to comply with the standards and procedures they develop.

3. For States and Local Governments

This policy is intended to inform State and local governments of the general approach that EPA will take to electronic reporting. However, this policy recognizes that as the Agency institutes electronic reporting for specific State-delegated programs, acceptable approaches will require consultation and collaboration with the affected organizations, as well as accommodation of their needs, capabilities and available resources.

D. Background

For the purposes of this policy statement, electronic reporting includes any case of transferring data (or information) using an electronic or magnetic medium. Electronic media include telecommunications transmissions as well as direct computer-to-computer links; magnetic media include disks, diskettes and tapes. In both business and government, electronic reporting is rapidly becoming the preferred mode for transferring data: compared to paper, it greatly increases the speed at which the transferred data can be processed, if only because it eliminates the need to rekey the data into a recipient's database. By eliminating this rekeying—and by eliminating the need to produce and physically transfer paper—electronic reporting usually reduces the cost and improves the accuracy of the data transfer as well.

Some of the most familiar cases of electronic reporting deal with financial data: in the world of business, the electronic transfer of funds that we now take for granted, and, in government, the increasingly prevalent electronic filing of tax returns. We may also find examples of electronic reporting at EPA: the electronic transmission of laboratory measurement results, codified in EPA Order 2180.2, and the magnetic submission of Form R Toxic Release Inventory data, provided for in the EPA guidance document, "Toxic Chemical Release Inventory Magnetic Media Submission Instructions". Nonetheless, the dominant force in electronic reporting lies elsewhere—in the realm of what is commonly known as "electronic data interchange" or "EDI".

Historically, EDI arose from efforts in certain industries to standardize the formats and procedures for the electronic transmission of such routine business documents as purchase orders and invoices. The goal was to create common definitions for these electronic

transactions that would allow participants to avoid the unacceptable costs of making and maintaining unique arrangements with each of their trading partners. As industry participation has broadened, the EDI standard setting process—at least in this country—has come to reside in the American National Standards Institute (ANSI), a nongovernmental body founded in 1918 to coordinate America's voluntary standards system. In 1979, ANSI established the X12 committee to oversee the development of standards for the electronic exchange of business information. The resulting EDI standards are known generally as the "ANSI X12" standards. There is a similar, though discrete, standard setting effort for international transactions, sponsored by the United Nations. The standards in this case are referred to as the "UN/EDIFACT" standards (for United Nations EDI for the Administration of Commerce and Transport). Given current business trends, it appears likely that over the next decade or so the use of ANSI X12 in the United States will give way to EDIFACT.

While they are not identical, ANSI and EDIFACT standards are structurally parallel, each consisting of four interrelated groups of standards: the data element dictionaries and tables, the segment directory, the transaction set standards, and the transmission control standards. Very generally:

- The data element dictionaries define the data elements that can be transmitted and provide standard codes for them; tables provide the range of standards values for certain of the data elements;
- The segment directory defines the syntax for allowable combinations of data elements—segments often corresponding to a single line on a hardcopy document;
- The transaction set standards define electronic documents—such as invoices and purchase orders—in terms of the allowable sequences of segments out of which they may be formed; and
- The transmission control standards provide the "envelope" for the electronic documents—defining such things as how the transaction sets are identified, how they may be grouped, how the sender and receiver are identified, how beginnings and ends of documents are marked, and providing procedures for starting transmission and acknowledging receipt.

Both within ANSI and EDIFACT, these four groups of standards serve primarily to set a uniform format for the content of messages. For the most part, they are equally applicable to messages

sent in a magnetic medium—such as floppy disk or 9-track tape—or transmitted via telecommunications. Moreover, by design, the standards are compatible with a very broad range of hardware/software configurations housing the senders' and recipients' databases—in environments that range from mainframe to personal computer.

Both the ANSI and EDIFACT standards are supported by an extensive array of commercially available software products that automate the process of translating between these standard formats and the particular structure of a sender's or recipient's database. Some of these products also automate the process of transmission, sometimes in conjunction with the services of a commercial network or electronic mail service. In addition, the increasing private and public sector use of the X12 standards has generated a body of accepted practices to deal with confidential business information (CBI), signature/certification, data security, and other issues raised by electronic reporting.

At least within the ANSI framework, standards now exist for a broad range of transaction sets, and includes the material safety data sheet (MSDS). The segment and data element standards are correspondingly broad; moreover, they are readily amended to cover additional data. It is especially easy to add new data element values; indeed, a number of industries maintain their own data element tables, which they can amend without prior ANSI approval.

Given these features, it is likely the ANSI framework—and increasingly the EDIFACT framework as well—will accommodate virtually all of the reporting to which this policy statement applies. Even as they stand, the ANSI data dictionaries and segment directories include a rich vocabulary and syntax for such tasks as identifying facilities, describing chemical properties and characterizing hazards, and the existing transaction sets will accommodate many if not all of EPA's standard reports. In addition, the transmission standards—to the extent that they are applicable—can probably be used without modification. The ANSI and EDIFACT frameworks, therefore, appear to offer the Agency a ready-made and widely accepted approach (or approaches) to electronic reporting. To take this approach in a particular case would largely be a matter of specifying how the particular EPA report was to be translated into existing ANSI or EDIFACT segments and data elements—as necessary, adding data elements/

values and segments conforming to ANSI/EDIFACT syntax.

This approach offers many advantages. Most obviously, it should minimize the cost and burdens of electronic reporting for the regulated community, since it would largely accommodate their existing EDI capabilities. Comments received in response to the **Federal Register** notice, "Formation of an Agency Workgroup to Consider Uniform Standards for the Electronic Reporting of Data", make this clear. 54 FR 15251 (April 17, 1989). Equally important, it should help minimize costs for EPA, providing much of the specification for electronic reports ready-made, and allowing programs to avail themselves of relatively inexpensive, commercially available EDI translation software in lieu of custom programming. In addition, it should put the Agency in better position to meet future needs for electronic dissemination of EPA information products, and give Agency programs easier access to each other's data. These considerations motivate many of the provisions of this policy.

E. General Policy on Electronic Reporting

1. Offering the Option of Electronic Reporting

EPA programs should strive to the extent practicable to offer their regulated communities the option of submitting reports electronically wherever it is reasonable to expect that this will:

• Reduce the costs and/or burdens of reporting for a significant segment of the regulated community, or

• Reduce the overall costs and/or increase the overall benefits of this reporting for the program.

2. Keeping Electronic Reporting Optional

While the cost/benefit ratio for electronic reporting generally improves as the number of participants increases, programs must do nothing to coerce members of their regulated communities into reporting electronically. Where offering the option of electronic reporting, programs must always allow paper submissions as an alternative.

F. Policy on the Approach to Electronic Reporting

1. Conforming to Established National or International Frameworks for Electronic Data Interchange (EDI)

Except in the cases described under paragraph F.3, programs implementing electronic reporting should do so within one or both of the two frameworks

provided by the ANSI X12 and UN/EDIFACT standards. Within the ANSI X12 or UN/EDIFACT frameworks, programs should ensure that:

• The format specifications they set for the electronic submission of particular reports conform—to the extent practicable—to existing data dictionaries and value tables, and to existing segment and transaction set structures;

• The submission procedures they set conform to existing transmission control standards to the extent that such standards apply;

• Any data elements or values, or segments or transaction sets that their implementations add to existing standards are—to the extent practicable—compatible with the syntax of the framework.

2. Considering Electronic Reporting Within the EDIFACT Framework

Programs implementing electronic reporting should consider the possibility that EDI participants in their regulated communities may—or may soon—be migrating from ANSI X12 to the UN/EDIFACT framework. A program should adopt whichever framework they reasonably expect will predominate in their regulated community within the relatively near term. Programs may also wish to weigh the relative costs and benefits of implementing electronic reporting within both ANSI and EDIFACT—in effect maintaining a "bi-lingual" electronic reporting system.

3. Cases Outside the ANSI and EDIFACT Frameworks

This policy recognizes that there are cases where the practices or capabilities that predominate in a regulated community—or other factors, such as the unique or nonrecurring nature of the data collection—indicate that electronic reporting outside the ANSI and EDIFACT framework is or will be the most cost-effective approach for all parties concerned. In such cases, this policy encourages programs to implement electronic reporting in a manner that best fits the special circumstances at hand—whether or not the implementation conforms to ANSI or EDIFACT standards. Programs that take such an approach, however, should ensure that:

• Their implementation is compatible with the electronic reporting practices or capabilities that predominate in their regulated community, at least for the kind of data being transferred; and

• In any case, their approach does not involve more expense or burden than

A

would implementation within ANSI or EDIFACT.

G. Policy on Hardware/Software Options for Electronic Reporting

1. Ensuring Compatibility With the Broadest Possible Range of Sender Hardware/Software

Regardless of the approach taken, programs should minimize the need for members of the regulated community wishing to report electronically to reconfigure their systems.

2. Allowing for a Broad Range of Electronic/Magnetic Media

In the interest of allowing the widest participation in electronic reporting, programs should not arbitrarily restrict the acceptable media for electronic reporting. However, programs may legitimately impose such restrictions where accepting certain media would involve unacceptable costs or administrative burdens for the Agency.

H. Taking Account of the Needs of State/Local Government Recipients of Data

This policy will apply in many cases where the recipients of data from the regulated community will include State and local government agencies. In these cases, programs should:

• Include the effects on State and local government agencies in considering the costs and the benefits of adopting electronic reporting, and

• Refrain from taking any approach to electronic reporting that will impose unacceptable costs on State or local government agencies.

I. Coordinating the Implementation of Electronic Reporting

1. Establishing an Agency Coordinating Committee for the Implementation of Electronic Reporting

When this policy becomes effective, the Electronic Reporting Standards workgroup will be reconstituted as a Coordinating Committee for the Implementation of Electronic Reporting.

2. Setting Format Specifications for Agency Reports

To the extent practicable, programs should consult with the Coordinating Committee as they set format specifications for Agency reports, to ensure that segments and data elements and values used by more than one program are represented uniformly—within the ANSI or EDIFACT frameworks—across the Agency.

3. Obtaining the Hardware and Software to Support Electronic Reporting

To the extent practicable, programs should consult with the Coordinating Committee as they obtain the hardware, software and related services that they will use to conduct their ANSI- or EDIFACT-based electronic transactions—so that they make the most of opportunities to share resources.

4. Liaison With EDI Standards Committees

The Coordinating Committee will serve as the Agency liaison for electronic reporting with other Federal agencies and with the standards committees sponsored by the ANSI and EDIFACT organizations. Where programs amend ANSI or EDIFACT standards in the course of setting specifications for particular Agency reports, they should inform the relevant standards committees of their activities by communicating through the Coordinating Committee. On a case-by-case basis, the Coordinating Committee may decide to seek formal ANSI or EDIFACT recognition for Agency amendments. However, the adoption and use of particular reporting standards are under the control of the Agency, and are not dependent upon formal ANSI or EDIFACT recognition.

J. Responsibilities for Implementation

The Office of Policy, Planning and Evaluation (OPPE), the Office of Information Resources Management (OIRM), and the programs will share the

responsibilities for implementing electronic reporting under this policy, with the advice of the Coordinating Committee established under paragraph I.1. Specifically:

• OPPE will convene and chair the Coordinating Committee, will ensure that this policy receives appropriate consideration in regulation development through the Agency Steering Committee process, and will foster compliance with this policy through the Paperwork Reduction Act clearance process;

• OIRM will provide the Coordinating Committee with technical leadership—particularly in addressing issues of systems coordination—will provide technical support to programs as they implement electronic reporting, and will foster compliance with this policy through their various information resource management program authorities;

• The programs will incorporate electronic reporting considerations into their information collection and system life cycle planning processes, with review by their senior information resources management officials (SIRMOs). The programs will develop the specifications and procedures for the electronic submission of particular Agency reports—to meet their individual information needs—in compliance with the relevant provisions of this policy. The programs will also be responsible for developing whatever procedures they need to assure CBI, data security, and privacy, or to deal with other concerns. They will be responsible for communicating these specifications and procedures to their regulated communities and to affected States and local governments.

Dated: July 20, 1990.

Dan Beardsley,

Acting Assistant Administrator.

CHAPTER 3 - STATE/EPA DATA MANAGEMENT

1. PURPOSE. This policy establishes the principles that govern the management and sharing of data between EPA and State environmental agencies and the information systems that handle these data. This policy also defines roles and responsibilities for implementing and ensuring adherence to these principles.

2. SCOPE AND APPLICABILITY. This policy is applicable to all EPA programs and Regional offices that develop and operate information systems that are used by the States or that contain data reported to EPA by States.

3. BACKGROUND.

   a. The underlying rationale for EPA's policy on State delegation includes a recognition that more effective environmental protection results when Federal goals and regulations are implemented in a fashion that is respon- sive to the diversity of local conditions. EPA's policies on information management must reflect this same balance of compliance with Federal statutes and priorities and responsiveness to local diversity.

   b. Federal policy, as most recently set forth in OMB Circular A-130, specifies that Federal agencies may "not require Federal information systems that unduly restrict the prerogatives of heads of State and local government units..."

   c. EPA remains responsible and accountable to the President, the Congress and the public for progress toward meeting national goals and for ensuring that Federal statutes are adequately enforced. In accordance with "EPA Policy on Oversight of Delegated Environmental Programs," April 4, 1984, the Agency has the responsibility to oversee the conduct of delegated inter-governmental programs, to enhance State capabilities to administer environmental protection programs and to analyze the status of State, regional and national environmental quality through ongoing monitoring and data collection efforts.

   d. EPA's policy of delegating program implementation responsibility to States means that the ultimate effectiveness of the Agency depends, to a very large

extent, on the effectiveness of State program managers. Among the several factors that determine the success of State program managers is their capacity to obtain and use management and environmental information.

e.   EPA's ability to oversee and support State performance of delegated programs, and to report on these programs to the President, the Congress and the public, is also heavily dependent upon accurate and timely State information resources and systems.

f.   EPA seeks to improve environmental decisions by more consistent and reliable estimation of health risk based on sound data and analysis methods and by integrating permitting, regulatory and compliance efforts across program lines. Improvement in the information management systems will result in more timely, quality assured data, a more integrated risk assessment and overall better State/EPA program management.

g.   Although each has requirements that differ in detail and emphasis, there are substantial benefits to EPA and to State agencies if both have timely, reliable access to the same basic management and environmental information.

h.   Most EPA programs have developed data systems to receive State reports and to provide the reports and analysis required by national program managers. There are substantial benefits to EPA when States agree to meet Agency reporting requirements by entering data directly into these systems. In at least some cases, States also benefit by gaining access to data and information systems capabilities that they cannot develop on their own. However, the benefits to States from using EPA information systems to report or to process data depend on several factors:

     (1)   The existing State investment in its own information systems

     (2)   The accessibility and reliability of the EPA systems for both entering and retrieving data

     (3)   The reliability and quality of EPA user support

(4)   The extent to which EPA systems contribute to State
      management objectives as the integration of environ-
      mental and management data, both across programs
      delegated from EPA and other State programs

(5)   The costs in using such systems both in actual
      dollars and resources necessary for use.

4. AUTHORITIES.

    a.   OMB Circular A-130, Management of Federal Information
         Resources.

5. POLICY.   It is EPA policy that Agency reporting requirements
    and information systems will be responsive to the information
    needs of State environmental agencies and will take into
    account the diversity among States in terms of organization,
    resources and program responsibilities. EPA systems that
    process and store data obtained from States will adhere to
    data management policies that avoid duplication of data and
    effort and promote integrated environmental program planning
    and management, both within States and between States and
    EPA. EPA will assure timely and reliable State access to
    any Agency information system that contains data obtained
    from States in response to EPA reporting requirements.

    a.   As required by OMB Circular A-130, EPA will adhere to
         reporting and information systems policies that do not
         unduly restrict State prerogatives to plan and manage
         information resources in response to State policy and
         management priorities.

    b.   EPA information systems that process and store data
         provided by States in response to EPA reporting require-
         ments will, insofar as practical, be developed and operated
         to accommodate State management needs. More specifically:

         (1)   EPA will ensure that States are afforded an active
               role in developing, improving and modifying informa-
               tion systems through the establishment of user
               groups, policy groups and other mechanisms which
               promote continuing State/Federal interaction.

         (2)   EPA will, insofar as practical, design such systems
               with the flexibility to accommodate State needs for
               related data standards that facilitate State informa-
               tion systems planning and the integration of data
               across EPA and State program lines.

     (3)   EPA will develop such systems in adherence to technology and data standards that facilitate State information systems planning and the integration of data across EPA and State program lines.

     (4)   EPA will design such systems to accept direct, electronic transmission of data from States that operate their own information systems.

     (5)   EPA will design such systems to support direct, electronic transmission of data to States from EPA systems to support local data analysis.

     (6)   EPA will strive to achieve consistency in design and access methods consistent with current industry technology.

c.   New EPA systems and data bases developed to process and store data obtained from State environmental agencies shall be designed to support timely and reliable State access to these data. Existing EPA systems that contain State data should allow for timely and reliable State access. Timely and reliable State access will vary according to the nature of the data and the system; however, for EPA's major national systems and data bases, it means:

     (1)   Direct, on-line State access to current data files

     (2)   The use of software and data communications technologies that adhere to Agency standards and that support efficient State access for reporting and retrieval of data

     (3)   The provision of documentation and user assistance to State users on a consistent and current basis.

d.   For those States which agree to meet EPA reporting requirements by directly entering data into EPA systems, the Agency will regard such data as the official State record of the delegated program. EPA will not unilaterally change these data, since doing so would force the State to maintain a separate system of records.

e.   EPA will allow the States at their option to enter data regarding non-delegated programs into the EPA systems.

However, States are not mandated to meet the same
requirements in the non-delegated programs that they are
obliged to meet for the delegated ones.

f.   EPA will support the use of State grant funds to develop
     State information resources and technology to the extent
     that doing so is consistent with the purposes for which
     these funds were appropriated. EPA will seek State
     proposals which assign funds from one or more EPA grants
     for information resources and technology that:

     (1) Promote the integration of environmental planning
          and management across State and EPA program lines

     (2) Foster improved data sharing between EPA and the
          State.

g.   EPA will design and manage its computing and data
     communications network to support timely and reliable
     State access to EPA systems and data bases. EPA's pursuit
     of this goal will be based on the following assumptions:

     (1) The achievement of this goal is dependent on the
          constraints of available resources.

     (2) EPA does not seek to be the primary or the "first
          choice" computing resource for any State environmental
          agency.

     (3) EPA does not seek to provide computing and
          telecommunications services to States in lieu of
          or in competition with either State or commercial
          sources.

h.   EPA recognizes one of the advantages of sharing data is
     reduced reporting by the States. Therefore, if a State
     is entering data directly into the EPA system, EPA will,
     insofar as practicable, adhere to data management policies
     that avoid duplication of data and effort and not require
     that the State report this information in additional
     formats.

6. RESPONSIBILITIES.

    a.   The Office of Information Resources Management shall:

         (1) Develop guidelines and programs to ensure that

Agency reporting requirements and information systems
are defined and implemented in accord with this
policy.

(2) Provide guidance and assistance to Assistant
     Administrators, Associate Administrators and Regional
     Administrators in implementing the requirements
     of this policy.

(3) Plan and oversee the acquisition, deployment and
     use of information technology within EPA to ensure
     support for effective management and sharing of
     data by EPA and State environmental agencies.

(4) Ensure EPA compliance with Federal statutes and
     regulations governing the acquisition, operation
     and use of information technology employed to share
     data between EPA and State agencies.

(5) Evaluate and report on the effectiveness of Agency
     activities in achieving the goals of this policy.

b. National Data Processing Division shall:

(1) Design and manage the acquisition and operation of
     data processing and telecommunications resources to
     support effective management and exchange of data
     between EPA and State environmental agencies.

(2) Develop standards for EPA data processing and
     telecommunications technology services that support
     the goals of this policy.

(3) Provide technical advice and assistance to EPA and,
     upon request, to State environmental agencies
     concerning the acquisition and implementation of
     information technology to achieve the goals of this
     policy.

c. Assistant Administrators and Associate Administrators
    shall assure:

(1) That State agency requirements for information and
     information technology are addressed in the design
     and implementation of EPA programs.

      (2)   That the information systems and data management
            practices of programs and activities under their
            direction are in accord with this policy.

      (3)   Effective State participation in the design and
            operation of national information systems and data
            bases that contain data reported by States and
            provide timely and reliable access by States to
            such data bases.

d.   Regional Administrators shall assure that:

      (1)   State requirements for information and information
            technology are effectively addressed in State delega-
            tion agreements, State grants and other agreements
            between EPA and States.

      (2)   Regional procedures for handling and validating
            State-reported data guarantee the integrity and
            accessibility of such data as required by this
            policy.

      (3)   The Regional Office has an effective program to
            foster and support State/EPA data management and
            sharing that meets at a minimum EPA Federal reporting
            requirements.

e.   The Office of Administration shall:

      (1)   Develop and implement policies and procedures to
            assure that information collection and processing
            activities performed by EPA contractors and grantees
            comply with this policy.

7. DEFINITIONS.

a.   "Data" refers to a collection of unorganized facts that
     have not yet been processed into information.

b.   "Data Base" is a collection of integrated data that can
     be used for a variety of applications.

c.   "Data Communications" refers to computer-to-computer,
     computer-to-device, device-to-computer communications
     and other communications such as a record, tele-processing
     and telemetry.

      d.   "Information Technology" refers to the hardware and
         software used in connection with government information,
         regardless of the technology involved, whether computers,
         telecommunications, micrographics or others.

      e.   "Software" refers to computer programs, procedures,
         rules and associated documentation pertaining to the
         operation of a computer system.

      f.   "Telecommunications" is the transmission and/or reception
         of information by telephone, telephone lines, telegraph,
         radio or other methods of communication over a distance.
         The information may be in the form of voice, pictures,
         text and/or encoded data.

8.   **PROCEDURES AND GUIDELINES.** Procedures and guidelines will
    be issued under separate cover.

CHAPTER 5 – DATA STANDARDS

1. PURPOSE. This policy establishes the EPA Data Standards Program. The purpose of this program is to provide consistent definition of data and to facilitate cross-media use of data. This policy sets forth Agency principles on data standards and assigns organizational responsibilities for implementing and administering common data standards.

2. SCOPE AND APPLICABILITY. This policy applies to all Environmental Protection Agency (EPA) organizations and their employees. It also applies to the facilities and personnel of agents (including contractors and grantees) of EPA who design, develop, operate or maintain Agency information and information systems. This policy applies to automated and manual systems developed for programs or administrative purposes. The requirements of this policy apply to existing data elements as well as new data elements.

3. BACKGROUND.

   a. Integration of information and data bases is difficult because program offices use disparate formats and names for similar data elements.

   b. There is a need to make and support decisions based on standard information and data collected that cuts across the Agency's programs.

   c. Specific programs, such as the Ground-water program, have an increasing need to share data from other programs, other agencies, States and local governments. This adds credence to the need for acceptable data standards to facilitate exchange of information.

   d. Information technology has reached a point at which the sharing of data among automated systems is technically feasible.

   e. The Agency has implemented standards for hardware and software that facilitate the sharing of data among programs.

   f. To support effectively the use of common definitions of environmental data with State programs, EPA must have common definitions for data elements and an intra-agency capability to share data.

g. Organizations outside EPA have been establishing data
    standards which are accepted nationally or internationally.
    These pre-existing standards, such as Chemical Abstract
    Service (CAS) registry numbers, may serve as the best
    data standard for certain data elements.

h. There is a growing need for agreement on the definition
    of Agencywide parametric data entities such as "site"
    and "facility."

i. The Agency has a facilities inventory system that lists
    facilities regulated by the various programs in EPA.
    The inventory includes the different names and addresses
    for a single facility. This system will be a critical
    part of the Agency data standards effort.

j. At a minimum, there are six major areas which would
    benefit from the use of data standards: data used in
    more than one program, facilities and site data, geographic
    data, measurement data, health and environmental effects
    data and core office systems data.

4. **AUTHORITIES.**

a. 15 CFR, Part 6 Subtitle A, Standardization of Data Elements
    and Representations.

b. OMB Circular A-130, Management of Federal Information
    Resources.

5. **POLICY.**   It is EPA policy to create and maintain consistency
    in the form of data elements that have more than one applica-
    tion within the Agency. This consistency will permit the
    cross media approach necessary to achieve environmental
    results. The data standards will reflect the Agency's program
    priorities.

a. As required by OMB Circular A-130, EPA will adhere to
    Federal Information Processing Standards (FIPS), except
    where it can be demonstrated that the costs of using a
    standard exceed the benefits of the standard or will
    impede the Agency in accomplishing its mission.

b. All organizational components of EPA, their contractors
    or grantees will promote the full utilization of Federal
    and Agency data standards and representations in the
    design and development of information systems.

    c.   Data elements, codes and representations already in use
        by the Agency will be evaluated and adopted as Agency
        standards wherever practicable.

    d.   Data elements, codes and representations may be recommended
        for standardization by any program office within EPA.

    e.   Geographical information systems developed by the Agency
        must conform to an established set of appropriate data
        standards which permit the use of the system by all
        relevant programs and State agencies.

    f.   All relevant facilities or sites data must be stored in
        the Agency's facility or site inventory systems.

6.   RESPONSIBILITIES.

    a.   The Office of Information Resources Management (OIRM)
        shall:

       (1) Provide effective leadership in developing,
            promulgating and enforcing the policies of the
            Agency data standards program.

       (2) Coordinate the evaluation and approval process of
            all data standards with the Assistant Administrators,
            Regional Administrators, Office Directors and Senior
            Information Resources Management Officers.

       (3) Exercise final approval authority for the adoption
            of data standards. Grant waivers to the implementa-
            tion of approved Agency data standards.

       (4) Support other EPA data administration efforts, e.g.,
            encourage cross reference files for non-standard
            information. Encourage the use of data element
            dictionaries.

       (5) Propose and apply effectively data elements or
            representations for use by more than one organizational
            component of EPA as Agency standards.

       (6) Publish and promulgate approved Agency standards in
            an EPA Data Standards Catalog.

    b.   Assistant Administrators, Associate Administrators,
        Regional Administrators, Laboratory Directors, Headquarters
        Staff Office Directors, General Counsel, Inspector General,
        and SIRMOs shall:

(1) Implement approved Agency data standards that are published under the provisions of this policy.

(2) Establish an organization-wide data standards work group which reviews and provides information and comments on proposed data standards.

(3) Propose the adoption of data standards for Agency use within the environmental community.

(4) Submit requests for waivers or deferments to the use of Agency data standards to OIRM.

7. DEFINITIONS.

a. "Data Element" is a unit of information used to describe data characteristics and attributes, e.g., eyes - blue or BL.

b. "Data Standards" are standards used generally, but not exclusively, for automated systems to ensure that one type of data is defined the same way in all systems. A similar definition means having the same name, the same number of maximum characters and the same type and content of data in all systems where a specific data item appears.

c. "Information Technology" refers to the hardware and software used in connection with government information, regardless of the technology involved, whether computers telecommunications, micrographics or others.

d. "Media" means Water, Air, Hazardous Waste and Pesticides and Toxic Substances program offices.

e. "System" is the organized set of procedures used to collect, transmit and disseminate information whether automated or manual.

8. PROCEDURES AND GUIDELINES. Procedures and guidelines for the Agency data standards program will be issued under separate cover.

CHAPTER 13 - LOCATIONAL DATA

1.   PURPOSE. This policy establishes the principles for
     collecting and documenting latitude/longitude
     coordinates for facilities, sites and monitoring and
     observation points regulated or tracked under Federal
     environmental programs within the jurisdiction of the
     Environmental Protection Agency (EPA). The intent of
     this policy is to extend environmental analyses and
     allow data to be integrated based upon location, thereby
     promoting the enhanced use of EPA's extensive data
     resources for cross-media environmental analyses and
     management decisions. This policy underscores EPA's
     commitment to establishing the data infrastructure
     necessary to enable data sharing and secondary data use.

2.   SCOPE AND APPLICABILITY. This policy applies to all
     Environmental Protection Agency (EPA) organizations and
     personnel of agents (including contractors and grantees)
     of EPA who design, develop, compile, operate or maintain
     EPA information collections developed for environmental
     program support. Certain requirements of this policy
     apply to existing as well as new data collections.

3.   BACKGROUND.

     a.   Fulfillment of EPA's mission to protect and improve
          the environment depends upon improvements in cross-
          programmatic, multi-media data analyses. A need
          for available and reliable location identification
          information is a commonality which all regulatory
          tracking programs share.

     b.   Standard location identification data will provide
          a return yet unrealized on EPA's sizable investment
          in environmental data collection by improving the
          utility of these data for a variety of value-added
          secondary applications often unanticipated by the
          original data collectors.

     c.   EPA is committed to implementing its locational
          policy in accordance with the requirements
          specified by the Federal Interagency Coordinating
          Committee for Digital Cartography (FICCDC). The
          FICCDC has identified the collection of
          latitude/longitude as the most preferred coordinate
          system for identifying location. Latitude and
          longitude are coordinate representations that show
          locations on the surface of the earth using the
          earth's equator and the prime meridian (Greenwich,
          England) as the respective latitude and longitude
          origins.

d.   The State/EPA Data Management Program is a
      successful multi-year initiative linking State
      environmental regulatory agencies and EPA in
      cooperative action. The Program's goals include
      improvements in data quality and data integration
      based on location identification.

e.   Readily available, reliable and consistent location
      identification data are critical to support the
      Agencywide development of environmental risk
      management strategies, methodologies and
      assessments.

f.   OIRM is committed to working with EPA Programs,
      Regions and Laboratories to apply spatially related
      tools (e.g., geographic information systems (GIS),
      remote sensing, automated mapping) and to ensure
      these tools are supported by adequate and accurate
      location identification data. Effective use of
      spatial tools depends on the appropriate collection
      and use of location identifiers, and on the
      accompanying data and attributes to be analyzed.

g.   OIRM's commitment to effective use of spatial data
      is also reflected in the Agency's comprehensive GIS
      Program and OIRM's coordination of the Agency's
      National Mapping Requirement Program (NMRP) to
      identify and provide for EPA's current and future
      spatial data requirements.

4.   AUTHORITIES.

      a.   15 CFR, Part 6 Subtitle A, Standardization of Data
            Elements and Representations

      b.   Geological Survey Circular 878-B, A U.S. Geological
            Survey Data Standard, Specifications for
            Representation of Geographic Point Locations for
            Information Interchange

      c.   Federal Interagency Coordinating Committee on
            Digital Cartography (FICCDC)/U.S. Office of
            Management and Budget, Digital Cartographic Data
            Standards: An Interim Proposed Standard

      d.   EPA Regulations 40 CFR 30.503 and 40 CFR 31.45,
            Quality Assurance Practices under EPA's General
            Grant Regulations

5.   POLICY.

    a.   It is EPA policy that latitude/longitude
        ("lat/long") coordinates be collected and
        documented with environmental and related data.
        This is in addition to, and not precluding, other
        critical location identification data that may be
        needed to satisfy individual program or project
        needs, such as depth, street address, elevation or
        altitude.

    b.   This policy serves as a framework for collecting and
        documenting location identification data. It
        includes a goal that a 25 meter level of accuracy be
        achieved; managers of individual data collection
        efforts determine the exact levels of precision and
        accuracy necessary to support their mission within
        the context of this goal. The use of global
        positioning systems (GPS) is recommended to obtain
        lat/longs of the highest possible accuracy.

    c.   To implement this policy, program data managers
        must collect and document the following
        information:

        (1)   Latitude/longitude coordinates in accordance
                with Federal Interagency Coordinating
                Committee for Digital Cartography (FICCDC)
                recommendations. The coordinates may be
                present singly or multiple times, to define a
                point, line, or area, according to the most
                appropriate data type for the entity being
                represented.

                The format for representing this information
                is:

                +/-DD MM SS.SSSS (latitude)
                +/-DDD MM SS.SSSS (longitude)

                    where:

                    •   Latitude is always presented before
                          longitude

                    •   **DD** represents degrees of latitude;
                          a two-digit decimal number ranging
                          from 00 through 90

                    •   DDD represents degrees of
                          longitude; a three-digit decimal
                          number ranging from 000 through 180

- **MM** represents minutes of latitude or longitude; a two-digit decimal number ranging from 00 through 60

- **SS.SSSS** represents seconds of latitude or longitude, with a format allowing possible precision to the ten-thousandths of seconds

- **+** specifies <u>latitudes north</u> of the equator and <u>longitudes east</u> of the prime meridian

- **-** specifies <u>latitudes south</u> of the equator and <u>longitudes west</u> of the prime meridian

(2) Specific method used to determine the lat/long coordinates (e.g., remote sensing techniques, map interpolation, cadastral survey)

(3) Textual description of the entity to which the latitude/longitude coordinates refer (e.g., north-east corner of site, entrance to facility, point of discharge, drainage ditch)

(4) Estimate of accuracy in terms of the most precise units of measurement used (e.g., if the coordinates are given to tenths-of-seconds precision, the accuracy estimate should be expressed in terms of the range of tenths-of-seconds within which the true value should fall, such as "+/- 0.5 seconds")

d. Recommended labelling of the above information is as follows:

- "Latitude"
- "Longitude"
- "Method"
- "Description"
- "Accuracy."

e. This policy does not preclude or rescind more stringent regional or program-specific policy and guidance. Such guidance may require, for example, additional elevation measurements to fully characterize the location of environmental observations.

f. Formats, standards, coding conventions or other specifications for the method, description and accuracy information are forthcoming.

6.   RESPONSIBILITIES.

      a.   The Office of Information Resources Management
            (OIRM) shall:

            (1)   Be responsible for implementing and supporting
                  this policy

            (2)   Provide guidance and technical assistance
                  where feasible and appropriate in implementing
                  and improving the requirements of this policy

      b.   Assistant Administrators, Associate Administrators,
            Regional Administrators, Laboratory Directors and
            the General Counsel shall establish procedures
            within their respective organizations to ensure
            that information collection and reporting systems
            under their direction are in compliance with this
            policy.

            While the value of obtaining locational coordinates
            will vary according to individual program
            requirements, the method, description and accuracy
            of the coordinates must always be documented. Such
            documentation will permit other users to evaluate
            whether those coordinates can support secondary
            uses, thus addressing EPA data sharing and
            integration objectives.

7.   WAIVERS. Requests for waivers from specified provisions
      of the policy may be submitted for review to the
      Director of the Office of Information Resources
      Management. Waiver requests must be based clearly on
      data quality objectives and must be signed by the
      relevant Senior IRM Official prior to submission to the
      Director, OIRM.

8.   PROCEDURES AND GUIDELINES. The Findings and
      Recommendations of the Locational Accuracy Task Force
      supplement this policy. More detailed procedures and
      guidelines for implementing the policy are issued under
      separate cover as the Locational Data Policy
      Implementation Guidelines.

⊌EPA
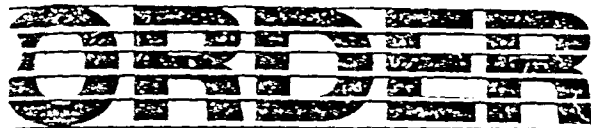
ORDER

ssification No.: 2180.1

oroval Date: 6/26/87

CHEMICAL ABSTRACT SERVICE REGISTRY NUMBER DATA STANDARD

1. PURPOSE. This Order establishes the policy and responsibilities related to the use of registry data, specifically the Registry Number, from the Chemical Abstracts Service (CAS) Division of the American Chemical Society in automated information systems containing data/information on specific, definable chemical substances. The CAS registry number provides consistent and unambiguous identification of chemicals and facilitates sharing chemical information across programmatic media.

2. SCOPE AND APPLICABILITY. This data standard applies to all EPA organizations and their employees as well as the facilities and personnel of agents of EPA who design, develop, operate or maintain Agency information and information systems having to do with specific, definable chemical substances. It also applies to EPA contractors and assistance recipients.

3. REFERENCE.

    Chapter 5 of the EPA IRM Policy Manual sets forth the general principles on data standards within the Agency. This Order relates to one of several data standards used by EPA.

4. BACKGROUND.

    a. Information technology makes efficient sharing of data among automated systems feasible, provided there exist common definitions for data elements among the systems.

    b. The CAS Registry Number is accepted nationally and internationally as an identifier for specific, definable chemical substances.

    c.   Agency programs need access to information about chemicals
       from other programs, agencies, States and local govern-
       ments. There is a need for a chemical identification
       data element standard to facilitate access to chemical
       information.

5. POLICY.

    a.   Any computer-based Agency system currently in use or
       being planned containing data/information on specific
       definable chemical substances shall contain the current
       CAS registry number for each chemical substance.

    b.   Additional data selected from the CAS chemical registry
       system, such as CA Index Names and Synonyms, Molecular
       Formulas and the CAS Chemical Registry Records, are
       optional.

6. RESPONSIBILITIES.

    a.   The Office of Information Resources Management (OIRM)
       shall:

       (1) Provide effective leadership in developing,
           promulgating and enforcing this data standard.

       (2) Provide guidance and technical assistance in
           implementing the requirements of this data standard.

    b.   Assistant Administrators, Associate Administrators,
       Regional Administrators, Laboratory Directors, Headquarters
       Staff Office Directors, the General Counsel, the Inspector
       General and Senior IRM Officials shall establish procedures
       within their respective organizations to ensure compliance
       with the requirements of this data standard.

    c.   The Procurement and Contracts Management Division and the
       Grants Administration Division shall establish procedures
       to ensure that contractors and assistance recipients are
       in compliance with the requirements of this data standard.

7. DEFINITIONS.
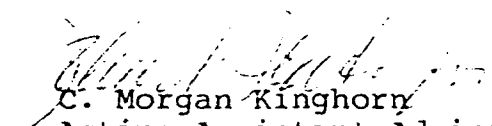
    a.   "CAS Registry Number" refers to a unique, identifying
       number assigned by CAS to each distinct chemical substance
       recorded in the CAS Chemical Registry System.

b.   "CA Index Names and Synonyms" refers to the names recorded
     in the CAS Chemical Registry System for each registered
     substance. These names include the CA Index Names and
     associated commercial names, author-designated names,
     acronyms, etc.

c.   "Molecular Formula" refers to the formula which lists the
     kind and number of atoms in a molecule.

d.   "CAS Unique Chemical Registry Record" refers to the
     machine record that represents the chemical structure in
     the CAS Chemical Registry System. The fundamental
     component is the connection table which is a detailed
     inventory of the atoms and bonds that comprise the two-
     dimensional representation of the substance's structure.
     Additional components of the structure record include
     the description of such items as stereochemistry, abnormal
     mass values and charges.

8.   PROCEDURES.

The Information and Management Services Division of OIRM
will provide any additional technical guidance for implementing
the requirements of this data standard.

C. Morgan Kinghorn
Acting Assistant Administrator
  for Administration and
  Resources Management

⊕EPA

ORDER

Classification No.: 2180.2

Approval Date: 12/10/87

DATA STANDARDS FOR THE ELECTRONIC TRANSMISSION OF
LABORATORY MEASUREMENT RESULTS

1. PURPOSE. The purpose of this Order is to issue standards for the electronic transmission of environmental measurement results from laboratories to EPA programs. These standards will provide a consistent definition of laboratory data and will facilitate cross-media use of laboratory data.

2. SCOPE. This Order applies to laboratories that supply measurement data for Agency, Regional or program office decisions.

3. BACKGROUND.

   a. Integration of information and dat bases is difficult because program offices use disparate formats and names for similar data elements.

   b. There is a need to make and support decisions based on standard information and data collected which cut across the Agency's programs.

   c. Specific programs have an increasing need to share data from other programs, other Agencies, States and local governments. This adds credence to the need for acceptable data standards to facilitate the exchange of information.

   d. Information technology has reached a point at which the sharing of data among automated systems is technically feasible.

   e. The Agency has implemented standards for hardware and software which facilitate the sharing of data among programs.

   f. Laboratory measurement results are commonly acquired by almost all the operating programs and Regions.

g. The large quantity of data that is received from
    laboratories mandates the use of automated systems of
    transmission to decrease errors of transcription, to
    increase the speed of reporting and to facilitate wide
    use of the data.

h. A standard approach to the transmission of laboratory
    data is required to ensure that all measurement data
    reported to Agency programs from laboratories will
    include common elements that define the sample type,
    the measurement technique and method, and the quality
    of the measurement, in addition to the measurement
    results.

i. These standards define data originally acquired for
    one specific purpose to other potential users. Use of
    these standards certifies the existence of qualifying
    information to second and third party users of the data.

4. AUTHORITIES.
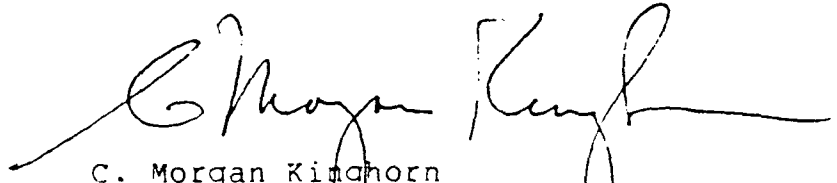
    a. 15 CFR, Part 6 Subtitle A, Standardization of Data
        Elements and Representations.

    b. OMB Circular A-130, Management of Federal Information
        Resources.

5. POLICY. The Standards for Electronic Transmission of Laboratory
    Measurement Results in Appendices A through C to this Order
    will be used to move measurement results from laboratories to
    program offices.

    a. Programs will adhere to the standards except where it
        can be demonstrated that the costs of using the standards
        exceed the benefits or will impede the Agency in
        accomplishing its mission.

    b. These standards provide a framework that can be adapted
        to the needs of each program. Addition or deletion of
        data elements is permissible within the standards.

    c. No timetables are set for conversion of existing automated
        data transfer mechanisms. The Office of Information
        Resources Management will ensure that new instances
        of the automated collection of laboratory measurement
        results will consider these standards as part of the
        workplan.

6. ASSISTANCE. Assistance in implementing this Order can be obtained from the Immediate Office, Office of Information Resources Management (OIRM).

7. PROVISION FOR WAIVER. OIRM recognizes that due to variances in mission needs, information requirements, and resource allocations, not all information systems can easily conform to the standards defined in this policy. In order to provide a reasonable amount of flexibility, this provision for waiver is included in the Order.

   a. An application for waiver should provide information to substantiate the problems encountered in adopting the standard. Also, the application should include the program's alternate plan of action for transmitting laboratory results.

   b. The application must be approved by the decision official in the requesting office and the respective Senior Information Resources Management Official prior to submission to the Director of OIRM, who has responsibility for final disposition. The applying office will be notified in writing of the disposition of the waiver within 30 days.

C. Morgan Kinghorn
Acting Assistant Administrator

CONTENTS OF APPENDICES

Appendix A

Format for Analytical Results Reports on Machine Readable Media

Introduction

This constitutes an EPA standard for media and record formats to be used in transmission of analytical results. The following points should be noted:

1. The standard describes transmission formats only. It is expected that processing systems will convert the input records into forms more convenient for storage and processing.

2. Spaces between fields permit these records to be prepared by programs written for laboratory automation systems in versions of BASIC which require this feature, as well as to be compatible with Agency standard statistical and database management systems (e.g., SAS, S2K, ADABAS, etc.).

3. Record formats contain sequence numbers and checksums to be consistent with requirements for a future error-free telecommunications format.

Media Format

The record formats are intended to be general for a variety of media, but some special considerations apply to certain media:

1. Magnetic Tapes shall be industry - standard 9-track, 800, 1600, or 6250 bits per inch, with no internal labels. Floppy diskettes shall be IBM-PC compatible and may be of any standard size. Telecommunications requirements will be defined as appropriate. Data compression or "squeezing" algorithms will be employed where appropriate for future telecommunications protocols.

2. Records shall be fixed-length 80-byte records consisting of ASCII characters. If the operating system producing the record requires an end-of-record code (such as carriage return and/or line feed), this code shall occupy record positions 79-80. Otherwise, positions 79-80 shall be blank.

3. Records on tape may be combined optionally into fixed-length blocks, with a blocksize not exceeding 4000 bytes. If the block includes a prefix or postfix supplied by the operating system in addition to the records, information about the presence and length of the prefix or postfix shall be included in the external label.

4. Tapes or diskettes shall consist of one or more files. Each tape file shall end with a tapemark; the last file on the tape shall end with two tapemarks. Diskettes shall have all files present in the root or parent directory.

5. Each tape reel or diskette shall bear one or more external labels, collectively supplying the following information: volume ID, number of files, creation date, and name, address and phone number of submitter. Magnetic tape labels shall also contain density, blocksize and recordsize. Individual Agency environmental monitoring programs may require additional external labels such as to provide linkage to other related data (e.g., field sampling data sheets or lab "chronicles").

6. The following media shall be compliant with Federal Information Processing Standards (FIPS) cited below:

FIPS	Subject
3	800 BPI, NRZI, 9-track tape
25	1600 BPI, PE, 9-track tape
50	6250 BPI, GCR, 9-track tape

Record Formats

There are six groups of record types in the standard, as shown below. Detailed record formats follow.

Type	Name	Contents
10	Run Header	Contains information pertinent to the whole production run (group, batch, etc. of samples or sample equivalents). See production run definition below.
20	Sample Header	Contains sample-identifying information or corresponding information for calibrations, QC samples, instrument performance checks, etc.
30	Results Record	Contains any final result on a sample, calibration or QC sample and identifying information.
40	Deleted Record	Signals a deleted record; record contents are undefined except for the record type code.
50	Special Record	Signals a header for other Agency Standard Data Base Records (e.g., STORET, SAROAD, SFC, AIRS, etc.).
90	Comments Record	Signals a record containing free-form comments.

Record types 10, 20 and 30 are mandatory, except when field sampling data only are being reported, in which case type 30 may be missing; other types are optional. Type 20, representing the sample, contains a Region and Sample ID which acts as an identifying label for the sample. The QC code indicates whether the data are from an environmental sample, calibration or QC sample; or other calculated run-wide data such as mean response factors. Type 30, representing an individual analyte, contains either a program or contract specified identifier or a CAS code and an indicator ("I" or "C" or another code) as to which code was used. Type 50 is used to include data from any other standard agency data base such as STORET or AIRS. It is required only when records from these other systems are being mixed with records from this standard. It should be noted that records which are optional in the standard may be considered mandatory in a given application (e.g., Contract Lab Program). See page C-5 for an example of the sequence of the record types.

Production Runs

Since, under this standard, a file contains the results for one production run, it is necessary to define a production run in terms applicable across a wide variety of analysis types. In general, a production run should represent a "group" or "batch" of samples that are processed in a continuous sequence under relatively stable conditions. Specific points characterizing a production run are:

° Calibration – initial and continuing checks. Typically all samples in a run use the same calibration data. (There will be a few exceptions, such as isotope dilution for GC/MS, where some of the calibration information is contained in each sample.)

° Method number – (see Appendix B) will be constant.

° Instrument conditions – are typically constant throughout a run. Results obtained on different instruments cannot be combined in one run.

The time span of a production run varies with the type of analysis. Many runs for inorganic analyses take a fraction of a day. Some organic analyses, such as GC and GC/MS, take a long time for each sample, so that the production run may contain data from many work shifts which could span days or weeks.

The first record in each file must be a record type 10, the Run Header. Positions 4-24 form an identifier for the run. Ignoring the blanks, this would read "8404011521GC/MS " for a GC/MS run started at 3:21 p.m. on April 1, 1984. If data from a single production run are split and reported on several files (presumably at different times), it is mandatory that this run identifier be identical on each such file. The measurement type is general and will be assigned by EPA. In runs completed during one work shift by one individual, the initials designate the responsible analyst. For runs which involve more than one instrument operator, it may be necessary to use the initials of a manager. In any case, the initials should indicate one individual responsible for the quality and consistency of the entire run.

Record Sequence   (see page C-5)

1. A Run Header (type 10) record must be present as the first record in the file. Further occurrences of the type 10 record in the file are not allowed.

2. Each environmental sample, calibration or quality control sample is represented by a group composed of a type 20 and 21 record, which holds sample level identifying information, followed by one type 30 record for each method analyte or standard. The region/client and EPA sample ID together should uniquely identify a single sample, but there is no separate requirement that the sample ID be unique on a national level. The type 20 record holds a count for the number of method analytes being determined. Type 20 records should occur in the order in which analytical results were obtained. The type 20 records for quality control items have further rules (see Appendix E, for definitions of QC types):

   a. LD1 must occur before the corresponding LD2 record, but the two records need not be adjacent. (Similar rule for FD1 and FD2)

   b. LF1 must occur before the corresponding LF2 record, but the records need not be adjacent.

   In addition, a type 20 record is used as a header for any additional run-wide data that must be reported for each method analyte (such as detection limits or interelement correction factors). Unique identifiers given on page E-6 are used in place of "QC codes" to indicate the types of data that follow.

3. Type 50 records are used to indicate the presence of data in formats of other existing agency data bases and may occupy any position. Each contains a counter to indicate the number of records from the other system that follow.

4. Type 90 records may be defined to occupy any position except before the type 10 (header) record, or between records following a type 50.

File/Record Integrity

All record types (excepting those following type 50) shall contain the following check fields to ensure file and record integrity:

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type or identifier	"10" or as appropriate
72-74	3	Record sequence number within file	000-999, repeated as necessary
75-78	4	Record checksum	Four hexadecimal digits; calculation algorithm to be supplied
79-80	2	Reserved for operating system use	Will contain blanks, or a code for CR and/or LF

Dates and Times

Wherever a date or time-of-day is required, the information consists of successive groups of two decimal digits each, separated by blanks. Dates are given in the order YY MM DD, and times as HH MM. All hours will be given as 0 to 23, right justified, using a 24 hour clock and will be local time. Since some computers generating the date and time sequence may have difficulty producing leading zeros, these will not be required. The program reading the file will convert leading blanks to leading zeros in all date and time fields.

Necessary Information

The exact list of reportable information will obviously vary considerably from one program to another. The information given on the following records is designed to be as general as possible, and not all of it will apply to any program or method. It is important to note that this standard is in no way attempting to determine, or even suggest, what data should or should not be reported for any given program; it is only defining how that data should be reported. Any data element that is not applicable should simply be left blank; if no data on a record type are applicable the entire record may be omitted. All of the definitions of the field contents should be considered to be general; specific programs and methods may further define any field, or may require the use of some fields to represent program or method specific information. Additional method dependent record types may be defined in the future to accommodate information which cannot be reported using the defined format.

Field Sampling Data

Field sampling data will also be reported using this standard. If the field sampling data are sent in separately, then the file will be structured in the same manner as an analytical analysis. There will be a type 10 record at the start which will have "FIELD" in columns 19-23 along with whatever information is appropriate. All type 20 records will have the appropriate field QC codes along with an appropriate sample qualifier, e.g., FLD". Much of the other information will be blank. Type 30 records will be present only if necessary (such as to report the amount of field spikes for each spiked analyte). If field sampling data are reported by the laboratory performing the analysis, using the same file, it will be necessary to have two type 20 records for each sample, one for the analytical results and one for field sampling data.

Multiple Volume Data

There is no requirement under this standard that all the data from an entire production run fit onto a single volume of the transmission medium. If data are being split into multiple volumes, then each program will define how this is to be performed. For example, if the multiple volumes are reported at different times, it may be necessary to repeat the transmission of all initial calibration data with each volume. On the other hand, if multiple volumes are utilized simply because all data will not fit onto a 360 K diskette, then there would be no need to repeat the initial calibration data on each volume. In all cases, the program will define when and where data may be split and how the files are to be named so that the sequence is unambiguous. What is necessary, is that all volumes start with a type 10 record, and that all type 10 records have the same run identifier as explained on page A-3. If it is necessary to split the data from a single sample into multiple volumes, then the type 20 (and following) type records for that sample must be repeated; in this situation, it is mandatory that columns 4-37, which collectively identify the sample, be identical in each volume.

General Instructions

1. All character data are to be upper case, except in comment fields where no restrictions are given or when using the symbols for chemical elements (one upper case letter or one upper case letter followed by a lower case letter).

2. Missing or unknown values are to be left blank.

3. All character fields are to be left justified.

4. All numeric fields are to be right justified. A decimal point is to be used with a non-integer if exponential notation is not used. Commas are not allowed.

5. All temperature fields are in centigrade and are presumed non-negative unless preceded by a minus sign (-).

Format of the Mandatory Production Run Header Record (Type 10)

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"10"
3	1	blank	
		Positions 4 through 17 contain the date/time of the start of instrumental analysis.	Positions 4-24 constitute the run ID. See instructions for record type 10, page A-3.
4-5	2	Year	YY
6	1	blank	
7-8	2	Month	MM
9	1	blank	
10-11	2	Day	DD
12	1	blank	
13-14	2	Hour	HH
15	1	blank	
16-17	2	Minute	MM
18	1	blank	
19-24	6	Measurement Type or Agency Code	General descriptor (e.g., ICAP, GC/MS, ASTM, USGS); or
25	1	blank	"FIELD" if field data only.
26-30	5	Method Number	Standard number defined by EPA or other Agency. (see page B-1 for examples).
31	1	blank	
32-34	3	Person responsible for run	3 initials of Manager.
35	1	blank	
36-41	6	Lab ID	From EPA standard list or Project Officer.
42	1	blank	
		Positions 43-51 contain the date report prepared.	
43-44	2	Year	YY
45	1	Blank	
46-47	2	Month	MM
48	1	Blank	

Format of the Mandatory Production Run Header Record (Type 10) cont.

Record Position	Field Length	Field Contents	Remarks
49-50	2	Day	DD
51	1	Blank	
52-61	10	Contract Number	Agency standard number.
62	1	blank	
63-68	6	Instrument ID	e.g., GC8312; provided by contract lab; must be unique and permanent within lab.
69	1	blank	
70	1	Security code	"S" = secure, "U" = unsecure Other codes may be defined to comply with additional contract requirements.

Format of the Chromatography Record (Type 11)

Use: To describe Chromatograph conditions. Applies to a group of samples in a run. Will be present for any method involving chromatography.

Position: Follows type 10

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"11"
3	1	blank	
4-11	8	Commercial Column name	e.g. SP2330
12	1	blank	
13-16	4	Column Length in meters	e.g., 100 or 99.5 or 3.5
17	1	blank	
18-21	4	Column inside diameter in mm.	e.g., 2 or .3
22	1	blank	
23	1	Type of Injector	S = split
24	1	blank	L = splitless
			O = on column
25-26	2	Carrier Gas	Chemical Symbol
27	1	blank	e.g. He, Ar, N, H
28-30	3	Carrier Gas flow rate in mL/min or Cm/sec	nnn
31	1	blank	
32-33	2	Units code	"ML" or "CM"
34	1	blank	
35-37	3	Initial Column Temp. in degrees C	e.g., 50 or 300
38	1	blank	
39-42	4	Initial Temp. Holding Time in min.	XX.Y
43	1	blank	
44-45	2	Number of Column Temperature Programs	Integer number
46	1	blank	
47-48	2	First (or only) Column Temperature Program in degrees C/min	e.g., 8
49	1	blank	

Format of the Chromatography Record (Type 11) (cont.)

Record Position	Field Length	Field Contents	Remarks
50-52	3	First* Column Temp. in degrees C.	e.g., 250 or 350
53	1	blank	
54-57	4	First* Temp. Holding Time in min.	XX.Y

*Note: When Number of Column Temperature Programs is "1", positions 50-52 and 54-57 will hold the final column temperature and holding time, and no type 12 record will follow.

Format of the Chromatography Record (Type 12)

Use: Continuation of type 11. Used only if multiple ramp column temperature programs are employed.

Position: Follows the type 11 to which it applies.

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"12"
3	1	blank	
4-14	11	Second Column Temperature Program	Use the same format as positions 47-57 of record type 11.
15	1	blank	
16-26	11	Third Column Temperature Program	Use the same format as positions 47-57 of record type 11.
27	1	blank	
28-38	11	Fourth Column Temperature Program	Use the same format as positions 47-57 of record type 11.
39	1	blank	
40-50	11	Fifth Column Temperature Program	Use the same format as positions 47-57 of record type 11.
51	1	blank	
52-62	11	Sixth Column Temperature Program	Use the same format as positions 47-57 of record type 11.

Format of the Mass Spectrometer Record (Type 13)

Use: To describe Mass Spectrometer conditions. Applies to a group of samples in a run. Will be present whenever mass spectrometry is used.

Position: Follows type 10

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"13"
3	1	blank	
4-9	6	Instrument model	First letter - manufacturer,
10	1	blank	1-5 characters for model.
11-13	3	Scan cycle time in sec.	1.3.
14	1	blank	
15	1	Scan Type	R - Continuous Scan Range
16	1	blank	S - SIM - mass range given
			U - Unknown scan type
			N - SIM - # masses given
17-20	4	Initial Mass Value or Number of Masses	Integer mass value. Lowest mass for "R", "S", or "U"
21	1	blank	(above); or number of masses monitored for "N".
22-25	4	Final Mass Value	Highest mass for "R", "S",
26	1	blank	or "U"; or blank for "N".
27-29	3	Pos. or Neg. ions	"POS" or "NEG"
30	1	blank	
31	1	Type of Instrument	M - magnetic, Q - quadrupole
32	1	blank	Other types may be defined
33-38	6	Mass Spectrometer Resolution or Peak Width *	Integer resolution value
39	1	blank	
40-41	2	Ionization Mode	FA, FD, FI, EI, TS, CI, AP
42	1	blank	
43-49	7	Reagent Gas	Chemical symbol or formula e.g., He, CH4, C9H18, NH3

*(Defined as M/delta M for magnetics, Peak Width in amu for Quads)

Format of the AA/ICAP Instrument Record (Type 14)

Use: To describe AA/ICAP instrument conditions. Applies to a group of samples in a run. Will be present whenever AA/ICAP is used.

Position: Follows type 10

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"14"
3	1	blank	
4-9	6	Instrument model	First letter - manufacturer,
10	1	blank	1-5 characters for model
11-15	5	Initial Wavelength in nm	e.g., 5000
16	1	blank	
17-21	5	Final Wavelength in nm	
22	1	blank	
23-28	6	Gas utilized	Chemical symbol or formula,
29	1	blank	e.g., C2H2, NO
30-32	3	Flow rate magnitude	
33	1	blank	
34-35	2	Flow rate units	
36	1	blank	
37-39	3	Other gas added	e.g., AIR
40	1	blank	
41-43	3	Flow rate magnitude	Assumes same units as in
44	1	blank	Positions 34-35.
45-47	3	Digestion time	
48	1	blank	
49-51	3	Digestion temperature	
52	1	blank	
53-58	6	Acid used	e.g., H2SO4 or HNO3
59	1	blank	
60-64	5	Oxidizer used	e.g., H2O2
65	1	blank	
66-70	5	Other digestion options	Program may assign a code, e.g., SW846 digestion method.

-Format of the Mandatory Sample Header Data Record (Type 20)

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"20"
3	1	blank	
4-5	2	Region or other client	Alphanumeric
6	1	blank	
7-14	8	EPA Sample I.D.	Raw Sample ID only;
15	1	blank	no suffixes
16	1	Sample Medium/Matrix Code (Z)	See page B-9. Examples are
17	1	blank	found in Appendix C.
18-20	3	QC code	Codes type of data to be
21	1	blank	reported; see page B-2
22-24	3	Sample Qualifier	Code to qualify the results
25	1	blank	of the entire sample analysis (see page B-10).
26-33	8	Project number	e.g., Case # for Contract
34	1	blank	Laboratory Program.
35-37	3	Batch/shipment number	Alphanumeric
38	1	blank	
		Positions 39 through 52 contain the date/time of instrument analysis.	Field samples use date of sample collection.
39-40	2	Year	YY
41	1	blank	
42-43	2	Month	MM
44	1	blank	
45-46	2	Day	DD
47	1	blank	
48-49	2	Hour	HH
50	1	blank	
51-52	2	Minute	MM
53	1	blank	
54	1	Work shift of sample analysis	"G", "D" or "S" for:
55	1	blank	graveyard, day, swing.

Format of the Mandatory Sample Header Data Record (Type 20) cont.

Record Position	Field Length	Field Contents	Remarks
56	1	Sample Units Code	"L" = liters
57	1	blank	"C" = cubic meters
			"K" = kilograms (wet wt.)
58-65	8	Sample Size	See note.
66	1	blank	
67-69	3	Analyte count	Numeric; 1-3 decimal digits.

Note: Sample Size is the volume in liters for liquids, the volume in cubic meters for air and the wet weight in kilograms for solids. The Sample Units Code indicates which units are in use for the current sample.

Format of the Sample Header Data Record (Type 21)

Use: Continuation of type 20.

Position: Follows the type 20 to which it applies.

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"21"
3	1	blank	
4	1	Method Variation Code (N)	Codes any sample method
5	1	blank	variations. See Appendix C.
6	1	Concentration level	Indicates possible method
7	1	blank	variations. (See Note 1)
			"L" = low
			"M" = medium
			"H" = high
8-10	3	Clean-up or other sample processing variation	Codes to be used will be defined by each
11	1	blank	program.
12	1	Extraction code	As defined in contract.
13	1	blank	
14-16	3	Initials of operator	Use whomever is responsible
17	1	blank	for the sample results.
18-23	6	General Administrative Reporting Number	Alphanumeric; e.g., S.A.S. Number for CLP program
24	1	blank	(if necessary).
25-35	11	Laboratory Data File Name*	File name in instrument data
36	1	blank	system or other descriptor.
		Positions 37 through 44 contain the date/time that sample preparation began.	
37-38	2	Year	YY
39	1	blank	
			MM
40-41	2	Month	
42	1	blank	
43-44	2	Day	DD
45	1	blank	

Note 1: The Concentration level is an estimate of overall level for all analytes.

* The file name is the identifying code for sample data in a laboratory data system. In laboratories without data systems, the file name will be any code used for sample data identification.

Format of the Sample Header Data Record (Type 21) cont.

Record Position	Field Length	Field Contents	Remarks
46	1	Work shift for sample prep	"G", "D" or "S" for:
47	1	blank	graveyard, day, swing.
		Positions 48-55 contain date sample received at lab.	
48-49	2	Year	YY
50	1	blank	
51-52	2	Month	MM
53	1	blank	
54-55	2	Day	DD
56	1	Blank	
57-63	7	Source of Compound (if not unknown sample)	Company or EPA from which compound was obtained.
64	1	blank	
65-70	6	Volume of Sample Analyzed; Units determined by Contract	50 or 0.5; e.g., Injection Volume in uL for CLP.

Format of the Sample Conditions Record (Type 22)

Use: Continuation of type 20. Used to describe additional sample conditions.
Position: Follows the type 20 and 21 to which it applies.

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"22"
3	1	blank	
		Positions 4-17 contain the date/time of associated calibration. See Note 1. (Date of Source of the response factors used)	
4-5	2	Year	YY
6	1	blank	
7-8	2	Month	MM
9	1	blank	
10-11	2	Day	DD
12	1	blank	
13-14	2	Hours	HH
15	1	blank	
16-17	2	Minute	MM
18	1	blank	
19-29	11	Calibration data File Name*	See Note 2. Data File Name of associated calibration or "AVERAGE" in positions 21-27 (if mean used).
30	1	blank	
31-34	4	Sample pH	XX or XX.X
35	1	blank	
36-37	2	Percent moisture	For organic, dioxin
38	1	blank	
39-40	2	Decanted percent moisture	For organic dioxin
41	1	blank	

Note 1: For average, use the date and time average was calculated.
Note 2: This field must match positions 25-35 of record type 21 for the associated QC injection.

* The file name is the identifying code for sample data in a laboratory data system. In laboratories without data systems, the file name will be any code used for sample data identification.

Format of the Sample Conditions Record (Type 22) cont.

A-49

Record Position	Field Length	Field Contents	Remarks
42-46	5	Extract Volume in ml.	e.g. 1.0 or 0.050
47	1	blank	
48-54	7	Concentration/dilution factor	e.g., 2000 or .001
55	1	blank	
56-59	4	Method Detection Limit Method	Established per project by Project Officer.
60	1	blank	
61	1	Code for quantification report type	Program will specify when desired.
62	1	blank	
63-70	8	Sample Dry Weight or percent solids	If necessary, contract will define required value.

Format of the Associated Injection and Counter Record (Type 23)

Use: Continuation of type 20. Used to identify associated QC injections and to provide for program specific counters. May not be required in all programs.

Position: Follows the type 20, 21, and 22 to which it applies.

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"23"
3	1	blank	
4	1	Type of First QC Injection	Identifies injection type.
5	1	blank	"P" – performance check, "B" – blank, etc. Other codes may be defined.
		Positions 6 through 19 contain date/time of associated QC injection. (Acquisition date and time of QC injection to be linked with this sample.)	
6-7	2	Year	YY
8	1	blank	
9-10	2	Month	MM
11	1	blank	
12-13	2	Day	DD
14	1	blank	
15-16	2	Hour	HH
17	1	blank	
18-19	2	Minute	MM
20	1	blank	
21-31	11	QC injection File Name*	See Note 1.
32	1	blank	
33	1	Type of Second QC Injection	Identifies Second injection type; same as position 4.
34	1	blank	

Note 1: This field must match positions 25-35 of record type 21 for the associated QC injection.

* File name is the identifying code for sample data in a laboratory data system. In laboratories without data systems, the file name will be any code used for sample data identification.

Record Position	Field Length	Field Contents	Remarks
		Positions 35 through 49 contain the date/time of associated QC injection. Positions 33 to 60 have the same format as positions 4 to 31 for the second type of QC injection. If more than two types of injections must be linked with the sample then use additional records.	
35-36	2	Year	YY
37	1	blank	
38-39	2	Month	MM
40	1	blank	
41-42	2	Day	DD
43	1	blank	
44-45	2	Hour	HH
46	1	blank	
47-48	2	Minute	MM
49	1	blank	
50-60	11	QC Injection File Name*	
61	1	blank	
62	1	Description Code of First Counter	Program may define any necessary sample-wide coun-
63	1	blank	ters to be reported here.
64-65	2	First Counter	nn
66	1	blank	
67	1	Description Code of Second Counter	Two counters may be entered on each record.
68	1	blank	
69-70	2	Second Counter	

* The file name is the identifying code for sample data in a laboratory data system. In laboratories without data systems, the file name will be any code used for sample data identification.

Format of the Field Sampling Auxilliary Data Record (Type 24)

Use: Continuation of type 20. Used for Field Sampling Data to describe additional sample conditions.

Position: Follows the type 20 and 21 to which it applies.

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"24"
3	1	blank	
4-39	36	Project Name	
40	1	blank	
41-55	15	Sample Station Number	From Standard List.
56	1	blank	
57	1	Type of Sample	G grab; T time composite;
58	1	blank	S space composite
59-61	3	Preservative Added	From standard list - blank
62	1	blank	if none.
		Positions 63 through 70 contain the date the sample was shipped to the lab.	
63-64	2	Year	YY
65	1	blank	
66-67	2	Month	MM
68	1	blank	
69-70	2	Day	DD

Format of the Field Sampling Auxilliary Data Record (Type 25)

Use: Continuation of type 24. Used for Field Sampling Data for additional
      descriptive information - exact·format defined by each program.

Position: Follows the type 24 to which it applies.

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"25"
3	1	blank	
4-36	33	Station Name, Location, and/or Description	Field contents and formats may be defined further by
37	1	blank	the individual program.
38-70	33	Names of Samplers	As above. May also contain chain-of-custody data.

Format of the Field Sampling Auxilliary Data Record (Type 26)

Use: Continuation of type 24. Used for Field Sampling Data to record any numerical values which indicates where or how the sample was collected. The exact format is defined by each program.

Position: Follows the type 24 to which it applies. (Record will only be required for some programs.)

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"26"
3	1	blank	
4-6	3	Description of First Value	Program specified descriptor, e.g., "FLW" - flow rate; "TMP" - temperature; "LAT" - latitude; "LON" - longitude; "ALT" - altitude.
7	1	blank	
8-11	4	Magnitude of First Value	Fixed or Scientific notation (.XXXXEYYY). Program will define appropriate measurement and applicable units.
12	1	blank or 'E'	
13-15	3	Exponent	Blank field will be interpreted as "+000".
16	1	blank	
17-19	3	Description of Second Value	Each value has the same format as positions 4-15. Up to five values may be given on each record. Additional records may be added if necessary
20	1	blank	
21-24	4	Magnitude of Second Value	
25	1	blank or 'E'	
26-28	3	Exponent	
29	1	blank	
30-32	3	Description of Third Value	
33	1	blank	
34-37	4	Magnitude of Third Value	
38	1	blank or 'E'	
39-41	3	Exponent	
42	1	blank	

Format of the Field Sampling Auxilliary Data Record (Type 26) cont.

Record Position	Field Length	Field Contents	Remarks
43-45	3	Description of Fourth Value	
46	1	blank	
47-50	4	Magnitude of Fourth Value	
51	1	blank or 'E'	
52-54	3	Exponent	
55	1	blank	
56-58	3	Description of Fifth Value	
59	1	blank	
60-63	4	Magnitude of Fifth Value	
64	1	blank or 'E'	
65-67	3	Exponent	

Format of the Results Data Record (Type 30)

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"30"
3	1	blank	
4	1	Type of Identifier Used	"I" = General Identifier
5	1	blank	(e.g. chemical symbol, program code). "C" = CAS Registry Number. Other codes may be defined.
6-14	9	Identifier Code or CAS #	Identifier codes may only be
15	1	blank	used when no acceptable CAS # exists. (Use right justification in either case.
16-24	9	Identifier Code or CAS # of internal standard utilized.	For internal standard, if measurement uses internal
25	1	blank	standards; otherwise leave blank.
26-30	5	Units of measure	Established per project by
31	1	blank	Project Officer.
32-34	3	Non-numeric result	See page B-10; also called
35	1	blank	a result qualifier.
36-41	6	Numeric analytical result	Fixed point or scientific
42	1	blank or 'E'	notation.
43-45	3	Exponent	
46	1	blank	
47	1	Calculated Value Descriptor	Describes following value:
48	1	blank	"S" - surrogate; "F" - spiked analyte; "N" - # of points in mean. Other codes may be defined.
49-54	6	Related Calculated Value	Value represents amount
55	1	blank or 'E'	added or other calculated or theoretical value. Format same as 36-46.
56-58	3	Exponent	
59	1	blank	
60	1	QC or Limit Value Descriptor	Describes following value:
61	1	blank	"D" - method detection limit "S" - surrogate % recovery.
62-66	5	Related QC or Limit Value	Value is method detection
67	1	blank or 'E'	limit; surrogate % recovery or other type defined by the appropriate program.
68-70	3	Exponent	

Format of the Instrumental Data Readout Record (Type 31)

Use: To describe a specific instrument readout value (raw data), for a specific
sample where both the instrument setting and the associated value must be
reported; exact nature of the value will be program dependent.

Position: Follows type 30. (Record will only be required for some programs.)

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"31"
3	1	blank	
4	1	Type of Data	Code for description of property
5	1	blank	being measured or Instrument Setting; e.g., M - mass (GC/MS), or W - wavelength in nm.
6	1	Type of Value Recorded	Code for Value Recorded;
7	1	blank	e.g., A - area, B - absorbance, H - height, P - percent abundance, I - intensity.
8	1	Method of Data Entry	C - computer, M - manual
9	1	blank	(could be a sequence #).
10-17	8	First Instrument Setting	e.g., 320 or 320.0736 for
18	1	blank	mass, or 4973.61 for wavelength.
19-28	10	First Instrument Value	Up to 10 decimal digits.
29	1	blank	
30-37	8	Second Instrument Setting	Up to three readouts may be
38	1	blank	given on each record provided that positions 4, 6 and 8 are
39-48	10	Second Instrument Value	the same for all.
49	1	blank	
50-57	8	Third Instrument Setting	
58	1	blank	
59-68	10	Third Instrument Value	

Format of the Auxilliary Data Record (Type 32)

Use: To describe qualifying data for calibration or analytes in samples. Indicates where in the analysis data are located or how data were found or measured CLP program will report analyte scan number and retention time (in minutes). Other projects may use this record for any numerical sample qualifying data.

Position: Follows type 30. (Record will only be required for some programs.)

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"32"
3	1	blank	
4-5	2	Match Score (if matching of any sort was utilized)	0 to 99%; e.g., used for tentatively identified
6	1	blank	compounds in GC/MS
7	1	Match Score Specifier	How score was obtained; one
8	1	blank	alphabetic char; program will specify code when it is appropriate.
9-10	2	Description Code of First Value	Program specified descriptor, e.g., "RT" for GC/MS
11	1	blank	retention time; "IT" for integration time; "QM" for quantitation mass.
12-17	6	Magnitude of First Value	Fixed or Scientific notation
18	1	blank or 'E'	as in Record Type 30. Program will define appropriate measurement and applicable units.
19-21	3	Exponent	
23-24	2	Description Code of Second Value	Each value has the same format as positions 9-21. Up to four
25	1	blank	values may be given on each record. Additional records may be added if necessary.
26-31	6	Magnitude of Second Value	
32	1	blank or 'E'	
33-35	3	Exponent	
36	1	blank	

Format of the Auxillary Data Record (Type 32) cont.

Record Position	Field Length	Field Contents	Remarks
37-38	2	Description Code of Third Value	
39	1	blank	
40-45	6	Magnitude of Third Value	
46	1	blank or 'E'	
47-49	3	Exponent	
50	1	blank	
51-52	2	Description Code of Fourth Value	
53	1	blank	
54-59	6	Magnitude of Fourth Value	
61	1	blank or 'E'	
61-63	3	Exponent	

Format of the Name Record (Type 33)

Use: To carry an analyte name and any other necessary identifying information
Different programs may define further information to be reported.

Position: Follows type 30. (Record will only be required for some programs.)

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record type	"33"
3	1	blank	
4-70	67	Name of compound	Different programs may define this field further.

Format of the QC Limit Record (Type 34)

Use: To report QC limit values that were in effect for the indicated measurement,
or for the entire production run, depending on program requirements.

Position: Follows type 30. (Record will only be required for some programs.)

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record Type	"34"
3	1	blank	
4-7	4	Type of data present	QC chart type, or any other
8	1	blank	descriptor. See page B-8.
9-11	3	Type of Value(s) present	Limit Type (MIN, MAX, A, B,
12	1	blank	LWL, LCL, AVE, UCL, UWL), or other descriptor. See page B-8.
13	1	Method for calculating limit	M = manual, C = computer Other codes may be defined.
14	1	blank	
15-22	8	Instrument Setting	Only if appropriate; (e.g.
23	1	blank	wavelength value).
24-29	6	First QC or Limit Value	May be a mean. Use fixed or
30	1	blank or 'E'	scientific notation.
31-33	3	Exponent	
34	1	blank	
35-40	6	Second QC or Limit Value	May not be necessary. Use
41	1	blank or 'E'	format of positions 24-29.
42-44	3	Exponent	
45	1	blank	
46-51	6	Standard Deviation	
52	1	blank or 'E'	
53-55	3	Exponent	
56	1	blank	
57-59	3	Number of points used for mean	Integer.
60	1	blank	

Format of the QC Limit Record (Type 34) cont.

Record Position	Field Length	Field Contents	Remarks
		Positions 61 through 68 contain the date the QC limits were computed.	
61-62	2	Year	YY
63	1	blank	
64-65	2	Month	MM
66	1	blank	
67-68	2	Day	DD

Format of the Correction Data Record (Type 35)

Use: To record any correction data required. Different programs may define
      further information to be reported.

Position: Follows type 30. (Record may be required only for some programs.)

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record Type	"35"
3	1	blank	
4-6	3	Type of Correction	"ICP" for ICP interelement
7	1	blank	correction factors.
9-12	5	Type of Value or Units	If necessary - describes
13	1	blank	factor or gives units.
		Positions 14 through 22 contain the date the factor was determined.	
14-15	2	Year	YY
16	1	blank	
			MM
17-18	2	Month	
19	1	blank	
			DD
20-21	2	Day	
22	1	blank	
23-31	9	CAS # of interfering element	
32	1	blank	
33-40	8	Instrument setting in nm	Wavelength for ICP
41	1	blank	
42-47	6	Correction factor	Use fixed or scientific notation.
48	1	blank or 'E'	
49-51	3	Exponent	

Contents of rest of record may be defined further by other programs.

Format of the Deleted Data Record (Type 40)

Use: To delete any record.

Position: May occur anywhere.

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record Type	"40"
3	1	blank	
4-70	67	Contents undefined	

Note: Any record type may be logically deleted by changing Record Type field to "40". Remaining contents of record are unchanged and should be ignored by all processing software.

Format of the Special Data Record (Type 50)

Use: To indicate the presence of any data records from other Agency Data Base
       Systems (e.g., STORET, AIRS, etc). This record may be used to report data in
       any other format without having to convert the data.

Position: May occur anywhere.

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record Type	"50"
3	1	blank	
4-12	9	EPA (or other agency) Project Type	e.g., STORET, SAROAD, AIRS, SFC
13	1	blank	
14-18	5	Counter	Indicates the number of records from the indicated system that will follow.
19	1	blank	
20-70	51	Comments	Any free-form comments may appear here.

This record is necessary only if records from another system are being mixed with
records from this standard. The record may appear multiple times if data from more
than one additional system are present, or if all such data are not contiguous. The
counter will give the number of records in the alternate format that follow. These
alternative records have no defined format within this standard, and therefore no
check for any contents will be made. Record types, sequence numbers, and checksums
will not be present in the expected fields, and the sequence number counter will
simply ignore these records. Processing programs are expected simply to pass these
records as received to the appropriate system.

Format of the Comment Record (Type 90)

Use: To provide any other necessary comments. Different programs may define this further and may require its presence in various places.

Position: May occur anywhere (see above).

Record Position	Field Length	Field Contents	Remarks
1-2	2	Record Type	"90"
3	1	blank	
4-70	67	Any Comment	Any program may use this record for any purpose and may further define field contents.

Appendix B

---- Definitions of Various Codes

STRUCTURE OF THE METHOD NUMBER

The Method Number

The method number is a five character alphanumeric code. The purpose of the method number is to define concisely the target analytes and the details of the method of analysis. The method number has the form:

XXXXY

Where:

XXXX   defines one or more target analytes plus the analytical method. This part of the code is identical with the method numbers defined in EPA methods manuals, the code of Federal Regulations and the private standard setting organizations, e.g., ASTM.

Y   is an alphanumeric modifier which specifies that an allowed option in the method has been implemented or specifies fractions of analytes in the method. The defined values of Y are dependent on the value of XXXX, that is, a Y = 5 in the 200 series methods may have a different meaning than Y = 5 in the 300 series methods. As an example, Y may distinguish total and dissolved phosphorus measured by the same method but with or without the optional method filtration. Another example is the use of Y to distinguish the acid and base/neutral fractions in method 625. If Y is not defined in a method, the default value is one.

The method number is validated as alphanumeric for XXXX and Y. It is stored right justified in the 5 digit method number field. Appendix C gives examples for organic and inorganic analyses.

Quality Control and Related Codes (QCC) in Type 20 Records

Note: These QCC appear in the QC code fields of type 20 records. They are used
        to indicate the type of data that are being reported. See page A-12.

QCC	Name	Definition
LD1	LABORATORY DUPLICATE FIRST MEMBER	The first of two aliquots of the same environmental sample. Each aliquot is treated identically throughout a laboratory analytical procedure; and each is carried through the entire laboratory analytical method as applied to all other samples analyzed with the same method.
LD2	LABORATORY DUPLICATE SECOND MEMBER	The second of the two aliquots described under LD1.
LD3 to LD9	LABORATORY REPLICATE Nth MEMBER	The 3rd through the 9th additional aliquots which logically follow LD1 and LD2. If more than two aliquots are used, all names are changed from duplicates to replicates. Codes do not change.
LRB	LABORATORY (REAGENT) BLANK	An aliquot of reagent water or equivalent neutral reference material treated as an environmental sample in all aspects in the laboratory including addition of all reagents, internal standards, surrogates, glassware, apparatus, equipment, solvents, and analyses.
LDB	LABORATORY (DRY) BLANK	Exactly the same as the LRB except the aliquot of reagent water or equivalent neutral reference material is omitted.
LSB	LABORATORY (SOLVENT) BLANK	Exactly the same as the LDB except any internal standards or surrogates are omitted.
LCB	LABORATORY CALIBRATION BLANK	An aliquot of reagent water, possibly adjusted in pH, but without addition of other reagents.
LCM	LABORATORY CONTROL SOLUTION	An aliquot of reagent water or equivalent neutral reference material to which a known quantity(s) of method analyte(s) was added in the laboratory. The LCM is treated as an environmental sample in all aspects in the laboratory including addition of all reagents, internal standards, surrogates, glassware, equipment, solvents, and analyses.
LVM	LABORATORY CALIBRATION VERIFICATION SOLUTION	Exactly like LCM; used for calibration verification.

QCC	Name	Definition
LIM	LABORATORY INTERFERENCE CHECK SOLUTION	Exactly like LCM; used to verify inter-element and background correction factors.
LFM	LABORATORY FORTIFIED BLANK	An aliquot of sample matrix, known to be below detection limits for an analyte(s), to which a known quantity(s) of method analyte(s) was added. The LFM is treated as an environmental sample in all aspects in the laboratory including addition of all reagents, internal standards, surrogates, glassware, equipment, solvents, and analyses.

LSO	LABORATORY SPIKED SAMPLE BACKGROUND (ORIGINAL) VALUES	An environmental sample which is analyzed according to the analytical method, and a <u>single</u> independent aliquot of the same sample is taken for fortification (spiking) with the method analyte(s).
LSF	LABORATORY SPIKED SAMPLE - FINAL VALUES	An environmental sample in which the analyte(s) was measured in an independent sample aliquot before spiking (LSO), a known concentration increment was made, and the measurement(s) of the final concentration(s) were made according to the analytical method (LSF).

LDO	LABORATORY DILUTED SAMPLE BACKGROUND (ORIGINAL) VALUES	An environmental sample which is analyzed according to the analytical method, and a single independent aliquot of the same sample is taken and diluted according to the analytical method.
LDF	LABORATORY DILUTED SAMPLE - FINAL VALUES	An environmental sample in which the analyte(s) were measured in an independent sample aliquot before dilution (LDO), a known dilution was made, and the measurement(s) of the final concentration(s) were made according to the analytical method (LDF).

LSD	LABORATORY SPIKE DUPLICATE	An environmental sample exactly like the LSO except that <u>two</u> independent aliquots of the same sample are taken for fortification (spiking) with the method analyte(s).
LF1	LABORATORY SPIKED SAMPLE - FINAL - FIRST MEMBER	An environmental sample exactly like the LSF except that duplicate aliquots were spiked, and the measurement(s) of the final concentration was made according to the analytical method (LF1).
LF2	LABORATORY SPIKED SAMPLE - FINAL - SECOND MEMBER	The second member of the LF1/LF2 duplicate pair.

QCC	Name	Definition
LPS	LABORATORY CONFIRMATORY SCAN	The measurement of the spectrum or partial spectrum of an analyte(s) in an environmental sample or extract to obtain additional qualitative evidence when the analyte(s) identification and measurement were obtained from other techniques.
LPC	LABORATORY PERFORMANCE CHECK SOLUTION	A solution of method analyte(s), surrogate(s) and/or internal standard(s) used to evaluate the performance of an instrument with respect to a defined set of criteria.
LDX	LABORATORY DOUBLE PURPOSE PRECISION AND ACCURACY SAMPLE	An environmental sample which is used for both the LSO (background level before spike) and LD1 (first member of a duplicate).
CAL	CONCENTRATION CALI-BRATION SOLUTION (Type Unspecified)	A solution of method analyte(s) used to calibrate the instrument response in terms of concentration of analyte(s). Response factors rather than concentrations will be reported on the following type 30 records.
CLM	INITIAL CALIBRATION - MULTI POINT	A calibration solution as above used to determine the initial calibration of an entire production run where a group of calibrations are required at different levels of method analyte concentrations.
CLS	INITIAL CALIBRATION - SINGLE POINT	Exactly the same as CLM except only a single level of method analyte concentrations are utilized.
CLC	CONTINUING CHECK CALIBRATION	A calibration solution as above used to verify whether the initial calibration data are still currently valid. Will be run several times throughout the duration of the production run.
CLD	DUAL PURPOSE CALIBRATION	A calibration solution as above used both as an initial calibration (CLM or CLS) and a continuing check (CLC).
IDL	INSTRUMENT DETECTION LIMIT SOLUTION	A calibration solution (not necessarily the same solution as above), where the data are to be used to calculate instrument detection limits only.
blank		Unknown sample, not associated with any quality control item.

The following QCC will only apply to field data.

QCC	Name	Definition
FD1	FIELD DUPLICATE - FIRST MEMBER	The first of two environmental samples taken at the same time and place under identical circumstances. Each sample is treated identically throughout field and laboratory analytical procedures; and each is carried through the entire laboratory analytical method as applied to all other samples analyzed with the same method.
FD2	FIELD DUPLICATE - SECOND MEMBER	The second of the two samples described under FD1.
FRB	FIELD BLANK	An aliquot of reagent water or equivalent neutral reference material treated as an environmental sample in all aspects in both the field and the laboratory including addition of all preservatives, reagents, internal standards, surrogates, glassware, apparatus, equipment, solvents and analyses.
FCM	FIELD CONTROL SOLUTION	An aliquot of reagent water or equivalent neutral reference material to which a known quantity(s) of method analyte(s) was added in the field. The FCM is treated as an environmental sample in all aspects in both the field and the laboratory, including addition of all preservatives, reagents, internal standards, surrogates, glassware, equipment, solvents and analyses.
FRM	FIELD REFERENCE SOLUTION	An aliquot of a sample (submitted by the requestor) having a certified value. These samples are usually obtained from the NBS, EMSL, etc. The concentration measured by the same analytical procedure used for other samples is the "found" value.
FFM	FIELD FORTIFIED BLANK	An aliquot of sample matrix, known to be below detection limits for an analyte(s), to which a known quantity(s) of method analyte(s) was added in the field. The FFM is treated as an environmental sample in all aspects in the field and in the laboratory, including addition of all preservatives, reagents, internal standards, surrogates, glassware, equipment, solvents and analyses.

QCC	Name	Definition
FSO	FIELD SPIKED SAMPLE BACKGROUND (ORIGINAL) VALUES	An environmental sample which was split in the field. The portion represented by FSO is analyzed according to the analytical method without fortification (spiking).
FSF	FIELD SPIKED SAMPLE - FINAL VALUES	The second portion of the environmental sample which was split in the field, and to which a spike was added in the field with a known concentration increment. The measurement(s) of the final concentration(s) was made according to the analytical method (FSF).

The following QCC values do not refer to actual samples or calibrations for which laboratory results are obtained. Instead they are used on type 20 records which act as a header and indicate that additional (usually calculated) analyte specific data will be present on type 30 (and following type) records. Usually these data will apply to an entire production run, in which case they will appear immediately following the type 10 record. If the data apply to only a portion of the samples in the run, they should be placed immediately preceding the samples to which they applies. Much of the rest of the information in the type 20 record may be blank, indicating that this data does not apply to these results. Many of these codes are method specific, and more codes will be added as additional methods require additional data.

MNC	MEAN VALUES FROM CALIBRATIONS	The data following represent mean values and percent RSD's from several calibration solutions. Data will be present for each method analyte for which a mean has been determined.
SID	SAMPLE INDEPENDENT (i.e. INSTRUMENT) DETECTION LIMITS	The data following represent sample independent detection limits for each method analyte calculated according to the method being utilized.
ICF	INTER-ELEMENT CORRECTION FACTORS	The data following represent ICP interelement correction factor measurements for each method analyte.
SDR	SPIKE/DUPLICATE CALCULATED RESULTS	The data following represent calculated QC results for any QC samples involving multiple injections. Data will consist of percent recoveries and the percent RSD values for each appropriate method analyte that was analyzed according to the analytical method.

Laboratory Quality Control Codes Which Do Not Involve Real Samples

QC Code	Name	I/O	Internal Stds	Surrogates	Clean Matrix Present	Clean Matrix Analyzed	Clean Matrix Spiked
LRB	Lab Reagent Blank	O	yes	yes	yes	yes	no
LRB	Lab Reagent Blank	I	no	no	yes	yes	no
LDB	Lab Dry Blank	O	yes	yes	no	N/A	N/A
LSB	Lab Solvent Blank	O	no	no	no	N/A	N/A
LCB	Lab Calibration Blank	I	no	no	yes	yes	no
LFM	Lab Fortified Blank-Measured	O	yes	yes	yes	no	yes
LCM	Lab Control Solution Measured	I	no	no	yes	no	yes

NOTE 1: All except LCB prepared in the laboratory and treated exactly like a sample for the value being measured, including all preanalysis treatments.

NOTE 2: Entries in I/O column: I = inorganic, O = organic

NOTE 3: LVM and LIM differ from LCM in their QC role in the run.

Quality Control Codes in Type 34 Records

Note: Type 34 records are used to record the limit values which were in force during the run. Other programs may define other codes. All codes should be left justified.

Record Field	Code	Meaning
QC Chart Type	LSPK	Statistical data from Laboratory Spikes
	LSSP	Statistical data from Laboratory Surrogate Spikes
	LRBL	Statistical data from Laboratory Reagent Blanks
	LDUP	Statistical data from Laboratory Duplicates
	LCST	Statistical data from Laboratory Control Standards
Limit Type	A	Critical Range (R_C) Slope
	B	Critical Range (R_C) Intercept

Note: Upper limit for duplicates is expressed by the critical range linear equation: $R_C = AX + B$

	MIN	Minimum concentration for which duplicates limit is applicable
	MAX	Maximum concentration for which duplicates limit is applicable
	LCL	Lower control limit
	LWL	Lower warning limit
	AVE	Mean
	UWL	Upper warning limit
	UCL	Upper control limit

Note: LCL, LWL, AVE, UWL and UCL apply to all QC chart types except LDUP.

Codes For Sample Medium (Matrix, Source)

Medium	Code
All Media, Don't Know or Don't Care	0
Water, Type Unknown or Not Specified	1
Drinking Water	2
Ambient Surface Fresh Water	3
Raw Wastewater	4
Primary Effluent Wastewater	5
Effluent Wastewater (Secondary - Tertiary)	6
Industrial Wastewater	7
Salt, Ocean or Brackish Water	8
Ground Water	9
Leachate	A
Air, Type Unknown or Not Specified	B
Ambient Air	C
Source or Effluent Air	D
Industrial Workroom Air	E
Solids, Type Unknown or Not Specified	F
Bottom Sediment or Deposit	G
Soil	H
Sludge	I
Hazardous Wastes, Dumps	J
Fish, Shellfish Tissue	K
Plants, Algae Tissue	L
Commercial Product Formulation	M
Gasoline	N
Waste Oils	P
Field Sampling Equipment Solvent Washings	Q
Atmospheric Deposition (Direct only)	R

LIST OF SAMPLE and RESULT QUALIFIERS

Definition: A sample qualifier or a result qualifier (also called a non-numeric result) consists of 3 alphanumeric characters which act as an indicator of the fact and the reason that the subject analysis (a) did not produce a numerical result, (b) produced a numeric result but it is qualified in some respect relating to the type or validity of the result or (c) produced a numeric result but for administrative reasons is not to be reported outside the laboratory. Qualifiers related to STORET remarks are indicated in the list below. This list is not intended to be complete, and it is assumed that individual projects will add additional qualifiers to cover project specific circumstances.

Qualifier	Full Name	Definition
BDL	BELOW DETECTABLE LIMITS	There was not a sufficient concentration of the parameter in the sample to exceed the lower detection limit in force at the time the analysis was performed. (No result; STORET "W" remark) Numeric results field, if present, is at best, an approximate value.
FPS	FAILED PRELIMINARY SCREENING	A preliminary screening of the sample for the subject parameter was conducted. The result of the screening indicated that it would not be useful to determine the concentration of the parameter. (No result; no STORET remark)
NSQ	NOT SUFFICIENT QUANTITY	There was not a sufficient quantity of the sample to conduct an analysis to determine the concentration of the subject parameter. (No result; no STORET remark)
LAC	LABORATORY ACCIDENT	There was an accident in the laboratory that either destroyed the sample or rendered it not suitable for analysis. (No result; STORET "O" remark)
FAC	FIELD ACCIDENT	There was an accident in the field that either destroyed the sample or rendered it not suitable for analysis. (No results; no STORET remark)
ISP	IMPROPER SAMPLE PRESERVATION	Due to improper preservation of the sample, it was rendered not suitable for analysis. (No results; no STORET remark code)
PNQ	PRESENT BUT NOT QUANTIFIED	The subject parameter was present in the sample but no quantifiable result could be determined. (No result; STORET "M" remark)
CMP	USED AS PART OF A COMPOSITE	The sample was not analyzed for the subject parameter, instead it was used as part of a composite sample. (No result; STORET "E" remark)

Qualifier	Full Name	Definition
NAI	NOT ANALYZED DUE TO INTERFERENCE	Because of uncontrollable interference the analysis for the subject parameter was not conducted. (No result; no STORET remark)
NAR	NO ANALYSIS RESULT	There is no analysis result required for this subject parameter. (No result; no STORET remark)
PRE	PRESUMPTIVE PRESENCE	Presumptive evidence of presence of material; tentative identification (No result; STORET "N" remark)
UND	ANALYZED BUT UNDETECTED	Indicates material was analyzed for but not detected. (No result; STORET "U" remark)
FQC	FAILED QUALITY CONTROL	The analysis result is not reliable because quality control criteria were exceeded when the analysis was conducted. Numeric field, if present, is estimated value. (Result; no STORET remark, non-reportable; or report with STORET "J" remark)
RNA	RELEASE/REPORT NOT AUTHORIZED	The analysis result is not authorized (by laboratory management) for either forwarding to a National Database or presentation in Engineering tabulations (No STORET remark)
AVG	AVERAGE VALUE	Average value - used to report a range of values (STORET "A" remark)
CNT	NON-ACCEPTABLE COLONY COUNTS	Results based on colony counts outside the acceptable range. (STORET "B" remark)
CAL	CALCULATED RESULT	Calculated result. (STORET "C" remark)
FLD	FIELD MEASUREMENT	Field measurement. (STORET "D" remark)
FEM	FEMALE SEX	In the case of species, indicates female sex. (STORET "F" remark)
KIT	FIELD KIT DETERMINATION	Value based on field kit determination - results may not be accurate. (STORET "H" remark)
EST	ESTIMATED VALUE	Present above detection limit but not quantified within expected limits of precision. (STORET "J" remark)
CAN	CANCELLED	The analysis of this parameter was cancelled and not performed. (No result; no STORET remark)
MAL	MALE SEX	In the case of species, indicates male sex. (STORET "M" remark)

Qualifier	Full Name	Definition
LTL	LESS THAN LOWER DETECTION LIMIT	Actual value is known to be less than value given - lower detection limit. (STORET "K" remark)
GTL	GREATER THAN UPPER DETECTION LIMIT	Actual value is known to be greater than value given - upper detection limit. (STORET "L" remark)
LTC	LESS THAN CRITERIA OF DETECTION	Value reported is less than the criteria of detection (which may differ from instrument detection limits). (STORET "T" remark)
UNK	UNDETERMINED SEX	In the case of species, indicates undetermined sex. (STORET "U" remark)
RET	RETURN(ED) FOR RE-ANALYSIS	The analysis result is not approved by laboratory management and reanalysis is required by the bench analyst with no change in the method. (No STORET remark)
EER	ENTRY ERROR	The recorded value is known to be incorrect but a correct value cannot be determined to enter a correction. (No STORET remark)
REQ	REQUEUE FOR REANALYSIS	The analysis is not approved and must be re-analyzed using a different method. (No STORET remark)
CBC	CANNOT BE CALCULATED	The calculated analysis result cannot be calculated because an operand value is qualified.
LLS	LESS THAN LOWER STANDARD	The analysis value is less than the lower quality control standard. (Result; STORET "J" remark)
MPR	MIDPOINT OF RANGE	The analysis value is the midpoint value of a range of concentrations.
MSL	"EMSL" DETECTION LIMITS	Instrument Detection Limits were computed using a "T" test on two or more calibration samples.
TIE	TENTATIVELY IDENTIFIED - ESTIMATED VALUE	The subject parameter was not in the contract-defined list of parameters to be analyzed for; however its value has been estimated. (No STORET remark)
RIN	RE-ANALYZED	The indicated analysis results were generated from a re-analysis (injection) of the same sample extract or aliquot.
REX	RE-PREPARED	The indicated analysis results were generated from a re-preparation (extraction) of the same sample.

A-78

Qualifier	Full Name	Definition
REJ	REJECTED	The analysis results have been rejected for an unspecified reason by the laboratory. For any results where a mean is being determined, this data was not utilized in the calculation of the mean.
SPL	SPLIT RESULTS	The indicated environmental sample or calibration has been split into more than one analysis, and the analysis results will be reported as more than one group of results (multiple type 20 records).
SRN	SPLIT RESULTS – RE-ANALYZED	A combination of "SPL" and "RIN"
SRX	SPLIT RESULTS – RE-PREPARED	A combination of "SPL" and "REX"
STD	INTERNAL STANDARD	The subject parameter is being utilized as an internal standard for other subject parameters in the sample. There is no analysis result to report, although the theoretical and/or limit value(s) may be present.
STB	INTERNAL STANDARD BELOW DETECTION LIMITS	A combination of "STD" and "BDL"
BAC	BACKGROUND CORRECTION	Background correction has been applied to this value.
FBK	FOUND IN BLANK	The subject parameter had a measurable value above the established QC limit when a blank was analyzed using the same equipment and analytical method. Therefore the reported value may be erroneous.
CON	CONFIRMED	The subject parameter has been confirmed using an auxilliary analytical technique as specified in the analytical method.
TFB	TENTATIVELY IDENTIFIED AND FOUND IN BLANK	A combination of "TIE" and "FBK"
ALC	ALDOL CONDENSATION	The indicated compound is suspected by the analyst of being a product of an aldol condensation reaction.
ALT	ALTERNATE MEASUREMENT	The subject parameter was determined using an alternate measurement method. Value is believed to be accurate but could be suspect.
AFB	ALTERNATE AND FOUND IN BLANK	A combination of "ALT" and "FBK"

Appendix C

Example Method and Matrix Codes for Dioxin,
General Organic and Inorganic Methods

The codes in Tables 1 and 2 are examples of method number designations for dioxin, general organics and inorganics. In all of these, the Z position refers to the matrix code and should be interpreted with the aid of page B-9. The generic value of 1, which represents "water, type unknown or not specified", is used for water analysis. Each of these generic matrix codes represents a group of specific codes, with Z values of 2 through 9, and A through R.

Solid samples are represented by two specific codes, with Z values of G (bottom sediment or deposit) and H (soil).

Dioxin rinsate samples use the value of Q (field sampling equipment solvent washings).

Each method code shown occurs in a type 10 record and acts as the header for the appropriate list of method analytes.

Method variations are designated by the N position. For example, Method 613 (Table 1) is run in three variations - full scan, partial scan and high resolution.

Table 1

Examples of Method and Matrix Codes for Dioxin and General Organics

XXXX	Y	N	Z	Definition
680	1	1	1	Pesticides and PCBs - water
680	1	1	G	Pesticides and PCBs - sediment
680	1	1	H	Pesticides and PCBs - soil
613	1	1	1	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - water
613	1	1	G	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - sediment
613	1	1	H	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - soil
613	1	1	Q	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - rinsate
613	1	2	1	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - water partial scan
613	1	2	G	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - sediment partial scan
613	1	2	H	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - soil partial scan
613	1	2	Q	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - rinsate partial scan
613	1	3	1	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - water - high resolution scan
613	1	3	G	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - sediment high resolution scan
613	1	3	H	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - soil - high resolution scan
613	1	3	Q	2,3,7,8-Tetrachloro-dibenzo-p-dioxin - rinsate - high resolution scan
624	1	1	1	GC/MS - Purgeables - water, internal/external standard
624	1	1	G	GC/MS - Purgeables - sediment
624	1	1	H	GC/MS - Purgeables - soil
624	1	2	1	GC/MS - Purgeables - isotope dilution - water
625	A	1	1	GC/MS - Acid Fraction - water, internal/external standard
625	B	1	1	GC/MS - Base/Neutral Fraction - water, internal/ external standard
625	C	1	1	GC/MS - combined acid and base/neutral fractions water, internal/external standard
625	C	1	G	GC/MS - combined fractions - sediment
625	C	1	H	GC/MS - combined fractions - soil
625	A	2	1	GC/MS - Acid Fraction - water, isotope dilution
625	B	2	1	GC/MS - Base/Neutral Fraction - water, isotope dilution

Notes:

1. See Page B-1 for the structure of "XXXX" and "Y".

2. For each water sample, the appropriate value of "Z" should replace the generic value of 1.

3. The values of "Z" and "N" are sample dependent and may vary within a production run. They are reported on type 20 and 21 records (pages A-12 and A-14).

Table 2

Examples of Method and Matrix Codes for Inorganics

XXXX	Y	N	Z	Definition
200	4	1	1	Generic code for analysis of total metals in water, after method – defined digestion, by atomic absorption, flame.
200	4	2	1	Generic code for analysis of total metals in water, after method – defined digestion, by atomic absorption, furnace.
200	4	7	1	Generic code for analysis of total metals in water, after method – defined digestion, by ICP.

Note 1: For specific matrix codes, replace Z with specific value for type of sample (from page B-9), and XXXX with value for metal:

202	Aluminum
204	Antimony
206	Arsenic
208	Barium
210	Beryllium
213	Cadmium
215	Calcium
218	Chromium
219	Cobalt
220	Copper
236	Iron
239	Lead
242	Magnesium
243	Manganese
249	Nickel
258	Potassium
270	Selenium
272	Silver
273	Sodium
279	Thallium
282	Tin
286	Vanadium
289	Zinc

Note 2: Exception to above:

245	1	1	1	Analysis of mercury in water by the manual cold vapor technique.
245	1	2	1	Analysis of mercury in water by the automated cold vapor technique.
245	4	5	G	Analysis of mercury in sediment, after method-defined digestion, by the manual cold vapor technique.

XXXX	Y	N	Z	Definition
245	4	5	H	Analysis of mercury in soil, after method-defined digestion, by the manual cold vapor technique.
245	4	5	I	Analysis of mercury in sludge, after method-defined digestion, by the manual cold vapor technique.
335	1	2	1	Analysis of total cyanide in water by titrimetric, manual spectrophotometric, or semi-automated spectrophotometric means.
335	1	2	G	Analysis of total cyanide in sediment by titrimetric, manual spectrophotometric, or semi-automated spectrophotometric means.
335	1	2	H	Analysis of total cyanide in soil by titrimetric, manual spectrophotometric, or semi-automated spectrophotometric means.

Table 3

Example of the Sequence of Record Types in a Production Run

10   Contains the Run Header information
   11   Contains Additional Run-Wide Information as Required.
   12
   13

   20   Occurs once for each sample, calibration, mean response factors,
               instrument detection limits, etc. - Acts as a header.
   21   Will usually be present
      22   Contains additional information for samples.

      30   Occurs once for each final analytical result. Will give
                  whatever value is being determined as defined by the type 20.
         31   Reports any instrumental data necessary.
         32   Reports any auxilliary data necessary.
         33   Reports component names if necessary.
         34   Reports QC Limit information if necessary.
         35   Reports Corrections to results if necessary.

      30   Values for the next analyte or parameter being measured.
         31   Additional data may vary for each parameter, and records
         32   may occur in any order. Multiple occurrences of the
         32   same record type, however, must be consecutive.
         33
      30   Continues for as many as are necessary.
         31
         32
         33
      30
         31
         32
         33

   20   Next Sample Header record - the following applies to the next
   21   sample or other group of data.
      22
      30
         31
         32
         33
      30
         31
         32
         33
                  etc.

   20
   21
      30
         31
         32
         33
                  etc.

⊕EPA

ORDER

Classification No.: 2180.3

Approval Date: 4/9/90

FACILITY IDENTIFICATION DATA STANDARD

1. **PURPOSE**. This Order establishes a data standard for unique facility identification codes to be maintained in all EPA data collections containing information on facilities regulated by EPA under authority of Federal environmental legislation. Standardization of the format and content of facility identification codes will enhance data integration capabilities and increase the utility of all EPA data on facilities.

2. **SCOPE AND APPLICABILITY**. The requirements of this Order apply to all programs responsible for data on regulated facilities reported to EPA and kept in automated or manual information collections developed for programmatic, research, or administrative purposes. The Order applies to programs operating both existing or future Agency systems in support of Federal environmental regulations.

   The principles of the standard can be extended to cover nonregulated facilities at program office discretion. Excluded from the standard are data monitoring or observation points, unless they are associated with a facility. In that case, the information collection must allow association of such data with the relevant facility.

3. **REFERENCE**. Chapter 5 of the EPA IRM Policy Manual sets forth the general principles on data standards within the Agency. This Order defines one of several data standards for use by EPA in implementing the policy.

4. **BACKGROUND**.

   a. Adoption of a consistent, Agency-wide coding scheme for facility identification will enhance the utility of EPA data by increasing access to and integration of facility information.

b.   Since the early 1980s, EPA has striven to create a
      standard ID coding scheme for facilities that can be
      used not only in individual programs, but across the
      Agency as a whole.

c.   There now exists an urgent need for improved data
      integration capabilities. These capabilities provide
      the underpinnings for the comprehensive analyses of
      environmental conditions that increasingly guide EPA's
      initiatives in protecting and improving the environment.
      Examples of such analyses include risk assessment,
      compliance behavior determination, vulnerability
      assessments, "hot-spot" identification, research and
      modeling, and special inter-program studies.

d.   Environmental legislation and regulation often define
      specifically the meaning of the term "facility" for EPA
      programs. The resulting differences in use of facility
      identifiers make it difficult to compare and integrate
      information on the same facility in different data
      bases.

e.   A variety of descriptors, such as SIC (Standard
      Industrial Classification) codes and DUNS (Data
      Universal Numbering System) numbers, are available to
      programs for describing the corporate characteristics
      of a site. A new data standard is needed to establish
      that, for EPA information management purposes, the
      uniqueness of a facility is based upon its location
      rather than corporate characteristics.

f.   EPA has implemented other data standards, as well as
      standards for hardware and software. Adoption of a
      standardized facility identification code will help the
      Agency realize the potential benefits of these
      standards for information integration and analysis.

5. AUTHORITIES.

    a.   15 CFR Subtitle A, Part 6, Standardization of Data
          Elements and Representations.

    b.   OMB Circular A-130, Management of Federal Information
          Resources.

6. POLICY. This Order establishes an EPA standard that each
   facility regulated by EPA will have a unique facility ID
   code, which will be the same for that facility across all
   EPA collections of information. Therefore:

   a.   Any collection of information assembled by or for EPA
        with data that describe a facility regulated by EPA
        under authority of Federal environmental statutes,
        including facilities regulated by state programs with
        delegated authority from EPA, shall contain in each
        record on or related to a single facility the facility
        identification code described in this Order and its
        appendix. The identification code for any one facility
        will be the same in all EPA collections of information.

        At the discretion of program offices, nonregulated
        facilities may also receive EPA IDs.

   b.   The facility identification code shall be comprised of
        a unique 12-character identification code controlled
        and issued through the EPA central facility data base.

   c.   The central facility data base will be operated by the
        Office of Information Resources Management. The
        objectives in maintaining this central data base are:

        (1) To provide a concise, comprehensive inventory of
             facilities regulated by EPA

        (2) To provide users with a simple method for
             determining a facility's ID code and ascertaining
             which program systems keep information on each
             facility.

   d.   The data element field used to store the identification
        code in program systems should be readily accessible to
        system users. The data element need not be a required
        field for initial data entry, but should always
        eventually be filled.

   e.   Program personnel may continue to use any
        program-specific identifiers (including DUNS numbers)
        needed to support the program mission, provided that
        such identifiers are kept in addition to the facility
        identification code established in this Order.

f.   In the interest of maintaining the confidentiality of
     Confidential Business Information (CBI) data kept by
     the Agency, information collections with CBI data will
     be considered as special cases.

g.   The Agency goal is to complete implementation of this
     standard by 1995, in recognition of the extensive time
     and resource commitments required by programs and OIRM
     to implement the standard.

h.   Once this Order becomes effective, adherence to the
     standard will become a key step in the development of
     new information collections. Existing information
     collections may be made consistent with the standard
     through phased implementation. This phasing will take
     into account system capabilities and needs.

7. RESPONSIBILITIES.

a.   The Office of Information Resources Management (OIRM)
     shall:

     (1) Develop, implement, and ensure adherence to this
          data standard.
     (2) Develop a management plan describing steps for
          implementation of the standard.
     (3) Provide guidance and technical assistance in
          meeting the requirements of this standard.
     (4) Provide unique facility identification codes in an
          efficient and responsive manner.
     (5) Maintain a central facility data base with
          identification codes and basic information
          associated with each facility.
     (6) Oversee resolution of conflicts regarding
          applicability or other issues relating to the
          standard.

b.   Assistant Administrators, Associate Administrators,
     Regional Administrators, Laboratory Directors, the
     General Counsel, and Heads of Headquarters Staff
     Offices shall establish procedures within their
     respective organizations to ensure compliance with the
     requirements of this data standard. Such procedures
     shall include the following:

     (1) Oversee development of individual program
          implementation plans.

     (2) Ensure that program representatives supply the central facility data base with program-specific IDs and with the key information necessary to assign the EPA ID code (e.g., name, address, etc.).

     (3) Ensure that each system containing facility information maintains a data element field for the standardized facility identification code.

     (4) Ensure that each facility record -- both new and existing -- carries the standardized facility identification code from the central facility data base.

     (5) Inform facility representatives of their ID codes, as appropriate, and incorporate the codes into reporting forms.

     (6) Notify OIRM of any difficulty in meeting the requirements of the data standard within the time frame specified in the program implementation plan.

c. The Procurement and Contracts Management Division shall work with EPA Program and Regional Offices and Laboratories to ensure that, where appropriate, the requirements of this standard are incorporated into EPA contracts.

d. The Grants Administration Division shall work with EPA grants management offices to require a special condition in future award documents mandating assistance recipients to use facility ID codes for any facility-related information collected under the assistance agreement, in cases where recipients are acting for EPA under delegated authority.

8. DEFINITIONS.

a. A "FACILITY" is a locational entity, deliberately established as a site for designated activities, but not primarily for habitation (even though on-site habitation may be necessary to the execution of the primary activities). Examples include a factory, a military base, a college, hospital, national park, office building, or prison. (Adapted from FIPS Pub. 55.)

In some cases, a facility with complex, multiple functions may have several plants or establishments operating within its property boundaries. For these facilities, ID codes will generally be assigned to the most comprehensive "level." However, a complex facility with multiple establishments or operations may receive several IDs, if more than one code is appropriate.

A facility may include wells or pipes located off the facility property. Although these "sub-units" will not ordinarily receive separate facility ID codes, program systems should be able to associate data on the sub-units with the record for the facility itself.

b. "FACILITY IDENTIFICATION CODE" is a 12-character code that uniquely identifies a facility. The appendix to this Order describes the format of the code in detail.

c. A "FACILITY IDENTIFICATION DATA STANDARD" is the requirement, in terms of format and content, that every record of information referring to a particular facility contain a data element field with a unique facility identification code. This code is to be used consistently across all collections of information containing information on the same facility.

9. PROVISION FOR WAIVER. In general, OIRM will attempt to work with program offices to develop a feasible implementation schedule for the standard. There may exist, however, cases in which exceptions to the requirements of this Order are warranted. In these cases, program offices shall demonstrate reasons for waiver. The process to apply for a waiver is as follows:

a. Draft an application for waiver to OIRM outlining the reasons why the facility identification data standard should not be implemented in the information collection.

b. Obtain approval by the decision official in the requesting office and the respective Senior Information Resources Management Official (SIRMO).

c. Submit application to the Director of OIRM, who has responsibility for final disposition.

The applying office will be notified in writing of the disposition of the waiver.

10. PROCEDURES.   The appendix to this Order contains preliminary information relevant to implementation of the standard. OIRM will also issue a management plan, which will describe the general steps and overall schedule for implementation of the standard over the next five years.   In addition to the Agency management plan, programs will develop individual implementation plans together with OIRM.   These plans will consider program-specific capabilities and needs.

Charles L. Grizzle 4/9/90

Charles L. Grizzle
Assistant Administrator for
Administration and
Resources Management

APPENDIX TO THE FACILITY
IDENTIFICATION DATA STANDARD

1.0 INTRODUCTION

    1.1 Purpose of Appendix

        The purpose of this appendix is to provide further
    detail on the facility identification data standard
    announced in the preceding Order. While the Order
    introduces the data standard, it does not contain the full
    level of detail necessary for programs to form a working
    understanding of the standard.

        In addition to this appendix, an Agency-wide
    implementation plan will be issued through the Office of
    Information Resources Management (OIRM). This
    implementation plan will cover in detail such issues as
    overall schedule for implementation, instructions for
    applying for waivers, procedures adopted by OIRM to ensure
    adherence to this standard, and the role of the central
    facility index system (FINDS).

    1.2 Background of the Facility Identification Data
         Standard

        The Agency has long striven to create a standard
    coding scheme for facility ID codes that could be used not
    only in individual programs, but across the Agency as a
    whole. Such a scheme would help EPA more readily generate
    facility-specific responses to public inquiries and
    determine patterns of compliance behavior across programs.

        These reasons, however, are only part of the
    justification for adopting this new standard. EPA is
    experiencing a vigorous trend towards cross-media analyses
    such as site characterizations, risk assessments and other
    environmental analyses that require the integration of data
    on individual facilities from diverse sources. The net
    result is an even greater need for data integration and
    sharing across different environmental media and programs.
    This trend is often acknowledged by staff in single-media
    programs, who now face increasing demands for data sharing
    and integration.

OIRM, in order to support the Agency's needs for data sharing and integration, has established an information management policy on Data Standards. Substantial implementation costs might be expected to be associated with promulgation of numerous retroactive data standards, and in the start-up phase many programs responsible for implementing these standards would derive few immediate benefits. For this reason, OIRM is taking a very measured approach and has identified only two standards as being of paramount importance.

The facility ID data standard is one. The other is embodied in the locational data policy, which establishes latitude and longitude, in an internationally-compatible format, as the Agency's preferred locational coordinate system. Both are undergoing concurrent green border review and will be complementary when implemented. Adoption of these selected standards at this time is absolutely essential if the Agency is to respond effectively to the anticipated data integration needs of the future.

As discussed above, OIRM will issue a management plan for the facility ID data standard, which will describe the general steps and overall schedule for implementation of the standard over the next five years. In addition to the Agency plan, programs will develop individual implementation plans together with OIRM. These plans will consider program capabilities and needs with respect to implementation.

OIRM has attempted to furnish program personnel with ample opportunity to provide input on the development of the facility identification data standard. To this end, program and system managers from EPA Headquarters, Regions, and states have been asked to contribute both to the development of the standard and to the formulation of this Order. Only with full program support and participation can OIRM establish a data standards program that will meet EPA's needs both now and in the future.

2.0 STRUCTURE OF THE FACILITY IDENTIFICATION CODE

The facility identification code is the key feature of this facility identification data standard. This code will consist of a 12-character standardized identification (ID) code. The first two characters will be the Federal Information Processing Standard (FIPS) two-letter abbreviation for the state or territory in which the facility is located. Abbreviations will also be used for facilities located outside of the U.S., for example in Canada or Mexico.

The remaining ten characters of the code will include a
check digit, will have no inherent meaning, and will not attempt
to describe, categorize or classify the facility in any way.
This applies only to new facility ID codes issued after the
effective date of implementation. Identification codes issued
previously under OIRM's Facility Index Data System (FINDS) may
be retained.

3.0   SOURCE OF THE STANDARDIZED FACILITY ID CODE

OIRM will maintain a central facility data base with basic
information on each facility such as name, address, etc. This
data base will serve as an inventory of facilities of interest
to EPA. It will also be the means through which the unique,
standardized facility ID codes will be assigned. OIRM, in
cooperation with program offices, will develop methods and
procedures for assigning unique ID codes efficiently. These
procedures will be defined in the program-specific
implementation plans, and will be automated as fully as possible

To create or identify a pre-existing ID code, key information
on the facility, including particularly the facility name and
address or other location information, must be made available to
the central facility index system via these procedures. OIRM
will develop effective and efficient procedures for ID
assignments.

4.0   SCOPE OF THE FACILITY IDENTIFICATION DATA STANDARD

Section 2 of the Order, Scope and Applicability, presents
the official scope of the facility identification data standard.
 This section augments that discussion in an attempt to
anticipate questions that program managers may have regarding
applicability of the standard.

4.1   Definition of a Facility

Successful implementation of a facility identification
data standard hinges on a consistent understanding of what
"facility" means. A common understanding is necessary for
determining to which facilities the standard applies. An
Agency-wide definition of "facility" is difficult to
establish, however, because of differences in Federal
legislation, which prescribe the program definitions of
"facility." As a result, perceptions of facilities differ
from program to program.

Differences notwithstanding, most program managers have developed a common-sense definition of a facility. In this approach, a facility is a single contiguous property--either an entire property or the portion thereof that is of regulatory concern. Variability in definitions of facilities usually occurs because different portions of facilities are regulated, not because of differences in facilities per se.

OIRM intends to establish a Facility Advisory Committee composed of Agency program office representatives to assist in implementing the data standard. Should further clarification of a standard facility definition be required, the committee may be called to resolve the issue.

4.2 Inclusion of Non-EPA-regulated Facilities

If desired, a facility ID code may be assigned to a facility not regulated by EPA. Such a scenario might arise, for example, if program personnel need ID codes for facilities of interest to EPA, but located outside of the United States. Additionally, ID codes can be assigned to state-regulated facilities, if appropriate.

4.3 Use of Concurrent Program Identifiers

Many program-specific codes are now used to identify facilities (e.g., permit numbers or site IDs). Program personnel may continue to use these identifiers in their own data collections under this standard, provided that the program information system accommodates both the standardized ID code and the program-specific identifiers.

If an information collection stores other identification codes in addition to the standardized ID code, these codes should be provided to the central data base under the procedures outlined in each program implementation plan. The central data base can then store listings of both facility and program-specific ID codes, which will further facilitate data integration and sharing.

4.4 Use of Previously Assigned EPA ID Codes

Facilities regulated by some programs have already been assigned EPA ID codes through FINDS under the current, DUNS-based numbering scheme. (DUNS is the Dun & Bradstreet Data Universal Numbering System for corporations.) These codes do not have to be replaced, i.e., new standardized ID

codes for these facilities need not be assigned. However,
if an EPA ID number was assigned internally by program
staff but was not entered into FINDS, OIRM should be
notified of this number and the facility name and address
associated with it. This notification will enable the
central data base to keep a complete record of all EPA
facilities and their associated identification codes.

5.0 **RELATIONSHIP OF THE FACILITY IDENTIFICATION DATA STANDARD
     TO FINDS**

Assignment of EPA ID codes already occurs through a central
data base--FINDS--for several EPA programs. This data base
will continue to operate under the new data standard. Some
features of the current FINDS operation will remain the
same, while others will be modified to make the system more
responsive to program needs under the new standard.

5.1 **Similarities Between Old and New Operations**

Similarities between the old and the new FINDS are as
follows:

- Program data managers will obtain facility ID
  codes through the central data base.

- The central data base will continue to house
  basic descriptive information on each facility.

- The central data base will have the Dun and
  Bradstreet (D&B) file available to obtain
  descriptive information on facilities.
  Information supplied by the program offices will
  be used in conjunction with D&B data to ensure
  the accuracy of facility information in the
  central data base.

- Program personnel will work with the OIRM system
  manager in keeping the descriptive information
  up-to-date by notifying OIRM through automated
  procedures when they become aware of a change in
  this information. OIRM will work with program
  offices to accomplish regular and responsive
  updating.

5.2   <u>Differences Between the Old and the New FINDS</u>

        Differences between the current and the new FINDS are
as follows:

        •   The new identification code will no longer be
              based on the 9-digit DUNS identification code.

        •   EPA will actively encourage states to use FINDS
              and to supply information to it on facilities
              regulated by state law.

        •   To the extent possible, the system will store
              information on all regulated facilities reported
              to EPA, and will not exclude facilities based on
              size or type of operations.

5.3   <u>Use and Availability of DUNS Numbers</u>

        OIRM has based the decision to use a unique,
nondescriptive ID rather than a DUNS-based ID on a
determination that, due to facility definition problems,
the DUNS number does not fully meet EPA's needs for a
unique ID code. For example, when an establishment moves,
the DUNS number moves with it, and many EPA programs in
this case would require new permit numbers to be assigned.

        The fact that the facility ID code will no longer be
based on the DUNS number does not diminish the importance
of these numbers. OIRM continues to encourage programs to
collect and maintain DUNS numbers in their information
systems, if these numbers are needed for the program
mission. So that program personnel can obtain these
numbers more easily, EPA will continue to purchase on-line
access to the Dun and Bradstreet file of facilities.

5.4   <u>Relationship of Facility Identification Data Standard
      to Locational Data Policy</u>

        OIRM has previously issued two data standards, one for
use of Chemical Abstract Service numbers and the other for
electronic transmission of laboratory data (EPA Orders No.
2180.1 and No. 2180.2). OIRM is also planning to issue a
locational data policy, currently in green border review.
This policy establishes the principles for collecting and

documenting latitude/longitude coordinates for facilities,
sites and monitoring and observation points regulated or
tracked under federal environmental programs within EPA's
jurisdiction.

All of these efforts are related in that they increase
the potential for data sharing and integration. The
locational data policy, however, is directly related to the
facility identification data standard. These two efforts
complement each other by providing primary identification
and locational information on entities of interest to EPA.
An important difference in scope is that the facility ID
data standard applies primarily to regulated facilities,
whereas the locational data policy covers both facilities
and other locations, such as monitoring stations, where
environmental data are collected. FINDS will continue to
allow entry of locational information for facilities.

6.0   IMPLEMENTATION OF THE FACILITY IDENTIFICATION DATA STANDARD

OIRM intends to establish an advisory committee composed of
Agency program office representatives to assist in the
implementation of the data standard.

6.1   Schedule for Implementation of this Standard

In the forthcoming Agency implementation plan, OIRM
will publish the general schedule for implementation of the
standard over the next five years. This plan will be
followed by development of individual program
implementation plans, which programs will formulate
together with OIRM. For the individual program plans, OIRM
will work on a one-to-one basis with program staff to
develop a realistic implementation schedule that takes into
account such factors as mission priorities, system
capabilities, and the size of the information collection.
OIRM recognizes that implementation of the standard is a
large-scale undertaking, and it has already considered that
for some program offices, implementation should be phased
over several years.

6.2 Procedures for Adopting the Standard

The program implementation plans will discuss:

- procedures to obtain IDs for facilities that are being newly regulated by the program

- procedures to obtain new IDs for facilities with existing records in program data bases

- methods for communicating changes in facility descriptive information.

The plans will define individual procedures for the exchange of facility IDs and related information between specific program systems and FINDS.

6.3 Procedures Adopted by OIRM for Assuring Adherence to this Standard

In addition to the assurance that comes from extensive communication with program offices when developing and executing implementation plans, OIRM has a number of alternatives for assuring adherence to the information management requirements of the standard. These include:

- Exercising the right to refuse to endorse any computer-based collections of information that do not include a field for the standard facility ID code

- Not concurring on forms for systems-related information collection that do not reference the standard facility ID code.

Further potential enforcement measures as well as procedures for sharing enforcement responsibilities with other EPA management offices will be addressed in the implementation plan.

⊕EPA

ORDER

7500.1
9/11/89

MINIMUM SET OF DATA ELEMENTS FOR GROUND-WATER

1. PURPOSE. This policy statement establishes the minimum data elements to be collected and managed by the Environmental Protection Agency (EPA) ground-water data collection activities. The underlying purpose is to more efficiently manage and share data within the ground-water community, including States, local governments, the regulated community, EPA, and other Federal agencies.

2. SCOPE AND APPLICABILITY. This policy applies to all ground-water data collection activities directly carried out by EPA staff or EPA contractors, including research and development, and enforcement.

   The collection of ground-water data by EPA's delegated state programs and grantees, States, localities, the regulated community, and other Federal agencies is not within the scope of this policy. However, to encourage the efficient sharing of data within the ground-water community, these entities are encouraged to adopt this minimum data set where appropriate. In the future, some of these entities may be required by EPA to comply with this policy when, after opportunity for notice and comment during the Agency's normal regulatory or policy development processes, the Agency determines compliance is necessary for accomplishing its statutory mandates.

3. BACKGROUND. In 1986 EPA completed its Ground-Water Monitoring Strategy and in 1987 it completed a Ground-Water Data Requirements Analysis. Both of these reports provide a clear message to EPA Program and Regional Offices, and the States, on the importance of, and need for, ground-water data standards. The establishment of a minimum set of data elements for EPA ground-water data collection activities is a first major step in this process.

The Agency has defined the minimum set of data elements as those basic elements necessary for it to use data from wells and springs across ground-water related programs. The set is characterized by elements that:

o   are needed to communciate ground-water data across programs;

o   facilitate accessing data from automated information systems other than one's own;

o   provide to the users a road map for finding other related data; and

o   provide a link between water quality and well location information.

4.   POLICY. It is EPA policy to collect and maintain consistent ground-water data based on a minimum set of data elements. EPA Program Offices and Regions will incorporate this policy into their ground-water data collection activities through appropriate mechanisms such as regulations, policies, directives, orders, guidance, or procedures, when consistent with statutory and other significant policy, administrative, and technical considerations. When Agency programs employ alternative data collection schemes in future activities, regulatory or other documentation supporting the new activity will show how the alternative dlata collection scheme relates to the minimum set of data elements and will provide a rationale for the differences.

This policy does not require modifications to existing regulations to bring them into accord with policy; however, the policy should be considered during the reauthorization process for regulations.

a.   The minimum set shall consist of the following:

Latitude
Longitude
Method of Measure for Latitude/Longitude
Source Agency for Latitude/Longitude
State FIPS[1] Code

---

[1] Federal Information Procedures System (FIPS)

County FIPS Code
Altitude
Unique Well/Spring and Facility
  Identification
Use of Well
Depth of Well at Completion
Depth to Top of Open Interval
Depth to Bottom of Open Interval
Location of Log
Type of Log
Source Agency for Sample Data
Sample Date
Parameter Measured
Concentration/Value (in standard units)
Confidence Factor (Field and Lab quality
  assurance)
Sample Identification
Depth to Water
Measurement Quantification

b. A dictionary defining elements in the minimum data set
    will be developed by the Office of Ground-Water Protection
    (OGWP). Data standards and formats will also be developed,
    where appropriate, by the Office of Information Resources
    Management (OIRM).

5. RESPONSIBILITIES.

    a. The Office of Ground-Water Protection (OGWP), in
        cooperation with EPA Program Offices, shall provide
        guidance with technical assistance in implementing the
        requirements of this policy.

    b. Assistant Administrators and Associate Administrators
        shall ensure that Programs and activities under their
        direction are in compliance with this policy. EPA Program
        Offices and Regions can add additional elements to the
        minimum data to meet their special·program needs.

    c. Regional Administrators shall assure that the Regional
        Offices have an effective program to foster and support
        the minimum set of data elements for reporting ground-
        water data.

d. The Office of Information Resources Management (OIRM) shall provide effective leadership by helping other Agency Offices comply with this policy. OIRM shall be involved as part of the review process in cases where Agency Offices employ alternative data collection schemes. OIRM shall work with other Agency Offices to develop any requisite data standards for individual elements of the minimum set.

e. The Procurement and Contracts Management and Division shall work with EPA Program and Regional Offices and Laboratories to ensure that, where appropriate, the requirements of this policy are incorporated into EPA contracts.

Charles L. Grizzle
Assistant Administrator