EPA

2180 - Locational Data Policy Implementation Guidance

TABLE OF CONTENTS

EXECUTIVE SUMMARY i-1

Chapter 1 INTRODUCTION 1-1

1.1 Purpose of Locational Data Policy 1-1
1.1.1 Policy Requirements 1-1
1.1.2 Policy Benefits 1-1
1.1.3 Scope of Participation 1-2
1.1.4 Locational Accuracy Task Force 1-2
1.2 Background 1-2
1.2.1 Motivation for Policy Development 1-3
1.2.2 Concurrent Federal Agency Data Requirements 1-3
1.3 Contents of This Guidance Document 1-4

Chapter 2 REQUIREMENTS 2-1

2.1 Applicability and Scope 2-1
2.1.1 Data Types Affected by the LDP 2-1
2.1.2 Entities under the Scope of the LDP 2-3
2.1.3 Data Collection Activities Affected by the LDP 2-4
2.2 Data Documentation Requirements 2-6
2.2.1 Latitude 2-6
2.2.2 Longitude 2-7
2.2.3 Method 2-8
2.2.4 Accuracy 2-9
2.2.5 Description 2-13
2.3 Recommended Locational Data 2-13
2.4 Special Considerations 2-13
2.4.1 "Tiering" and Spatial Extent 2-13
2.4.2 Standard Datum 2-15

Chapter 3 RECOMMENDATIONS OF THE LATF 3-1

3.1 Introduction 3-1
3.2 Summary of the LATF Recommendations 3-1
3.3 Specifics of the LATF Recommendations 3-2
3.3.1 Accuracy Goal 3-2
3.3.2 Technology-Based Locational Data Accuracy 3-3
3.3.3 Waiver Applications 3-4
3.3.4 Incentives 3-5
3.3.5 Enhancements to FINDS 3-5

Chapter 4 RESPONSIBILITY AND AUTHORITY 4-1

4.1 The IRM Steering Committee 4-1
4.2 OIRM Responsibility and Authority 4-1
4.2.1 IMSD 4-2
4.2.2 PSD 4-5
4.2.3 Resources from OIRM 4-5
4.2.4 Authority 4-6

TABLE OF CONTENTS (continued)

4.3 Other EPA Offices and Programs 4-7
4.3.1 Non-Media Program Office Responsibilities 4-9
4.3.2 Media Program Offices 4-15
4.3.3 Individuals/Functions 4-17

4.4 State, Indian Tribe, and Trust Terrotory Delegates 4-20
4.4.1 Variability of State LDP Responsibilities under Different Types of Federal Environmental Laws 4-20
4.4.2 Reporting Requirements 4-22
4.4.3 EPA Support to States for LDP Participation 4-23
4.5 Designated Representatives 4-24

Chapter 5 AGENCY-WIDE SCHEDULE 5-1

5.1 Schedule and Priorities 5-1
5.1.1 Accuracy Goal 5-1
5.1.2 LDP Implementation Plans 5-1
5.1.3 Existing and Planned Systems 5-3
5.1.4 Historical Data 5-4
5.1.5 Form and Contract (Tool) Modifications 5-4
5.1.6 Role of FINDS 5-5
5.2 Coordination with Other Federal Agencies 5-6
5.2.1 Office of Management and Budget (OMB) 5-6
5.2.2 Federal Geographic Data Committee (FGDC) 5-6
5.2.3 Federal Geodetic Control Committee (FGCC) 5-7

Chapter 6 LDP IMPLEMENTATION PLANS FOR PROGRAMS AND ENVIRONMENTAL INITIATIVES 6-1

6.1 Components of LDP Implementation Plans 6-2
6.1.1 Identification of Entities to Which the LDP Applies 6-2
6.1.2 Definition of Activities During Which LDP Data Will Be Collected 6-3
6.1.3 Enhancements to Data Bases and Systems 6-3
6.1.4 Tool Redevelopment for Collecting and Documenting Locational Data 6-5
6.1.5 The Format of LDP-Required Data 6-9
6.1.6 Identification of Methods to be Employed 6-9
6.1.7 Definition of Roles and Responsibilities 6-7
6.1.8 Resources/Priorities 6-11
6.1.9 Quality Assurance Measures 6-12
6.1.10 Constraints 6-12
6.1.11 State Roles and Responsibilities 6-12
6.2 Process for the Developing LDP Implementation Plans 6-13
6.2.1 Coordination of Steps to Achieve 6-13
6.2.2 Phasing 6-13

TABLE OF CONTENTS (continued)

Chapter 7 DATA CONVERSION AND COLLECTION METHODS (GEOCODING) 7-1

7.1 Centralized vs. Incremental Data Conversion and Collection 7-1
7.2 Locational Data Collection Methods 7-2
7.2.1 Conversion from Non-Coordinate Geopolitical Locational Information to Latitude/Longitude 7-2
7.2.2 Conversion from One Coordinate System to Another 7-7
7.2.3 Conversion from One Datum to Another 7-8
7.2.4 Guidelines on How to Choose a Locational Data Conversion Method 7-9
7.3 Locational Data Collection Methods 7-10
7.3.1 Conventional Surveying 7-13
7.3.2 Map Interpolation 7-13
7.3.3 Photo Interpretation 7-15
7.3.4 LORAN-C 7-16
7.3.5 Global Positioning System (GPS) 7-18
7.3.6 Hybrid Systems 7-20
7.3.7 Guidelines on How to Choose a Locational Data Collection Method 7-20

Chapter 8 SYSTEM IMPLICATIONS 8-1

8.1 Location Identification Data Management 8-1
8.1.1 Relationship Between LDP Data Elements 8-1
8.1.2 Structure and Format of LDP Data Elements 8-2
8.1.3 Position and Function of Location Identification Data 8-4
8.1.4 Relationship to Other Locational Data 8-5
8.1.5 Edit Checks 8-5
8.2 Locational Data Sharing 8-6
8.2.1 State Entry in FINDS 8-6
8.2.2 GRIDS 8-6
8.2.3 National Spatial Data Transfer Standard 8-7

APPENDIX A - Locational Data Policy and LATF Recommendations A-1

APPENDIX B - Contacts for LDP Implementation Assistance B-1

APPENDIX C - Partial List of Forms Relevant to LDP C-1

APPENDIX D - Facility Definition Guidance from the Facility Identification Data Standard Implementation Plan (FIDSIP) D-1

LIST OF EXHIBITS

Exhibit 2-1 Codes to Document *Method* 2-9

Exhibit 2-2 Concept of Precision vs. Accuracy 2-11

Exhibit 2-3 Relationship of Method, Description, and Accuracy to Lat/Long Coordinates 2-16

Exhibit 4-1 OIRM Organizational Chart 4-3

Exhibit 4-2 OIRM Reponsibilities under the Locational Data Policy 4-4

Exhibit 4-3 EPA Office and Program Responsibilities under the Locational Data Policy 4-8

Exhibit 4-4 Interaction between the Regional GIS Work Group and the GIS User Community 4-12

Exhibit 4-5 Components, Goals, and Objectives of the Regional GIS Work Group Strategic Plan 4-14

Exhibit 4-6 Individual Responsibilities under the Locational Data Policy 4-18

Exhibit 5-1 Agency-wide LDP Implementation Schedule 5-2

Exhibit 6-1 LDP Implementation Plan Summary 6-4

Exhibit 6-2 Example of Table of Contents for LDP Implementation Plan 6-6

Exhibit 6-3 Review of the 637 Forms in the EPA Forms Catalog 6-10

Exhibit 7-1 Incremental vs. Centralized Data Converion and Collection 7-3

Exhibit 7-2 Comparison of Locational Data Conversion Methods 7-4

Exhibit 7-3 Comparison of Locational Data Collection Methods 7-11

Exhibit 7-4 Cost vs. Accuracy Curves for Collection Methods 7-12

Executive Summary

Background

The *Locational Data Policy* (LDP) was made effective on 17 May 1990 when, after formal Agency-wide review, it became an official directive under the 2100 series and Chapter 13 of the EPA Information Resources Management Policy Manual. The purpose of the LDP is to ensure the collection and documentation of accurate, consistently-formatted, fully-documented, latitude/longitude (lat/long) coordinates as part of all spatially-relevant data gathering activities. Under this policy, collection and documentation of locational information will be performed for all facilities, sites, monitoring points, and observation points regulated or tracked by EPA under Federal environmental laws. In addition to lat/long coordinates, the LDP requires documentation of specific information regarding the method used to measure lat/long, the accuracy of the measurement, and a description of the entity that the lat/longs represent.

Other notable points are as follows:

- All Agency-sponsored data collections and activities that define/describe attributes (environmental characteristics) *about a place* are within the scope of the LDP.
- As an established Agency-wide goal for new data, coordinate points are to be *accurate to within 25 meters*.
- The LDP establishes a technology-based standard (i.e., lat/long coordinates should be obtained from the best practicable geocoding technology).

Due to its potential for yielding consistent, highly accurate measurements, the Agency currently considers *Global Positioning Systems* (GPS) to be the best practicable technology. With implementation of the full constellation of GPS satellites scheduled for completion in 1992, and with the proliferation of commercial measuring devices, GPS should serve as the method of choice until a more accurate or less expensive technology (with similar accuracy) is introduced.

Policy Implementation Guidance

Locational Data Policy Implementation Guidance is provided in four documents: the *Guide to the Policy*, (i.e., this document) the *Guide to Selecting Latitude/Longitude Collection Methods*, the *GPS Primer*, and the *Summary of Comments on the February 1991 Guide to the Policy*. The *Guide to the Policy* explains the basic collection and documentation requirements of the LDP and the Agency-wide steps for its implementation. The *Guide to Selecting Latitude/Longitude Collection Methods* provides cost and accuracy estimates for

several common geocoding methods. It also offers procedures for selecting a geocoding method and for estimating resource requirements. The *GPS Primer* provides instructions on how to optimally apply GPS technology. The *Comments Summary* presents the comments that were received regarding a prior draft of the *Guide to the Policy*, upon which this version is based.

Policy Specifications

This *Guide to the Policy* provides specifications for collection and documentation of all information relevant to EPA's locational data policy:

- Complete locational data include lat/long coordinates, method, accuracy, and description. Documentation of lat/long, method, and accuracy have prescribed formats; description is free-format.

- Measurement technology, datum, and scale (if applicable) should be documented as part of "method" information; North American Datum (NAD) 1983 should be used at this time; GPS is the preferred technology.

- Accuracy should be 25m or better for all *new* data collected after 12/31/91 and all *existing* data by 12/31/95, and GPS should be phased in for wide-scale use by 12/31/95.

- Locational data should correspond to the tier (entity of environmental concern) for which attribute data are collected.

- The dimensional quality of locational data (i.e., points, lines, or polygons) should be determined by the use of the data and the scale of the map or analysis.

- Date of collection and source of locational data (i.e., who collected it) should be (but is not *required* to be) documented in addition to lat/long, method, description, and accuracy.

The Locational Accuracy Task Force (LATF)

The LATF was created in 1990 to assess the Agency's need for a minimum accuracy requirement as part of the LDP. The LATF presented its recommendations to the IRM Steering Committee in December of 1990, stating that:

- All Agency locational data should be accurate to within 25m by the end of 1995

- The best practicable geocoding technology, currently GPS, be phased in for wide-scale use

- A program of incentives be applied to encourage states and the regulated community to comply with the LDP
- EPA's Facility INDex System (FINDS) should be enhanced and used as a functional repository of locational data for regulated facilities
- A program of waiver applications from LDP requirements should be established for approval by EPA's IRM Steering Committee.

These recommendations were endorsed by the EPA Deputy Administrator and were incorporated into the 1991 revision of Locational Data Policy.

Roles and Responsibilities

Responsibilities for implementing the LDP are shared by many groups. A general template for activities is as follows:

- The IRM Steering Committee will oversee Agency-wide LDP adoption and evaluate any requests for LDP waivers.
- OIRM will provide information resources, policies, and implementation guidance.
- ORD will evaluate technological developments and provide DQO guidance.
- OPPE will provide assistance to programs where LDP-required data are collected from the regulated community.
- Managers of program offices and environmental initiatives will prepare LDP Implementation Plans and institute procedures within their programs (including delegated states) to assure LDP requirements are met.

EPA regional offices and their state counterparts will be key players in implementing the policy by ensuring that LDP requirements are followed in the field. State, Tribal, and Trust Territory delegates and EPA agents are subject to the same requirements as EPA program offices.

LDP Implementation Plans

Each program office or environmental initiative should complete a comprehensive LDP implementation plan that addresses the numbers and types of entities to which the LDP applies; the tools to be used for measuring and documenting locational data; definitions of priorities and resource allocations; quality control measures; and organizational roles and responsibilities. Data Quality Objective (DQO) reports are a valuable tool that can be used for program planning and waiver application, if

necessary. The IRM Steering Committee will monitor the development of all program LDP implementation plans and evaluate waiver requests based primarily on justifications demonstrated in the DQOs.

Agency-Wide LDP Implementation Steps and Schedule

Although the Policy Guidance specifies that full LDP implementation will be complete by December 31, 1995, *all newly generated data collections should comply with the LDP from this point forward.* As for existing data collections, general priorities have been established for bringing data into compliance with LPD requirements. Highest in priority are data collections with *geographic coordinates other than lat/long* (e.g., UTM and/or state-system coordinates). These data can be converted to the lat/long system using common algorithms. Of second priority are data collections with lat/long coordinates but *without the required metadata* (method, accuracy, and description). Of third priority are *data collections with no geographic coordinate data or metadata.* Although data collectors have flexibility to chose other implementation priorities based on their requirements or organizational processes, *all Agency locational data is expected to be in compliance with LDP requirements by the end of 1995.*

LDP Implementation Plans are to be completed and submitted to the IRM Steering Committee for review by June, 1992. Final approval should be targeted for September, 1992. Data collection should be accomplished by map interpolation beginning in 1992 and phased out as GPS becomes more available. Use of GPS should be widespread by the end of 1995. Collection and documentation tools (e.g., survey or permit application forms) should modified by March, 1993 to ensure LDP data are obtained in the course of normal field operations. Computer system enhancements for storing and accessing locational data should be planned for in 1992 and any redesigns should be completed by December 1995.

Methods of Locational Data Collection

This *Guide to the Policy* provides descriptions of a variety of methods used for collecting and converting locational data. These methods include address matching, map interpolation, LORAN-C, photo interpolation, surveying, and GPS (which should be used whenever possible, and phased in for wide-scale use by the end of 1995). For comparative purposes, each method is presented in terms of needed/anticipated expertise, effort, cost, and accuracy. The *Guide to Selecting Latitude/Longitude Collection Methods* also provides details on the various methods and on the procedures to follow for choosing a method and estimating resource requirements. The *GPS Primer* provides technical information on the use of GPS technology.

System Implications

Because the LDP affects an extensive number of Agency data collections, this document also offers general guidance for handling locational data in computer systems. The policy requires four pieces of information, as follows:

- *Lat/long coordinates* are a repeating set of numeric fields (two separate fields).
- *Method* is set of three separate, coded, alphanumeric fields identifying measurement technology, datum, and scale of map (if applicable).
- *Accuracy* is presented as a range of lat/long units (e.g., +/- degrees, minutes, or seconds).
- *Description* is a free-format, text field with specifications for the type of data to include.

In addition to containing the above data elements, it is highly recommended that data bases include information on the date of collection and source of data (e.g., EPA office, state organization, contractor, etc.). Computer systems with environmental data should offer users full capabilities to search through and select records based on any or all locational data elements. Edit checks for locational data are recommended. To facilitate LDP implementation, OIRM is working on providing general locational coordinates for all the facilities in the Facility INDex System (FINDS).

Chapter 1
INTRODUCTION

1. INTRODUCTION

Guidance for implementing the ***U.S. Environmental Protection Agency's Locational Data Policy (LDP)*** throughout the Agency and applicable regulated community is provided in a series of four documents produced by the Office of Information Resources Management (OIRM). These documents include the *Guide to the Policy,* (i.e., this document) the *Guide to Selecting Latitude/Longitude Collection Methods,* the *GPS Primer,* and the *Summary of Comments on the February 1991 Guide to the Policy*. The *Guide to the Policy* explains the basic collection and documentation requirements of the LDP and the Agency-wide steps for its implementation. The *Guide to Selecting Latitude/Longitude Collection Methods* provides cost and accuracy estimates for several common geocoding methods. It also offers procedures for selecting a geocoding method and for estimating resource requirements. The *GPS Primer* provides instructions on how to optimally apply GPS technology. The *Comments Summary* presents the comments that were received regarding a prior draft of the *Guide to the Policy,* upon which this version is based.

1.1 Purpose of the Locational Data Policy

The primary purpose of the Locational Data Policy is to ensure the collection of accurate, consistently-formatted, fully-documented locational coordinates for facilities, sites, monitoring points, and observation points regulated or tracked under Federal environmental programs within the jurisdiction of the U.S. Environmental Protection Agency (EPA). The intent of the policy is to allow data to be integrated based upon location, thereby promoting enhanced use of EPA's extensive data resources for cross-media environmental analyses and management decisions pursuant to the Agency's mission to protect human health and the environment. The objective of this policy is to improve the overall quality and compatibility of locational data throughout EPA.

1.1.1 Policy Requirements

The Locational Data Policy establishes principles for collecting and documenting geodetic coordinates defined in terms of ***latitude and longitude (lat/long)***. In addition to these locational data, the LDP requires documentation of specific information regarding the ***method*** used to measure lat/long coordinates, the ***accuracy*** of the measurement, and a ***description*** of the place where the lat/longs were taken. Further, the LDP commits OIRM to develop policy implementation guidance. This document, and its set of companions, fulfills that commitment.

1.1.2 Policy Benefits

Independent data collections can be integrated based upon location only after each set contains ***consistently-formatted, locational coordinate data of an expected or known level of quality.*** The capability to integrate data ***will***

enhance the utility of EPA's extensive data resources for secondary applications, such as cross-media analyses and decision-making. Known-quality locational data also will bolster the Agency's development of environmental risk management and pollution prevention strategies, methodologies, and assessments. In short, the ability to integrate data based upon location will increase the return on EPA's significant data investments.

1.1.3 Scope of Participation

This policy applies to all EPA organizations and agents, including state and local government personnel, directly responsible for, or who have delegated authority for, implementing Federal environmental laws. The policy applies to all organizations with responsibilities to support EPA, including contractors, universities, and grantees who design, develop, compile, operate, or maintain EPA information developed for environmental program support. Certain requirements of this policy apply to existing as well as new data collections.

In some instances, deviations (i.e., waivers) from specific requirements of this policy may be justified. Applications for waivers will be made to the IRM Steering Committee who will review the applications and grant or deny such requests.

1.1.4 Locational Accuracy Task Force (LATF)

Under the auspices of EPA's IRM Steering Committee, EPA and participating states (including Indiana, Minnesota, and Oregon) created the ***Locational Accuracy Task Force*** (LATF) to assess the Agency' need for a minimum accuracy requirement in the LDP. The LATF presented its final recommendations to the Steering Committee in December 1990. The recommendations of the LATF form a key component of the official Agency policy and guidance relating to locational identification for environmental data. The recommendations of the LATF are discussed in Chapter 3 and presented fully in Appendix A.

1.2 Background

EPA/OIRM has several major efforts underway to develop policies for sound management of information resources across the Agency. Wide-scale implementation of these policies will assure more uniform, consistent, and compatible information throughout all environmental programs. Agency-wide data standards adopted thus far include the use of *Chemical Abstract Service (CAS) numbers* to identify chemical substances and the assignment of *EPA Facility Identification Codes* to regulated facilities. Other Agency-wide data standards define the *Minimum Set of Data Elements for Collecting Ground Water Data* and the specific format for *Electronic Transmission of Laboratory Measurement Data.* As

another part of EPA's information management effort, the Agency instituted its *Locational Data Policy*. A discussion of the rationale and need leading to development of the LDP is provided below.

1.2.1 Motivation for Policy Development

The fundamental motivation in developing the LDP has been the growing need for integrated data to support activities associated with fulfilling EPA's mission to protect human health and the environment. In order for the Agency to satisfy demand for solid environmental assessment, risk-based decision making, and well-founded enforcement actions, EPA must be able to perform ***cross-programmatic, multi-media data analyses***. To perform these functions, a common denominator among program data collections must be developed and used as the basis for data integration. Uniform locational information of documented quality meets this need.

A tremendous opportunity to successfully conduct cross-programmatic, multi-media analyses has been created by the ***increasing availability and cost effectiveness of sophisticated technologies***. For example, Global Positioning Systems (GPS) now are available to collect extremely accurate locational data. Geographic Information Systems (GIS) are distributed throughout the Agency as a tool to perform integrated, location-based analyses.

Finally, ***data sharing is encouraged by EPA's State/EPA Data Management (SEDM) Program*** which links state environmental regulatory agencies and EPA in partnership. This program seeks to foster data sharing and to improve data quality and timeliness. The need for a common denominator exists for sharing data among EPA programs (i.e., Agency-wide) and between EPA and its state partners.

1.2.2 Concurrent Federal Agency Data Requirements

Another motivation for developing the LDP is the move to standardize collection and documentation of locational data throughout the Federal government. Several organizations are responsible for coordinating the digital cartographic and GIS activities of Federal agencies. For example, in two Director's Memoranda, dated 4 April 1983 and 18 March 1986, the Office of Management and Budget (OMB) charged the ***Federal Geographic Data Committee (FGDC)*** with recommending procedures and programs to:

- Facilitate the coordination of digital cartographic and geographic information system activities of Federal agencies
- Establish and promulgate standards and specifications for the production of digital cartographic data.

As a component of this mandate, the FGDC has created specifications for locational data, including the use of latitude/longitude as the preferred coordinate system.

Additionally, committees such as the ***Federal Geodetic Control Commission (FGCC)*** provide recommendations in other relevant areas, including the most appropriate geodetic datum to be used with locational data and approved conversion software.

OMB Circular A-16, dated 19 October 1990, established coordination mechanisms to promote development of a national digital spatial information resource. It identified specific responsibilities for various government agencies using locational data, such as the Department of State and the Department of Commerce. EPA's Locational Data Policy demonstrates its participation in the movement toward consistently-formatted, fully-documented spatial data across Federal agencies.

1.3 Contents of this Guidance Document

This document presents guidelines for implementing EPA's Locational Data Policy and contains the following chapters:

- ***Chapter 2 presents a detailed explanation of the policy requirements,*** including the types of entities, data collections, and activities to which this policy applies, and definitions of the required information.

- ***Chapter 3 outlines the recommendations of the Locational Accuracy Task Force,*** including the establishment of a 25 meter level accuracy goal for location identification data, endorsing use of GPS, establishing a waiver process, using incentives, and upgrading FINDS.

- ***Chapter 4 explains responsibilities of all participants for implementing the Locational Data Policy,*** including the resource commitments that must be made and the communication links that must be established.

- ***Chapter 5 lists priorities and schedules for implementing the policy Agency-wide,*** including historical data conversion, new data collection, and form/contract tool modification.

- ***Chapter 6 describes the creation of LDP implementation plans required of all participating environmental programs and initiatives,*** including discussion of a sample Table of Contents and required plan components.

- ***Chapter 7 provides information on available locational data collection and conversion methods,*** including guidance on selecting strategies appropriate for individual programs.

- ***Chapter 8 discusses the system implications of the Locational Data Policy*** including data field formats and system functional recommendations.

- ***Appendix A contains full text references*** of the Locational Data Policy and final recommendations of the LATF.

- ***Appendix B contains a list of contacts*** for management, technical, and other assistance in LDP implementation.

- ***Appendix C presents a partial list of EPA forms*** that may need to be modified to comply with the LDP.

- ***Appendix D contains guidance on defining a facility*** under EPA's Facility Identification Data Standard.

Chapter 2
REQUIREMENTS

2. POLICY REQUIREMENTS

Effective implementation of the Locational Data Policy depends, to a large degree, on a clear understanding of its requirements and specifications on the part of all participants. This Policy has several different types of requirements:

- Requirements for the types of data, entities, and activities to which this Policy applies, and
- Requirements for the values and formats that must be provided under this policy.

Each of these requirements is addressed in detail in the sections below. The goals for the level of 25 meter accuracy and wide-scale use of GPS are presented in Chapter 3.

2.1 Applicability and Scope

This section defines the types of data to which this policy applies, the entities for which locational data must be collected and documented, and the activities during which locational data collection should occur.

2.1.1 Data Types Affected by the LDP

The Locational Data Policy applies to all ***locationally-based information*** collected for environmental program support. As a guideline, ***whenever data are collected about a place,*** then corresponding lat/long coordinates and supporting metadata (method, accuracy, and description) also should be collected and documented.

Data Collections

The LDP applies to all EPA data collections that are locationally based (i.e., ***data about a place),*** both manual and automated. The LDP applies to the following types of data bases:

- ***Inventory data bases*** -- Data bases that contain inventories of regulated, tracked, or monitored entities. Examples include FRDS II, FURS, TRIS, FINDS, AIRS/AFS, IFD, etc.
- ***Compliance tracking data bases*** -- Data bases that contain data on compliance tracking for regulated entities. Examples include PCS, RCRIS, etc.
- ***Activity tracking data bases*** -- Data bases that track activities such as cleanup, inspections, or enforcement

actions at a place. Examples include CERCLIS, DOCKET, etc.

- ***Monitoring data bases*** -- Data bases containing ambient environmental quality data. Examples include STORET, AIRS/AQ, SCADS, ODES, etc.

Non-locationally based data collections, such as chemical reference systems (e.g., IRIS) or administrative systems (e.g., IFMS), are not affected by the Locational Data Policy.

Existing and/or New Data

The requirement for collecting and documenting lat/long, method, accuracy, and description applies to all existing as well as new locational data.[1] General types of existing data collections which may have to be modified are listed below (programs may assign their own priorities and approaches to improving the quality of their existing data):

- ***Data collections with geographic coordinates other than lat/long (e.g., UTM coordinates)*** -- Data element fields for lat/long, method, accuracy, and description should be created. Values can be supplied by conversion software. Other locational data (e.g., street address, elevation, etc.) are not superseded or precluded by information collected under the LPD.[2]

- ***Data collections with lat/long coordinates defining the locations of their entities, but without the lat/long qualifiers (method, description, accuracy)*** -- Data element fields for method, accuracy, and description should be created. Although "unknown" values may be documented as such, data managers should appreciate the restrictiveness of "unknown" data values and attempt to supply explicitly-defined values whenever possible.

- ***Data collections with no geographic coordinates*** -- Data element fields for lat/long, method, accuracy, and description should be created. Lat/long coordinates should be created by converting existing locational data

---

[1] "*The requirements of this policy apply to existing as well as new data collections.*" Ch.13, Sec. 2, p. 13-1, IRM Policy Manual

[2] "*... lat/long coordinates (are) to be collected and documented with environmental and related data ... in addition to, and not precluding, other critical locational data that may be needed to satisfy individual program or project needs, such as depth, street address, elevation, or altitude.*" Ch. 13, Sec. 5a, p. 13-3, IRM Policy Manual

(e.g., street address) to lat/long. Method, accuracy, and description fields should be labeled to indicate that the coordinate data were derived from conversion (see Chapter 7 for a description of conversion options).

OIRM has recommended a priority schedule for bringing these types of collections into compliance with the LDP. ***Data collections containing coordinates other than lat/long are assigned the highest update priority.*** The rationale behind this recommendation is that such collections are already coordinate-based and can most easily be made to conform to the LDP. Of ***second priority are existing data collections with incomplete or missing coordinate qualifiers.*** Of ***third priority, but to be completed by December 1995, are data collections with no coordinate data at all.*** Due to resource constraints, lat/longs for these third-priority data sets may have a fairly low level of accuracy.

All newly-collected environmental data must conform strictly to the formats and accuracy goals specified under the Policy. Requirements of the LDP and the nature of its minimum accuracy goal are discussed fully in Chapter 3.

2.1.2 Entities under the Scope of the LDP

The Locational Data Policy ***applies to all sites, facilities, points, or other entities regulated, monitored, or tracked under Federal environmental law.***[3] In short, ***any entity or place for which data are collected in support of Federal environmental law must have its locational coordinates defined according to the Locational Data Policy.*** In addition, the LDP applies to data collected to describe conditions of sub-units of a larger entity (e.g., a Superfund site). Examples of entities requiring LDP-compliant data include, but are not limited to:

- ***Facilities*** permitted, tracked, or monitored under Federal environmental law.
- ***Sites*** that are unique because of an environmental concern but which may not conform to a property boundary or other designation as a *facility*.
- ***Underground storage tanks*** (especially *leaking* tanks).
- ***Wastewater-discharge points*** as defined by and regulated under the National Pollutant Discharge Elimination System.

[3] *"...for facilities, sites and monitoring and observation points regulated or tracked under Federal environmental programs within the jurisdiction of ...EPA."* Ch. 13, Sec. 1, p. 13-1, IRM Policy Manual

- ***Public water supplies***, either well heads, key wells in a well field, reservoirs, or other points of water supply.

- ***Water bodies*** as defined by states under the Clean Water Act.

- ***Underground Injection Control (UIC) sites*** which, depending on the focus of the data collection activity, may be either a well field or a key UIC well in a well field.

- ***Emission sources*** of regulated air contaminants (whether stationary or portable) at permitted, registered, or exempt facilities.

- ***Monitoring sites*** where samples are collected to assess compliance.

- ***Sampling locations*** where samples are collected to determine environmental conditions.

Locational data requirements may not apply for entities that are transitory (i.e., those that change location frequently). Examples of such entities include waste haulers, mobile air emission sources, and portable operators. Program managers, however, may wish to capture the ***boundaries*** of the area within which the entity moves or the ***corridors*** along which the entity transports. Note that the LDP *does* apply to permanent records maintained on places that are of temporary environmental concern (e.g., spill sites).

2.1.3 Data Collection Activities Affected by the LDP

In the context of EPA's Locational Data Policy, ***"data collection activities" are the activities that require definition, collection, and documentation of locational coordinates.*** Again, the Locational Data Policy applies to activities ***about a place.*** Therefore, lat/long and supporting qualifiers should be collected and documented as an integral part of the following activities:

- ***Permit issuance*** -- Preparing and issuing permits to operate at a particular location within regulatory limits. Examples include permits for wastewater discharge, criteria pollutant emission, treatment/storage/disposal of hazardous wastes, etc.

- ***Compliance monitoring*** -- Monitoring sites to ensure compliance with Federal standards or permit limits. Examples include emissions monitoring points, wastewater discharge points, RCRA monitoring wells, applicable water supply facilities, etc.

- ***Enforcement*** -- Collecting environmental samples as part of case preparation.

- ***Reporting/notification tracking*** -- Tracking compliance with reporting requirements. Examples include EPCRA responders, incident reporting (e.g., spills), underground storage tank notifications, etc.

- ***Cleanup, environmental response*** -- Collecting data to determine the magnitude and extent of contamination at a site. An example would include an NPL site under Superfund.

- ***Environmental monitoring*** -- Routine environmental sampling to determine general environmental conditions for an area. Examples include ambient air or water quality monitoring, special studies (e.g., species diversity analyses), etc.

This policy does not apply to activities that are not necessarily locationally-based. For example, certain grants are awarded based on activity, not location (e.g., for methods evaluation), and certain enforcement actions are corporation-based, not location-based. Locational data are not required for these activities.

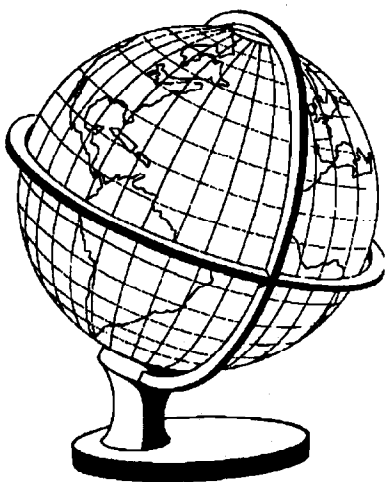
In addition, ***unless the activity generates information on previously undocumented "places" of concern, or is used as an opportunity to collect missing or more accurate locational information,*** it may not be an activity requiring collection of LDP data. Activities may be conducted at places that already have adequate locational data. For example, an inspection may occur at a facility for which locational data were collected and fully documented during permit issuance. Therefore, the inspection team does not have to re-collect locational data each time it visits the facility.[4]

[4]See description of "incremental" locational data collection in Chapter 7.

2.2 Data Documentation Requirements

This section defines specific requirements for locational data documentation, as well as specifications for data elements required by the LDP: latitude, longitude, method, accuracy, and description.

2.2.1 Latitude



Latitude can be defined as the distance north or south of the Equator as measured in terms of the 360 degrees of a circle. Each line of latitude is an imaginary circle around the earth, parallel to the Equator, therefore the lines are referred to as "*parallels of latitude.*" The Equator is located at 0 degrees latitude and is the starting point for measuring latitude. The North Pole is located at 90 degrees north latitude and is simply a point, as is the South Pole, at 90 degrees south latitude. Therefore, every other point falls somewhere between 0 and 90 degrees either to the north or south.

One degree of latitude is approximately 111 kilometers or about 69 miles. A degree of latitude can be divided into 60 minutes, and a minute can be divided into 60 seconds.

According to the Locational Data Policy, latitude will be reported in the following format:

+/-DDMMSS.SSSS[5]

where:

- ***DD represents degrees of latitude***; a two-digit decimal number ranging from 00 to 90.
- ***MM represents minutes of latitude***; a two-digit decimal number ranging from 00 through 59.
- ***SS.SSSS represents seconds of latitude***, with a format allowing up to ten-thousandths of a second of specificity.
- **+** ***indicates latitude north*** of the equator.
- **-** ***indicates latitude south*** of the equator.

[5]Ch. 13, Sec. 5c(1), p. 13-3, IRM Policy Manual.

At a minimum, ***values for latitude should always be complete to the second and in accordance with the 25 meter accuracy goal.*** Data systems, however, should be capable of handling latitude data to the full length of the format (i.e., +/-DDMMSS.SSSS) to accommodate more precise measurements likely in the future.

2.2.2 Longitude

Longitude measures distance east and west in terms of the 360 degrees of a circle. The lines of longitude meet at the poles and are referred to as meridians. The prime meridian, internationally accepted as the line of 0 degree longitude, runs through Greenwich, England. All other longitude measurements are reported as east or west of the prime meridian. Half of the world is measured in degrees east of the prime meridian, up to 180 degrees, and half is measured to the west, up to 180 degrees. The 180th meridian is designated as the International Date Line.

Unlike latitude, where the distance between units is always the same, the length of a degree of longitude differs greatly with latitude, or distance from the equator. The closer to the poles, the shorter the length of each degree of longitude, until it reaches zero at the poles. A degree of longitude, like latitude, is divided into 60 minutes, and each minute is divided into 60 seconds.

According to the Locational Data Policy, longitude is to be reported in the following format:

+/-DDDMMSS.SSSS[6]

where:

- ***DDD represents degrees of longitude***; a three-digit decimal number ranging from 000 to 180.
- ***MM represents minutes of longitude***; a two-digit decimal number ranging from 00 through 59.
- ***SS.SSSS represents seconds of longitude,*** with a format allowing up to ten-thousandths of a second of specificity.
- **+ *indicates longitude east*** of the prime meridian.
- **- *indicates longitude west*** of the prime meridian.

[6]Ch. 13, Sec. 5c(1), p. 13-3, IRM Policy Manual.

At a minimum, ***longitude values should always be complete to at least the second, in accordance with the 25 meter accuracy goal.*** Data systems, however, should be capable of handling longitude data to the full length of the format (i.e., +/-DDDMMSS.SSSS) to accommodate the more precise measurements likely in the future.

2.2.3 Method

The Locational Data Policy requires documentation of the method used to determine lat/long coordinates.[7] Locational data accuracy will be technology-driven to attempt to integrate the most sophisticated technological capabilities into EPA's data collection efforts. Until June 1992, data collectors may select any method that is most suitable for meeting Data Quality Objectives (DQOs) set for the effort. By June 1992, however, the Agency will move toward widespread use of Global Positioning Systems (GPS) for collecting new lat/longs when deployment of the full constellation of satellites necessary to use GPS is completed. A primer on the use of GPS is part of this implementation guidance package.

The data element *method* describes the procedure used to determine the latitude and longitude coordinates. In order for "method" to be most meaningful to secondary users, it must be documented so that there is sufficient information to independently reproduce the same locational coordinates. Therefore, ***this data element should include fields for technology (i.e., address matching, map interpolation, Loran-C, or Global Positioning System, etc.), reference datum (i.e., NAD27 or NAD 83) and map scale.***

It is likely that method will be an essential qualifier used to search lists and create subsets of coordinates in automated data bases. For this reason, it is essential to ensure consistency in this data field and establish a definitive list of the valid values for the field.

Standard documentation of "method" is done best by ***representing the method as a code and having qualifying data elements for datum and map scale.*** "Method" codes are presented in Exhibit 2-1.

The data element for *datum* should be in the format ***BB*** where BB is the year of the datum (e.g. 83). The data element for *scale* should be the ***"X" value of the 1:X ratio*** (e.g., if the scale is 1:24,000, the value of the scale data element should be "24,000").

---

[7] *"Specific method used to determine the lat/long coordinates (e.g., remote sensing techniques, map interpolation, cadastral survey)"* Ch. 13, Sec 5c2, p. 13-4, IRM Policy Manual.

Exhibit 2-1

Codes* to Document *Method*

Method	Datum	Spatial Reference (scale of source map or photo)	
		Scale	Data Element Value
Address matching -- **ADDMAT**			
Aerial photography w/ ground control -- **PHOTOGM**	**83** -- NAD 83	7.5' X 7.5' (1:20,000)	**20,000**
Cadastral survey -- **SUR-C**		7.5' X 15' (1: 20,000)	**20,000**
Conversion from state plane coordinate system -- **SPCSCONV**	**27** -- NAD 27	7.5' X 7.5' (1:24,000)	**24,000**
Conversion from township-section-range (etc.) system -- **TSRCONV**		7.5' X 15' (1:24,000)	**24,000**
Conversion from Universal Transverse Mercator (UTM) coordinates -- **UTMCONV**	**00** -- Datum unknown	7.5' X 7.5' (1:25,000)	**25,000**
Digital or manual raw photo extraction -- **PHOTORAW**		7.5' X 15' (1:25,000)	**25,000**
Geodetic quality Global Positioning System (GPS) Survey -- **SUR-GPS**		15' X 15' (1:62,500)	**62,500**
LORAN-C navigation device -- **LORAN-C**		7.5' X 20' (1:63,360)	**63,360**
Map interpolation, via digital source extractor or manual -- **MAP**		7.5' X 36' (1:63,350)	**63,350**
Method unknown -- **UNKNOWN**		1:15,840	**15,840**
Navigation- quality GPS -- **NAV-GPS**		1:20,000	**20,000**
Remote sensing -- **RMTSEN**		1:24,000	**24,000**
ZIP code centroid -- **ZIP**		Not Applicable	**NOT APPLICABLE**
		Unknown	**UNKNOWN**

*** Codes are maintained by EPA's GIS Program within OIRM; codes may be made available within the standards category on GISNET**

2.2.4 Accuracy

The Locational Data Policy requires documentation of the accuracy of the coordinates. The accuracy data element is intended to enable users to determine whether a lat/long measurement is adequate for their applications. ***"Accuracy" is a quantitative measurement of the amount of deviation from true value present in a measurement, (i.e., accuracy describes the correctness of a measurement).*** This term is different from *precision* which is a quantification of the range of variation normally present in a measurement technique (i.e., precision describes the likelihood of the same values being repeated in another measurement).

For example, ten people who read the coordinates of a particular point on a map could be fairly accurate as all ten could obtain coordinates which are very close to the actual coordinates of the point. Because each individual is applying different subjective judgments when reading the map, there could be considerable variation between individual readings resulting in poor precision for the technique. Conversely, a poorly calibrated LORAN receiver used by the same ten people at the same location would likely yield coordinates which were nearly identical, resulting in high precision. The coordinates obtained, however, would not be accurate because they would not be close to the actual location.[8] The difference between precision and accuracy is illustrated in Exhibit 2-2.

The issue of requiring that a particular level of lat/long accuracy has been addressed by the Locational Accuracy Task Force. The Task Force has recommended an accuracy *goal* of 25 meters. This visionary goal is discussed in detail in the next chapter.

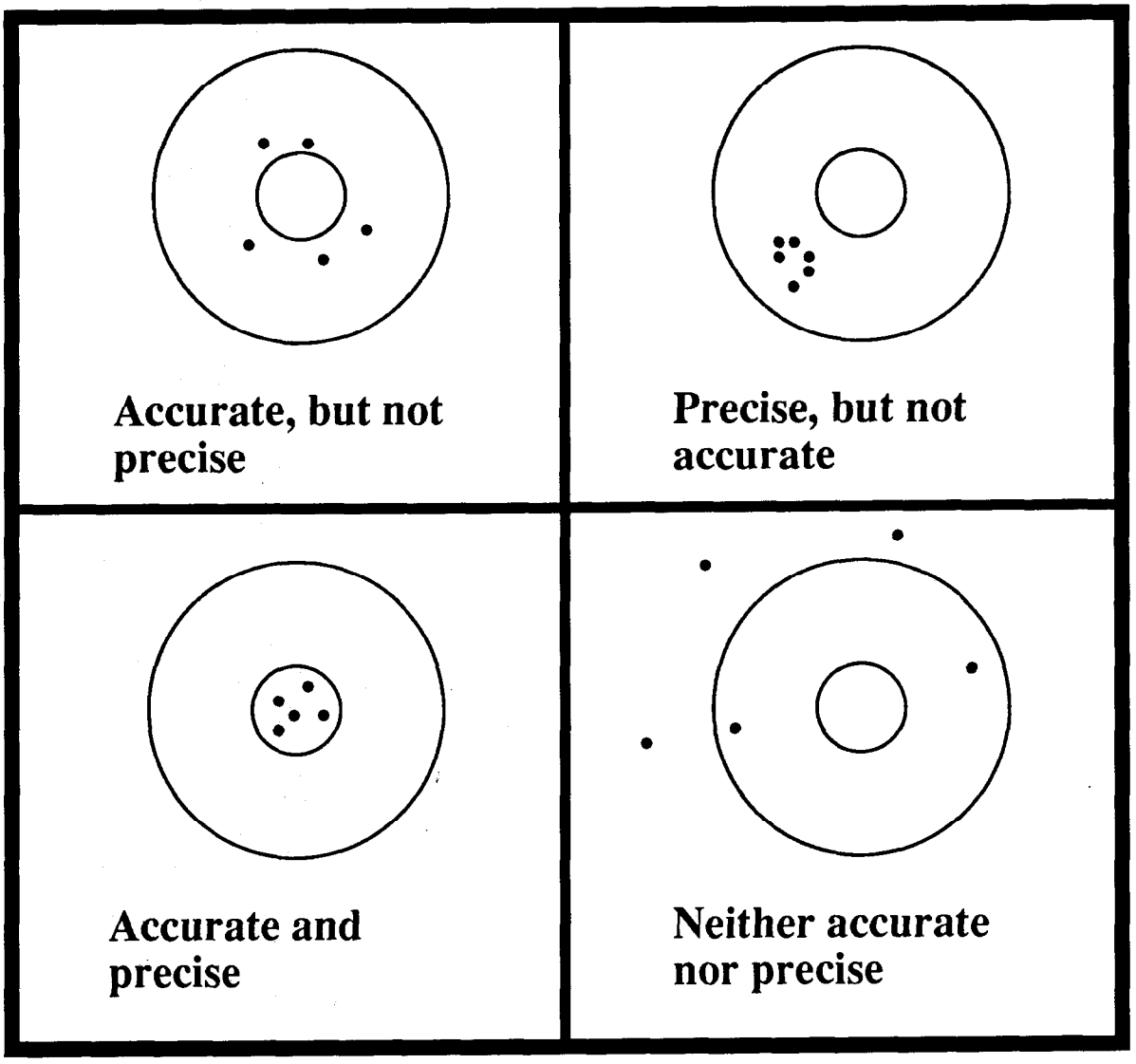
It has been noted that distance represented by a degree of latitude remains constant throughout the world whereas the distance represented by a degree of longitude varies from the poles to the equator. For example, the number of meters on the ground represented by a 1.0 second accuracy for longitude at the equator (0 degrees latitude) would be larger than $\pm$ 1 second accuracy at the poles (90 degrees latitude). Additionally, $\pm$ 1.0 second of accuracy for latitude and for longitude is similar only at the equator. Therefore, to be fully descriptive, coordinate pairs would require two accuracy measurements; one for latitude and one for longitude. Due to the additional burden on data storage, however, ***the LDP requires that only the lowest accuracy measurement be recorded, regardless of whether it is for longitude or for latitude.*** With such an arrangement, the user community will know that both coordinates are <u>at least</u> as accurate as the reported value.

---

[8]"No Minimum Required Accuracy! The Voice of Reason," Task Force on Locational Accuracy, White Paper #2, pp. 3-4.

Exhibit 2-2

Precision vs Accuracy



Accuracy is to be presented as a range within which there is 95% confidence that the true lat/long value falls. The format for presentation of accuracy is:

+/- X units

where units are degrees, minutes, seconds, or decimal fractions of a second.

Accuracy should be presented to one decimal place smaller than the units in which the lat/long coordinates are reported. Therefore, if coordinates are presented in whole-tenths-of-seconds, it is because they have been "rounded up" from some value in hundredths-of-seconds, and the accuracy is described as the range, in hundredths of seconds, within which the true value has a 95% chance of falling. To illustrate, lat/long coordinates reported as:

+432430.3,-1295720.8

which are presented to the whole tenths-of-seconds may have an accuracy of:

+/- .05 sec

In general, to meet the 25 meter goal, accuracy should be determined to within fractions of a second. ***The data collector must identify that value of the units*** (in this case, 5 hundredths of seconds) ***within which there is 95% confidence that the true location falls.*** More information on the 25 meter goal recommended by the Locational Accuracy Task Force is presented in the next chapter.

2.2.5 Description

Lat/long coordinates are often collected to be representative of an entity but are actually of a particular point or portion within the entity. Thus, the "lat/long of a facility" might actually be the lat/long of the entry-point of the property; or of the lat/long of an effluent sampling point at that facility. Because secondary users need to know exactly what the lat/long coordinates define, ***the Locational Data Policy requires a description of the exact place where the coordinates were collected.***

Because the exact locations to which coordinates refer can vary a great deal and are difficult to generalize, the format of the description data element is a free-format, text field. There should be, however, two components documented for "description:"

- ***The exact entity that the lat/long(s) are <u>of</u> (NOT what they represent, such as being <u>of</u> the driveway but <u>representing</u> the whole facility)***
- ***Whether the coordinates describe a point, line, or area.***

Data collectors should be consistent in their use of the description field. The exact place used to represent the location of the entity should be selected when planning the data collection process.

2.3 Recommended Locational Data

Although it is not mandatory under this policy, it is strongly recommended that all data bases containing locational coordinates also contain ***date of collection*** and ***source***. Date of collection is important because it enables secondary users to draw inferences about locational data. For example, if the original locational data were collected prior to the current revision of the base map, then the locational data may be incompatible with other points on the map.

To provide another basis for evaluating the data for re-use, the source of the locational data should also be documented. Source refers to the organization that actually collected the coordinates, (e.g., EPA office, a state environmental organization, a regulated entity reporter, or a specific contractor). This type of data gives secondary users more information by which to evaluate the quality of data because of what might be known about the collector's technique. It also enables the secondary user to contact the original collector with specific questions regarding compatibility.

Because they are not formal parts of the current policy, these elements are considered optional. If collected, they should be stored in additional fields. No specific formats for these fields have been established.

2.4 Special Considerations

Some important guidelines and principles for the collection and documentation of locational data are presented in this section.

2.4.1 "Tiering" and Spatial Extent

A common concern of data collectors is how to comply with the Locational Data Policy in terms of identifying an entity in the appropriate spatial terms.[9] This concern can be exemplified by three questions:

- *My "facility" is a complex entity; what portion of it (e.g., the back door, the driveway, the discharge pipe, the center of the building) is appropriate to document with lat/longs?*
- *My site is extensive; should the coordinates represent a point, line, or area?*

[9] *"...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..."* (Ch. 13, Sec. 5c(1), p. 13-3, IRM Policy Manual)

- *When is an area represented by a point? When by a polygon?*

The LATF has termed the first concern as "***tiering***," in recognition of the fact that an entity of environmental concern can have several "levels" of interest. For example, a site may have *several facilities* on it, each one with *several sources* of air emissions, each with *several points of emission*. Each level of interest on this site can be classified according to a different "tier." ***The tier to which the other attribute data belong (e.g., as monitoring or site description data) is the tier for which location identification data must be collected and documented***. The rationale behind this guideline is that every data collection project will have a specific objective in mind, and locational data will probably never be collected alone (except in some cases of *centralized* locational data collection, as defined in Chapter 7). Therefore, monitoring data that describe the composition of effluent from a particular discharge pipe should be accompanied by locational data identifying the geographic position of the pipe, not necessarily of the whole facility. Locational data to accompany general facility emissions and discharge data, such as required under EPCRA where the activities are described for the whole facility, should be of the facility in general, and not of each point within the facility.

It is likely that ***several sets of coordinates will be collected for a single site*** because many facilities are composed of sub-units, each having a distinct environmental concern. ***Locational data from the different tiers of that facility will be distinguishable from each other by the required DESCRIPTION field***, which must tell precisely to what the lat/longs refer (discussed below).

When the "facility in general" is the proper spatial scope, the data collector must determine which point (e.g., the entry point, northeast corner, etc.) will be used to represent the facility. This determination is made by taking into account *how the data being collected are to be used*. The data collector must determine the single point that is most representative of the entire facility and document it in the "DESCRIPTION" field. The *Guide to Selection of Latitude/Longitude Collection Methods* accompanying this LDPG provides information on using a single point to represent a whole facility. Appendix D of this document presents the definition of a facility as given in the 12/13/91 version of the FIDSIP.[10] In general, ***if a single point is used to represent an entire facility, it should be the most visible and accessible point*** (to ensure precision or "repeatability") and should be identified in the description field.

A similar principle guides determining the most appropriate spatial representation (i.e., point, line, or area) for an entity. In general, this determination is based upon how the data are to be used. For example, if the application of the data collection is national in scale, then a detailed spatial description (e.g., many points defining a polygon outlining the exact

[10]The Facility Identification Data Standard Implementation Plan (EPA/OIRM, 12/91)

boundaries of a site) need not be collected and documented; one set of coordinates to describe the entity location will suffice. If the data are collected to support a focused study (e.g., requiring delineation of exact boundaries of contamination), then the location definition should be reflective of the entity's geographic "shape" (i.e., an "area" or "polygon").

Another basis for determining whether an area should be represented by a point or a detailed polygon is the scale of the map used in the primary application of the data. If a polygon plotted on a map of the appropriate scale for the data's primary use is indistinguishable from a representative point for the same area at the national map accuracy standard[11] (i.e., the polygon will be so small that it appears as a point measured on the map at the scale being used) then only one set of coordinates must be defined to represent the site.

When *one set* of coordinates represents the location of an entity, that one set of coordinates is supported by one set of values for method, accuracy, and description. If *many sets of coordinates are collected to represent a single entity*, the same is true: all of those sets of coordinates are supported by a single value for method, accuracy, and description (i.e., lat/long coordinates, whether collected singly or as a set defining a polygon, are qualified by a single set of method-description-accuracy). This concept is illustrated in Exhibit 2-3.

2.4.2 Standard Datum

To make them most meaningful, locational identification data should be linked to a ***standard datum***. A standard datum is a network of monuments and reference points defining a mathematical surface from which geographic computations can be made. Documentation of the datum used to determine lat/long coordinates helps compensate for the fact that the actual shape of the earth is not a perfect sphere but instead is an oblate spheroid (flattened at the poles and bulging at the Equator). When the Earth is cut along its polar axis the cross-section approximates an ellipsoid.[12]

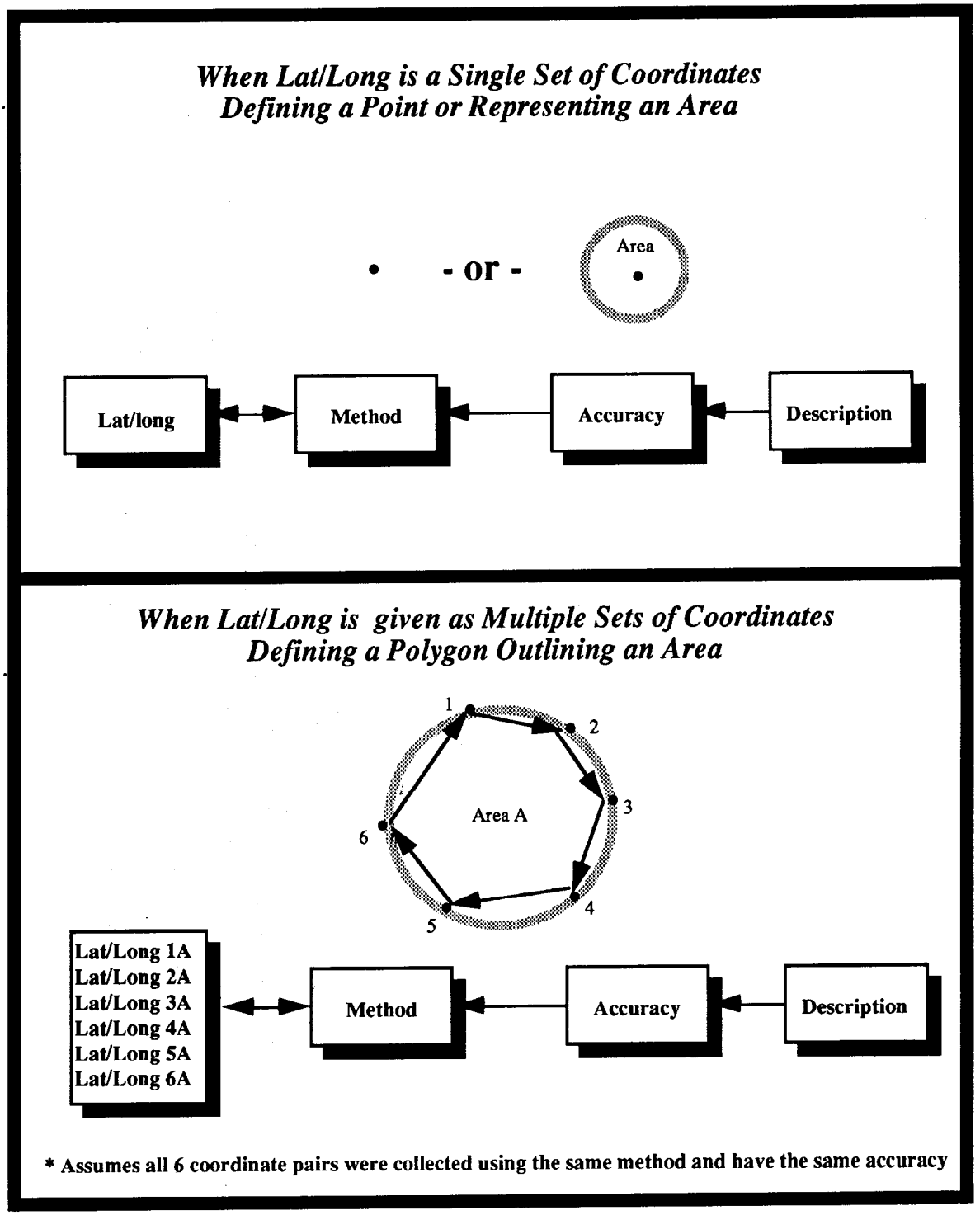
In the past, the *North American Datum of 1927* (NAD27), based upon the Clarke 1866 ellipsoid, has been used as the locational basis within the Agency. Recently, however, the National Geodetic Survey has redefined and readjusted NAD27, resulting in the *North American Datum of 1983* (NAD83). This system is an Earth-centered datum using the Geodetic Reference System 1980 ellipsoid as its basis. The datum shift from NAD27 to NAD83 has resulted in a change of latitude and longitude for all points on a map and a change in the positions of the grid coordinates. Specifically, the 1990 *Federal Digital Cartography Newsletter* estimated that this change caused coordinate shifts in the eastern United States ranging from 3 to 40 meters and in

[11]The true location is within a horizontal distance of 0.02 inches on the map being used.

[12]"Implementing the North American Datum 1983 for the National Mapping Program," United States Department of the Interior.

Exhibit 2-3

Relationship of Method, Accuracy, and Description To Lat/Long Coordinates



California and other parts of the western United States of more than 100 meters.

The National Geodetic Survey Division is encouraging all Federal agencies who use or produce spatial coordinate information to make the transition from NAD27 to NAD83. Even though the majority of EPA's data and most base maps for the United States are in NAD27, ***EPA data collectors should begin using NAD83 at this time if they are not already doing so.***

Conversion of old data should begin as soon as possible. EPA data collectors will be able to use datum conversion software known as *NADCON* (North American Datum Conversion), developed by the National Geodetic Survey. This program, written in FORTRAN 77, allows conversion of large amounts of coordinate data in ASCII format from NAD27 to NAD83. Conversion accuracy of approximately 0.15 meters is achievable within the conterminous United States and slightly lower accuracy levels (not more than 1.0 meter) in more remote areas of the country. EPA has developed software for conversion of locational data that is in ARC/INFO format from NAD27 to NAD83. This conversion software, known as CDATUM, is available from EMSL-Las Vegas (refer to Appendix B for a contact in EPA/ORD). Data conversion options are discussed in greater detail in Chapter 7.

Due to the transition to NAD83, it will be critical for data collectors to have documented the datum utilized in the collection process so that conversion to NAD83 can be made to only those coordinates needing it. Therefore, for each location definition (either a single set of coordinates or several sets representing a line or an area), ***the datum used as the basis will be recorded as part of the METHOD field.*** Complete guidelines on documentation of method are given in Sec. 2.2.3 of this chapter.

Chapter 3

RECOMMENDATIONS OF THE LATF

3. RECOMMENDATIONS OF THE LATF

Chapter 1 of this document introduced the *Locational Accuracy Task Force* (LATF), a group of senior management and technical professionals from EPA and state environmental agencies convened to develop recommendations on locational data accuracy for EPA's IRM Steering Committee. Over a period of six months, the LATF collected and weighed a considerable amount of information on geocoding technologies and programmatic requirements. Their efforts resulted in a series of recommendations which have been incorporated into the Locational Data Policy. All EPA personnel and personnel of its agents or state program delegates must address the LATF recommendations in order to comply with the LDP. Details of the LATF recommendations are summarized in the sections that follow and are presented in full in Appendix A.

3.1 Introduction

The consensus of the LATF was that accurate, well documented locational data are essential to risk management and multi-media decision making. They recognized that, even though some EPA programs, regions, states, and other Federal agencies have taken significant steps to develop locational policies, unless a clear goal is stated, the Agency's data collections will remain a "mixed bag." The group recommended a technology-based goal, centered around Global Positioning Systems (GPS) and photo interpretation/map interpolation to yield a goal of 25 meter accuracy for all coordinate points.

It is important to note that all EPA programs except RCRA and CERCLA already have regulatory requirements for collecting latitude and longitude as descriptors of location (OSWER has non-regulatory program directives for RCRA and CERCLA). These programs have a variety of accuracy specifications, ranging from 1 to 15 seconds (approximately 25 to 375 meters). With full implementation of the Locational Data Policy, EPA is striving to have all its lat/long information documented and at least as accurate as 25 meters.

3.2 Summary of the LATF Recommendations

The LATF recommended five general goals for implementing a locational standard as follows:

- Establish an Agency-wide goal of better than 25 meter accuracy for locational data by 1995.
- Encourage the use of the most practicable and accurate technology for determining lat/long coordinates, currently recommended to be GPS technology (with interim use of map interpolation between now and 1995).

- Institute of a process for administering waivers from LDP requirements for certain data collection efforts.
- Develop an incentive program to move the entire environmental community, including states and industry, toward the 25-meter locational accuracy goal.
- Enhance FINDS to be a source of general locational data for regulated facilities.

All of these recommendations were made to initiate a program whereby the quality of EPA's locational coordinate data continually improves. New data are of highest priority for compliance under the 25-meter accuracy goal. Beginning immediately, any collection effort during which new locational data are gathered (e.g., during site visits, permit renewals, or special studies) should yield data accurate to at least 25 meters. ***By 1995, all lat/long measurements in Agency data collections should have accuracies of better than 25 meters.*** Therefore, by 1995 all locational data will be collected in a consistent method, will be of documented quality, and will thus be more valuable in any application.

3.3 Specifics of the LATF Recommendations

The five components of the LATF recommendations are discussed in more detail below and are presented in full in Appendix A.

3.3.1 Accuracy Goal

The Agency-wide goal for all lat/long data is accuracy of better than 25 meters. A plan for reaching this goal will be initiated immediately to phase in the necessary procedural and system changes.

The goal was established at 25 meters because:

- Review of program requirements shows that ***EPA data will be more amenable to secondary use*** if the quality of its location identification data is improved to 25 meter accuracy.
- Examination of locational data accuracy requirements currently in EPA programs and in other agencies ***shows a target of 25 meters to be consistent.***
- Accuracy to the 25 meter level is becoming ***more achievable with improvements in technology.***

The accuracy limit was established as a ***goal*** and ***not a standard*** for several reasons. First, achievement of maximum locational data accuracy is necessarily technology-based (i.e., the quality of locational data should be as

good as the most practicable data collection technology). This technology constraint currently may limit the accuracy of locational data to 25-100 meters but will improve with time to be better than 25 meters (see the discussion of GPS technology below). Second, the complexity of the processes that must be put in place to achieve this level of accuracy makes the concept difficult to enforce, and participation-encouraging incentives were deemed more effective than enforcement mechanisms.

The target date for full implementation of the locational data accuracy goal is December 31, 1995. By this date, global positioning system technology will be mature, and enough time will have passed for updating to EPA data systems to accommodate the new, accurate locational data.

3.3.2 Technology-Based Locational Data Accuracy

The technology to produce highly accurate locational coordinates is improving rapidly. Techniques for map digitization, address matching, and global positioning are becoming more feasible every day. Therefore, the LATF recommendation is to have the best available technology applied to collection of locational data.

GPS technology has been determined to be an effective way of producing accurate locational data. When the constellation of satellites upon which this technology depends is fully deployed in 1992, means for collecting accurate locational data will be fully available to EPA and its partners. At that time, accuracies of 10 meters or better will be achievable with a high degree of confidence and precision. The use of GPS in all data collection efforts after June 1992 will ensure that:

- The most accurate locational data are obtained.
- Equipment purchases and subsequent training requirements are minimized.
- The type of information being reported is uniform.

Several issues that must be resolved prior to full utilization of GPS technology have been identified as:

- The requirements for the amount and type of global positioning devices needed.
- The absence of documented procedures for using GPS technology and the mechanisms for transmission of locational data from the field to the data bases.

- The amount of training and software necessary to properly apply GPS technology.
- The difference in approaches needed for arriving at a conclusion about locational data obtained in the field as opposed to in the office.
- The need to standardize a representative part of a facility for which locational data will be collected.

In addition, GPS technology may not prove to be the most suitable data collection technique in all cases. Selective availability of satellite signals may degrade the quality of results, weather or topographic conditions might influence the readings, and the availability of trained personnel or equipment might compromise accuracy. Therefore, although GPS is the preferred technology for acquiring location identification data, it is not the only acceptable method.

3.3.3 Waiver Applications

In some instances, it may not be appropriate for a data collection effort to meet the requirements of the LDP. Impediments to meeting policy requirements and regulatory restrictions, such as the type of location identification data required, its accuracy, the method used, or the schedule for implementation may be difficult to overcome. A process must therefore be developed for obtaining ***waivers to the LDP.*** This process must ensure uniform consideration of waiver requests without compromising the objectives of the LDP.

Program managers must prepare waiver applications to be presented to the IRM Steering Committee. The waiver applications need be nothing more than clearly defined Data Quality Objectives (DQO) reports. For each data collection effort, the waiver applications should present the following:

- The ***contrast in objectives*** between the data collection effort and the Agency-wide locational data accuracy goal.
- The ***approach that will be used*** to meet the locational accuracy needs of the data collection effort.
- ***Justification for not meeting LDP requirements.***

3.3.4 Incentives

Strict enforcement of a level of locational data accuracy is infeasible because of the wide range of participants (states, industry, etc.) and the different accuracy levels that can be achieved by each type of technology. ***Motivational***

incentives are much more effective than enforcement. The need to share data and the benefits of reduced data collection costs should be a sufficiently powerful incentive. Appendix A details some incentives that were identified by the LATF.

3.3.5 Enhancements to FINDS

EPA/OIRM's Facility INDex System (FINDS) is ideally positioned to be a gateway to all facility data. FINDS is currently undergoing an enhancement effort to accommodate changes that have occurred since its original design, such as the adoption of a Facility Identification Data Standard prompting the connection of many new data bases to FINDS.

FINDS contains basic identification information on every facility regulated by EPA (state use of FINDS is beginning on a pilot basis). In addition, each facility record in FINDS contains references to sources of more detailed program data. The current version of FINDS can house one set of locational coordinates per facility record.

Enhancements can be made to FINDS to make it a more comprehensive source of facility locational data. Such enhancements will reduce the need for collection of new locational data and serve as a way to double-check the quality of program locational data. ***The recommended enhancements to FINDS are:***

- Employment of address matching and other techniques to ***populate/improve the lat/long coordinate data elements in FINDS.***
- Inclusion of accurate ***lat/long coordinates for sub-facility tiers*** such as discharge pipes or stacks.
- ***Edit capabilities to detect changes*** in locational data within program systems.
- ***"Gateway" capabilities*** to directly link program source data bases.
- ***Structured Query Language (SQL) capabilities*** to enhance access to FINDS locational data.

A requirements analysis and DQO report may be necessary to detail the exact specifications and implications of each of these recommendations. In addition, an effort to make source data bases compatible with FINDS may have to be initiated.

Chapter 4

RESPONSIBILITY AND AUTHORITY

4. RESPONSIBILITY and AUTHORITY

Successful implementation of the Locational Data Policy will require the cooperation of many affected organizations. In this chapter, the specific authorities and responsibilities under the LDP of EPA's IRM Steering Committee, OIRM, other EPA offices and programs, the state delegates of Federal environmental laws, and other designated representatives will be discussed.

4.1 The IRM Steering Committee

The IRM Steering Committee, made up of individuals drawn from OIRM, EPA Regional and Program offices, as well as state governments, oversees activities and resolves issues relevant to environmental information management. For example, a subcommittee of the IRM Steering Committee, the LATF, developed a set of recommendations for establishing a goal for locational data accuracy which has been incorporated into the original LDP.

Waivers may be necessary for certain aspects of the LDP. To institute a waiver evaluation process, the IRM Steering Committee will:

- Design a procedure for waiver processing.
- Develop and publish waiver evaluation criteria.
- Define roles and responsibilities in the waiver process.
- Establish monitoring and enforcement mechanisms for LDP requirements.

Support will be provided to the IRM Steering Committee by OIRM/IMSD, various inter-organizational committees, and program managers (see below).

4.2 OIRM Responsibility and Authority

The Locational Data Policy describes OIRM's responsibilities in a general sense. The LDP states that *"OIRM shall:*

- *Be responsible for implementing and supporting this Policy.*
- *Provide guidance and technical assistance where feasible and appropriate in implementing and improving the requirements of this Policy."*

Responsibilities within OIRM for implementing the Locational Data Policy lie primarily within two OIRM divisions: *the Information Management and Services Division* (IMSD) and the *Program Systems Division* (PSD). Additionally, implementation guidance to managers of scientifically-oriented data systems can be

provided by OIRM's Scientific System Staff (SSS). There are several programs within theses groups that will have a key role or involvement with the LDP and include the Geographic Information Systems Program (GIS) and the SEDM Program.

"Tools" that OIRM will use to help implement the LDP include:

- *The IRM Policy Manual.*
- *The Catalog of Agency Data Policies and Standards.*
- *Standard EPA forms.*
- *FINDS.*
- *The Gateway.*
- *The GIS network.*
- *The IRM Steering Committee.*
- *Networks of senior IRM officials, Regional IRM Branch Chiefs, system managers, and SEDM Coordinators.*
- *Grants (e.g., SEDM).*
- *Cooperative agreements and memoranda of understanding.*
- *The LDP Guidance Documents.*

The following sections will define each of these OIRM responsibilities by division within OIRM. The current organizational structure of OIRM is displayed in Exhibit 4-1. Divisional responsibilities under the LDP are summarized in Exhibit 4-2.

4.2.1 IMSD

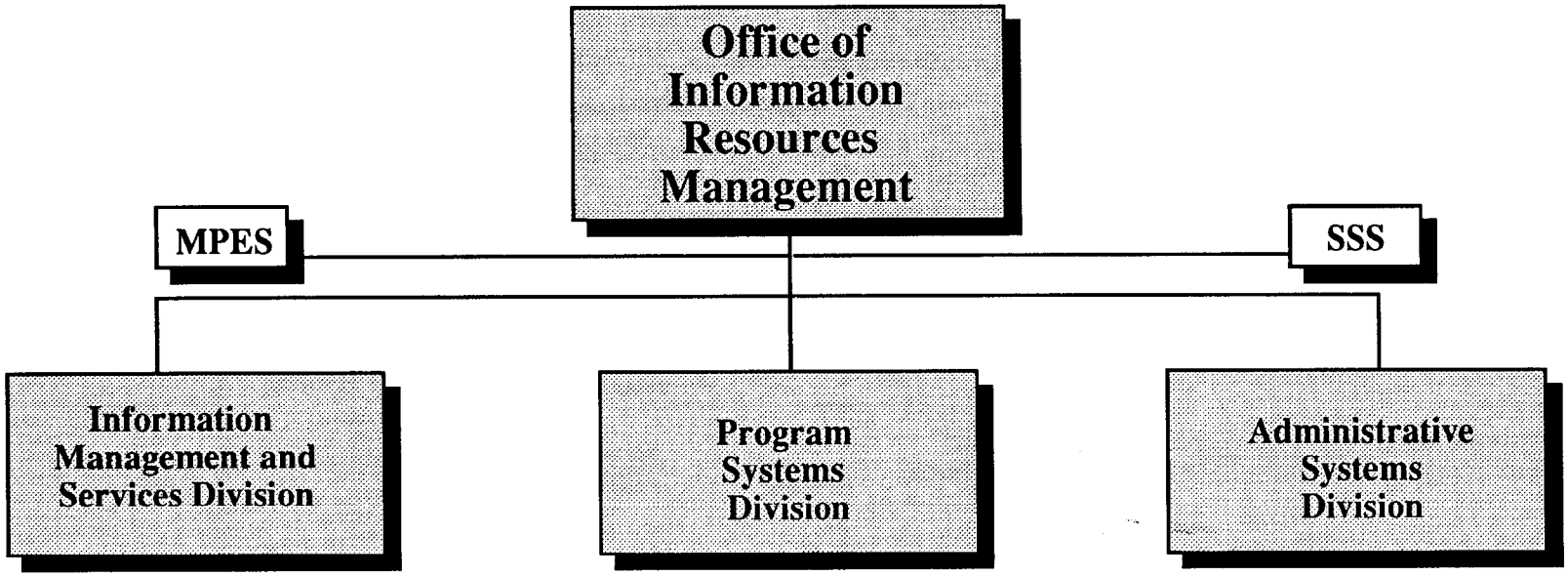
Two branches within IMSD, the Information Management Branch (IMB) and the Information Sharing Branch (ISB), will have responsibilities in LDP implementation. ***IMB has played a lead role in supporting the LATF and in developing the LDP and this implementation guidance.*** IMB also is responsible for the Agency **forms program**, which may be used to enhance standardization of locational data within EPA. IMB will continue to play a role in LDP implementation, including initiating program-specific planning processes, overseeing LDP progress, and coordinating development of information management standards, and policies associated with the LDP.

ISB is responsible for assisting and coordinating the implementation of this policy at the state level under the auspices of the ***State/EPA Data***

Exhibit 4-1

OIRM Organizational Chart

(1/92)



- Information Management Branch
- Information Access Branch
- Information Sharing Branch

- Systems Planning and Analysis Branch
- Systems Maintenance and Development Branch
- Information Integration and Management Branch
- Client Support Branch
- Technology Branch

- Technical Support and Development Staff
- Administrative Systems Support Branch
- Administrative Systems Development Branch

Exhibit 4-2

OIRM Responsibilities under the Locational Data Policy

IMSD Responsibilities	PSD Responsibilities
✔ Produce Agency-wide guidance for LDP implementation	✔ Provide technical leadership for LDP implementation through the GIS program
✔ Coordinate data standards and policies to assure consistency with LDP	✔ Promote locationally-based data integration through efforts such as the *Gateway*
✔ Conduct reviews of LDP implementation progress to assess effectiveness	✔ Ensure that EPA spatial and locational data policies are compatible with Federal government standards and policies by participation in the FGDC
✔ Provide assistance in development of Program or Environmental Initiative LDP Implementation Plans	✔ Provide access or reference to locational data via FINDS
✔ Aid states in identifying grant funds and available technical resources and identify data sharing and technology transfer incentives for non-EPA participants through State/EPA Data Management Program (SEDM)	✔ Ensure STORET compliance with LDP
✔ Provide guidance on form and procedure modifications	

Management (SEDM) Program. ISB will foster cooperation between states and EPA during implementation of the LDP. Mechanisms that may be used to encourage participation include technology transfer and application of financial assistance grants for data integration. Conforming to international standards, the Agency's approach to collecting and documenting locational data will be shared with other countries via ISB's International Data Sharing Program.

4.2.2 PSD

PSD provides general support and guidance to most EPA system development efforts. In this capacity, ***PSD is responsible for ensuring that LDP requirements are addressed in every system development effort.*** Most of PSD's responsibilities in the implementation of the LDP lie within the Information Integration and Management Branch (IIMB). ***The GIS Program in IIMB will have a lead technical role in LDP implementation.*** This program will be the focal point for spatial information resources, such as GRIDS, GIS policy, and national GIS activities. IIMB will encourage locational data sharing and exchange with all of its "clients" and contacts through its data integration function, including development of the EPA Gateway. Additionally, IIMB will continue to be responsible for ensuring that EPA is meeting Federal government standards for spatial data, as established by FGDC.

PSD's Systems Maintenance and Development Branch (SMDB) is responsible for ***upkeep of the FINDS System.*** SMDB will provide general lat/longs (when available) for each facility in the "master record" file, and *can* capture lat/longs from program systems for facility sub-portions in the FINDS "alias file." SMDB also will ensure STORET compliance with LDP requirements.

4.2.3 Resources from OIRM

Implementation of this Policy will require a partnership between OIRM, the programs, and other key participants (e.g. states, the regulated community). OIRM will work with managers in media programs and environmental initiatives to determine resource requirements and appropriate sources of assistance.[1] States may work through their SEDM and EPA program liaisons for possible assistance.

LDP participants are expected to extend their data collection resources to ensure availability of the locational data necessary to meet the requirements of the LDP. However, ***OIRM will provide certain resources that may reduce the financial burden.*** These resources include:

---

[1] A methodology for determining costs for LDP implementation can be found in the *Guide to Selecting Latitude/Longitude Collection Methods*, accompanying this LDPG.

- Guidance documents.
- LDP implementation plan development workshops.
- Data conversion assistance (software, technical assistance).
- System development guidelines, including available software for edit checks to assure locational data accuracy, completeness, and consistency (e.g., comparison of coordinates to county boundaries).
- Possible funding for systems enhancements through the System Development Center (SDC).
- A central repository or access capability to existing locational data for regulated facilities through the Facility INDex System (FINDS).
- News bulletins on developments which may affect implementation of the LDP or technology transfer ("GISNET").
- A program for data sharing, financial assistance, and technology transfer with states through the SEDM team including potential funding through SEDM's financial assistance program.

When substantial improvements to a system are required to implement the Locational Data Policy, and the requestor can demonstrate <u>not only mission criticality but broad benefits to other environmental activities</u> (i.e., "secondary usability" of data enhanced by more comprehensive locational data), this system will be given priority for resource allocation by OIRM.

4.2.4 Authority

OIRM's authority in implementing the LDP stems from Agency-wide concurrence with the LDP at the Assistant Administrator level. Additional authority is derived from the following sources:

- ***15 CFR, Part 6 Subtitle A, Standardization of Data Elements and Representations*** -- Establishes a Federal program for standardizing data elements and representations that are used and interchanged in government data systems, expanding the utility of every data collection and avoiding duplication of effort.
- ***Geological Survey Circular 878-B, A U.S. Geological Survey Data Standard, Specifications for Representation of Geographic Point Locations for Information Interchange*** -- Recommends formats for documenting location identification data.

- ***Federal Inter-Agency Coordinating Committee on Digital Cartography (FICCDC)/U.S. Office of Management and Budget, Digital Cartographic Data Standards: An Interim Proposed Standard*** -- Provides recommendations for a systematic and comprehensive set of digital representations of cartographic features to support various cartographic data systems and spatial data transfer.

- ***EPA Regulations 40 CFR 30.503 and 40 CFR 31.45, Quality Assurance Practices under EPA's General Grant Regulations*** -- Establish requirements for grantees involved with environmentally-related measurements or data generation to develop and implement quality assurance practices for producing adequate data meeting project objectives, and minimizing loss of data due to uncontrollable conditions or malfunctions.

- ***"Policy and Program Requirements to Implement the Quality Assurance Program" (EPA Order 5360.1)*** -- Requires users to specify data quality needs and the quality control necessary to assure that the resulting data satisfy intended uses. This order also requires thorough, verifiable, and defensible documentation so that data quality is known.

- ***Locational Accuracy Task Force Findings and Recommendations (December 13, 1990)*** -- Establishes the goal of 25 meter accuracy for lat/long to be implemented by 1995 and identifies the steps needed to achieve that goal.

- ***OMB Circular A-16 (revised 10/19/90)*** -- Establishes coordination procedures and assigns responsibilities to government agencies involved with spatial data analysis.

OIRM intends to monitor the effectiveness of the LDP to ensure that it is being properly implemented by all data generators and that it is meeting the goal of improving secondary data usability. Data collection efforts will be reviewed to assess whether documentation of lat/long data and corresponding attributes (method, description, and accuracy) are included.

4.3 Other EPA Offices and Programs

Because the LDP promotes consistency across the entire Agency, all offices and programs that collect locational data will have responsibilities in implementing it. These offices/programs and their responsibilities are summarized in Exhibit 4-3. More complete details are provided in the sections that follow.

Exhibit 4-3

EPA Office and Program Responsibilities under the Locational Data Policy

Non-Media Program Office Responsibilities

✔ **Office of Research and Development**

- Evaluate LDP data collection methods for cost and accuracy
- Conduct technical research for GIS/GPS and spatial data application and collection (including publication of the GPS Primer)
- Implement DQO Program as it relates to LDP
- Implement LDP in EMAP

✔ **Office of Policy, Planning and Evaluation**

- Aid programs in meeting LDP requirements when collecting data from the public
- Interact with OMB on behalf of programs for new information collection requirements

✔ **Regional Offices**

- Regional IRM organizations will serve as local sources for information, guidance, and monitoring during LDP implementation
- State/EPA Data Management Coordinators will work with other regional staff to facilitate implementation
- Environmental Services groups will provide technical assistance and guidance on locational data collection efforts

✔ **NDPD/OARM**

- Analyze need for Agency-wide vehicle for procuring GPS equipment
- Promote LDP compliance throughout client community

✔ **Regional GIS/GPS Work Groups**

- Promote adherence to LDP principles in activities involving spatial data

✔ **Environmental Initiatives**

- Ensure LDP is planned and fulfilled.

Media Program Office Responsibilities

✔ Develop Program LDP Implementation Plans

✔ Develop or redirect existing data collection and documentation policies and procedures to meet the requirements of the LDP

✔ Implement changes to documentation tools, such as forms, grant documentation, contracts, etc., including instructions on locational data collection

✔ Arrange for adequate resources to implement LDP within the program

✔ Address LDP requirements in planning processes

✔ Coordinate with OIRM to establish priorities for policy implementation

✔ Provide support for policy implementation throughout program with personnel, agents, grantees, delegates, and the regulated community

✔ Develop DQOs in support of data collection planning and waiver requests

4.3.1 Non-Media Program Office Responsibilities

Responsibilities for LDP implementation reside with several offices whose activities affect or involve media programs, but which are not media programs themselves.

Office of Research and Development

The Office of Research and Development (ORD) has several critical functions in effective implementation of the Locational Data Policy. ***ORD responsibilities include***:

- ***Methods evaluation*** -- ORD takes the lead in evaluating all measurement methods in EPA, including those for locational data. ORD will be responsible for reviewing existing and new technologies for collecting locational data and making recommendations on the costs, accuracy, benefits, and limitations of each method.

- ***GIS development activities*** -- ORD's EMSL facility in Las Vegas, which is a Center of Excellence, and its EPIC facility in Warrenton, Virginia, are responsible for conducting technical research in the effective use of GIS technology and spatial data application and collection.

- ***GPS technology expertise*** -- Because the locational accuracy goal of 25 meters is technology based, and because GPS is the recommended locational data collection method, ORD will be the Agency's contact for the application of GPS technology. ORD produced the GPS Primer accompanying this document and ORD will continue to assess technological developments and monitor the best methods for achieving the Agency's locational data accuracy goal.

- ***Data Quality Objectives (DQOs)*** -- ORD is responsible for assuring that the objectives of the LDP are addressed during the data-collection planning process, including DQO development. DQOs will be a key part of the policy waiver process. ORD offers guidance and technical assistance to all decision-makers and data collectors in the process of the DQO development.

- ***Environmental Management and Assessment Program (EMAP)*** -- ORD is responsible for developing and implementing EMAP and will ensure that LDP requirements are met in this program.

Office of Policy, Planning and Evaluation (OPPE)

Among other responsibilities, OPPE assures that EPA is in compliance with certain portions of the Paperwork Reduction Act. Toward that end, OPPE administers the Agency's information collection budget and works to facilitate OMB approval of information collection requests when gathering data from the public. Therefore, ***OPPE will be involved with implementing the LDP when data to satisfy LDP requirements must be collected from the public (i.e., regulated community).***

In particular, OPPE will provide guidance and assistance to all programs that depend on the regulated community to provide some or all of their locational data. Activities may involve ***creation or modification of forms used for data collection,*** as well as support in developing justifications to the Office of Management and Budget (OMB) for increases in information collection burdens. ***Where appropriate, OPPE will interact with OMB on behalf of the programs to help justify collection of information necessary to comply with the LDP.***

In its policy review capacity, ***OPPE will be responsible for developing and/or reviewing rules, regulations, policies, and guidelines to ensure compliance with the LDP.*** OPPE will notify and work with OIRM to promote modifications to these items as necessary to ensure consistency with the LDP.

Regional Offices

Most environmental programs are conducted through EPA's regional offices, either directly by regional staff or through delegation to states in the region. Each regional office has several organizations that will play a key role in LDP implementation. These organizations include:

- ***Regional IRM organizations*** -- The regional IRM organization will be responsible for working with regional program staff to promote LDP compliance during all appropriate activities, such as inspections, site cleanup management, enforcement actions, permit issuance, or compliance monitoring. The regional IRM organizations will be the local source for guidance and information on policy requirements and resource access, similar to OIRM within EPA/HQ. Regional IRM personnel also will promote awareness of the LDP at a regional level so that complete and accurate LDP data are entered into appropriate computer data bases in a timely manner.

- *SEDM Coordinators* -- SEDM coordinators will work with states to facilitate LDP implementation by reviewing plans and grant documents produced pursuant to the SEDM Program and by reviewing data collection activities involving the states. SEDM coordinators also will encourage states to adhere to LDP requirements and promote environmental data integration.

- *Environmental Services Departments* -- These departments provide technical services (e.g., laboratory analyses, field personnel, etc.) to media programs within the region, and frequently perform data collection or contribute to planning activities associated with data collection. The Environmental Services organization will provide technical guidance and/or assistance to regional media program personnel in complying with the LDP, including use of GPS. Environmental Services personnel will assure that locational data requirements are addressed and met by all spatial data-generating activities.

OARM/National Data Processing Division (NDPD)

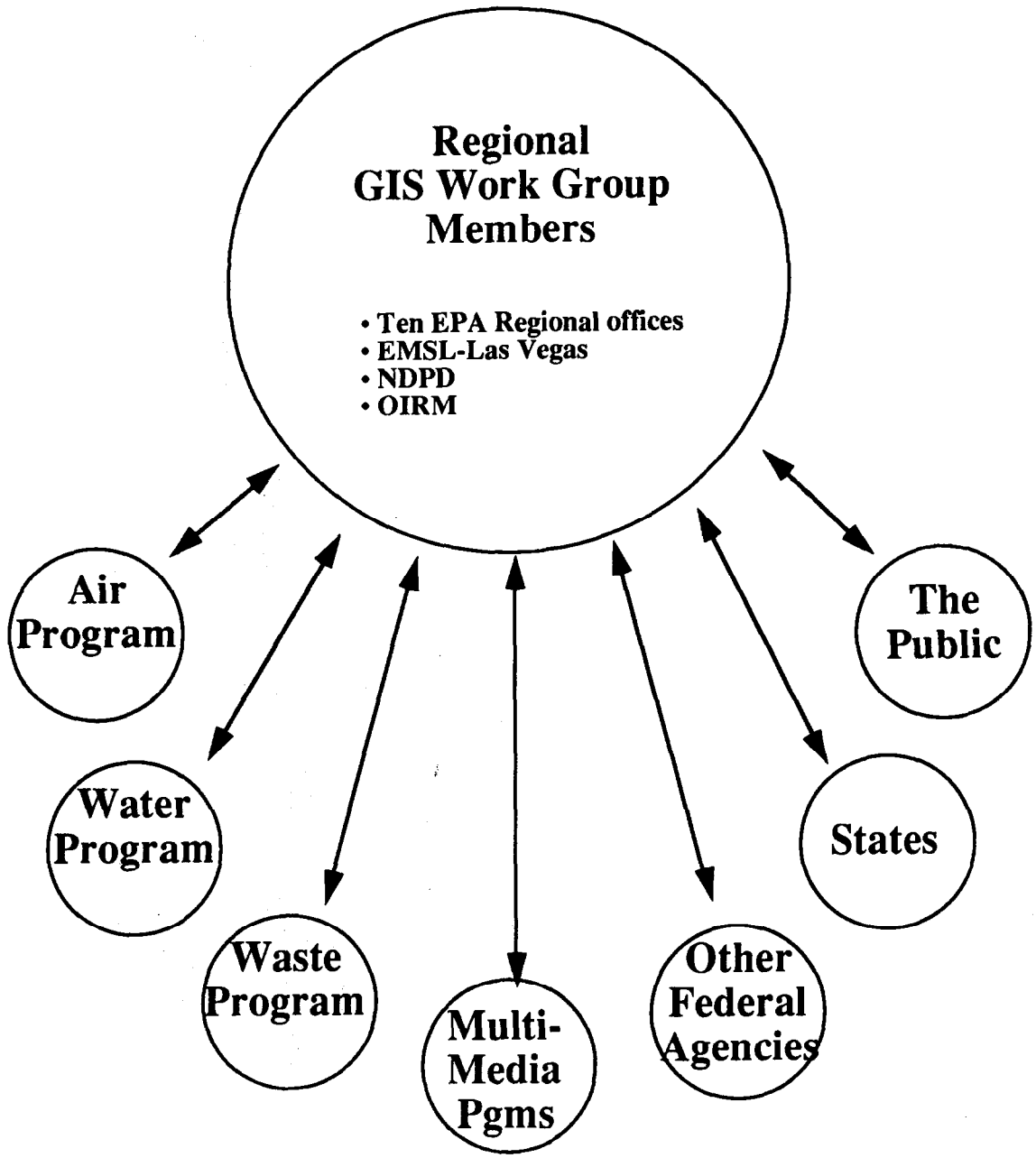
NDPD is responsible for overseeing the management and acquisition of computing and other technological resources for EPA. The endorsement of global positioning system technology to meet the goal of 25 meter locational data accuracy may necessitate the development of an Agency-wide plan to acquire certain equipment. ***NDPD will be responsible for examining and documenting requirements for a vehicle to obtain the necessary equipment, hardware, and software for full and consistent use of global positioning system technology.***

The Regional GIS Work Group

The *Regional GIS Work Group* is an organization of EPA professionals convened to address issues concerning the application of spatial analysis technology to environmental management. Members of the Work Group include representatives from the ten EPA regional offices, EMSL-Las Vegas, NDPD/OARM, and the Headquarters National GIS Program. The interaction between the regional GIS Work Group and the GIS user community is displayed in Exhibit 4-4. The purpose of the Regional GIS Work Group is to establish a network to identify regional-level issues to be addressed regarding environmental data and spatial analysis; to exchange ideas, progress, and experience; and to recommend activities that would increase the value of spatial data to all environmental protection participants in EPA, other Federal agencies, and states.

Exhibit 4-4

Interaction between the Regional GIS Work Group and the GIS User Community



The Regional GIS Work Group has developed a five-part Strategic Plan to ensure an effective and efficient GIS program and to promote spatial data acquisition/application capabilities for all end users. Each of the five components of the Strategic Plan is presented in Exhibit 4-5 along with associated priority goals and objectives. ***A high priority of the Regional GIS Work Group is to provide strategic direction and leadership for implementing the Locational Data Policy at the regional level.***

Regional GPS Work Group

Another inter-regional group that will have an increasingly important role in LDP implementation is the *Regional GPS Work Group,* formed in conjunction with the Regional *GIS* Work Group. The purpose of the GPS Work Group is to create a mechanism within EPA to identify, present, and discuss GPS issues.

The GPS Work Group has performed many activities since its inception in 1990. It has established a communication network for Agency personnel interested in GPS technology through Agency electronic mail and monthly conference calls. In addition, the Work Group nominated GPS coordinators for all 10 regional offices to act as regional contacts on GPS issues and to be responsible for regional cross-program GPS implementation and coordination of GPS activities.

The Regional GPS Work Group has encouraged nomination of Agency membership on the FGCC to allow EPA participation in inter-governmental discussions of many GPS issues. The information gathered from the FGCC is disseminated to the Work Group and to the regional GPS community through the nominated coordinators. The Regional GPS Work Group will provide valuable contacts to regional and state LDP participants. Their role will become increasingly valuable with the phase-in of GPS use.

Environmental Initiatives

Many interdisciplinary, inter-regional *environmental initiatives* have begun in EPA. These initiatives are large, complex programs targeting a specific environmental problem and approaches to solving it. Examples of such environmental initiatives include the Great Lakes National Program, the Chesapeake Bay Clean Up Program, and the Environmental Mapping and Assessment Program (EMAP).

These initiatives do not fall under the jurisdiction of a single media program office or regional office but, instead, involve the coordination of participants from many organizations including EPA regions, EPA media programs, state agencies, other Federal Agencies, and often

Exhibit 4-5

Components, Goals, and Objectives of the Regional GIS Work Group Strategic Plan

Component	*Provide effective management of Regional GIS Programs*	*Develop effective spatial analysis and data integration tools*	*Enhance information exchange and technical transfer*	*Provide effective coordination with states, other Federal agencies, and the public*	*Provide leadership to enhance spatial data quality*
Priority Goals	Identify, acquire, and manage priority data bases necessary to effectively meet the Agency's spatial analysis needs	Develop spatial analysis and data integration tools with broad application to Agency decision-making processes.	Establish and maintain effective mechanisms for technical transfer of GIS applications among all government GIS users.	Strengthen the State/EPA partnership for spatial data management	Acquire accurate and precise locational data for facilities and environmental entities of interest
Specific Objectives	Prepare and implement an Agency-wide spatial data plan to acquire and develop priority spatial databases. Improve the Agency's spatial data management infrastructure.	Focus spatial analysis and data integration tool development efforts to support Region-wide and Agency-wide decision processes. Promote and encourage effective documentation and dissemination of spatial analysis and data integration tools with broad application within the regions and the Agency.	Enhance and support opportunities for information exchange within the Agency, and between the Agency the states. Provide effective mechanisms to maintain and disseminate spatial analysis and data integration tools with wide application in the Agency. Develop and implement Agency-wide standards and conventions for spatial data development, exchange, and display. Enhance telecommunications infrastructure to facilitate communication between EPA and states.	Coordiante closely with the SEDM Program, and support its expansion to enhance integrated information needs. Emphasize media program and other grant processes to facilitate the development of effective, self-sustaining state GIS and data integration programs. Establish and maintain effective communication to enable direct coordination with state spatial data programs. Develop mutually compatible state and regional GIS programs, including hardware, software, standards, and data bases.	Establish operational GPS capability within the regional offices. Provide leadership and guidance for the implementation of the LDP. Establish locational data standards and practices as part of each national data system. Establish FINDS as the Agency's principal repository for facility/source locational data. Focus efforts to improve the quality of existing locational data.

times international governments. ***If the activities of these environmental initiatives include generation of locationally-based data (e.g., monitoring), those activities must comply with the requirements of the LDP,*** including the collection and documentation of lat/long, method, accuracy, and description. Further, they must encourage use of GPS and an accuracy goal of 25m accuracy for the resulting locational data. If the activities under these initiatives involve compilation of data from existing sources, ***initiative managers must accept only locational data from other programs that conforms to LDP requirements.*** In addition, ***implementation of these initiatives should begin with the development of a LDP implementation plan, similar to those prepared by media program offices,*** to ensure that LDP requirements are planned for upon project initiation.

4.3.2 Media Program Offices

All managers of media programs within EPA play a crucial role in implementing the Locational Data Policy. Full execution of this policy will require changes in practices and routines for almost all program data collection activities and will affect field personnel, program managers, and data managers. It is through their combined efforts that EPA will realize the substantial benefits of wide-scale LDP implementation.

Media Program responsibilities for LDP implementation include:

- ***Production of Program LDP Implementation Plans*** -- Every program will be responsible for laying out its plan to implement the LDP, detailing all activities required to assure policy requirements are met by all program participants, including states other members of the regulated community (Chapter 6).

- ***Development or redirection of existing data collection and documentation policies and procedures to meet the requirements of the LDP*** -- Standard procedures by which data supporting a program's mission are generated may require modification in order to assure LDP compliance. Program data managers will be responsible for identifying procedures which need to be modified to conform to the LDP and for implementing necessary changes.

- ***Implementation of changes to documentation tools, such as survey forms, grant documentation, contracts, etc., to ensure timely institution of the LDP*** -- Programs employ a variety of tools by which information is collected, including routine survey forms, electronic data submissions (on-line data access to EPA data bases or submissions on tape or diskette), application forms, notification forms, etc. Program managers will be

responsible for identifying the tools that require modification for implementing those changes.

- ***Dissemination of instructions*** -- Program offices are responsible for ensuring that adequate instructional material is included in data collection instruments or users' manuals for LDP compliance.

- ***Provision of adequate resources*** -- OIRM can make some limited resources available to programs for implementing the LDP (see section 4.2.3), but ***program offices must also redirect resources, in terms of personnel, funding, and equipment to meet the requirements of the LDP.*** An explanation of the expected costs and a plan for covering them should be included in the program's LDP Implementation Plan. A process for estimating resources for LDP implementation can be found in the *Guide to Selecting Latitude/Longitude Collection Methods.*

- ***Addressing LDP requirements in planning processes*** -- Planning documentation must be produced for a variety of EPA activities, including program planning, site sampling, project planning, and QA planning. ***Compliance with the LDP is a key element which media program managers must now include in preparing planning documents.***

- ***Coordination with the IRM Steering Committee to set priorities for policy implementation*** -- Program managers must work both with their program staff and the IRM Steering Committee to develop a sensible approach to implementing the LDP in their spatially-oriented projects, programs, and cooperative ventures. This coordination may be accomplished in phases, each phase may be prioritized according to various criteria such as need, availability of resources, opportunity, etc.

- ***Extension of the LDP to program personnel, agents, grantees, delegates, and regulated community*** -- Senior program managers must effectively institute compliance by***conveying the Policy to all other organizations, including states, other Federal agencies, sub-programs, contractors, and the regulated community with whom they are implementing their programs.***

- ***Development of DQOs for planning and waiver requests*** -- Program managers will develop DQOs when planning information collections. In addition, the ***DQOs will be used as the primary justification should a program wish to request a waiver from any portion of the LDP.***

The success of any venture depends on the individuals who play a part in it. Therefore, in addition to this discussion of *program* LDP implementation responsibilities, the responsibilities of key individuals are presented below.

4.3.3 Individuals/Functions

Within each office in EPA, certain individuals hold functional responsibilities that require them to be involved with LDP implementation. The individuals responsible for the LDP, summarized in Exhibit 4-6, include:

- ***Assistant Administrators, Associate Administrators, Regional Administrators, Laboratory Directors, Managers of Environmental Initiatives, and the General Counsel*** -- As stated in the LDP, these individuals shall "establish procedures within their respective organizations to ensure that information collection and reporting systems under their direction are in compliance with this Policy." Implementation of the LDP within each program or environmental initiative is the ultimate responsibility of its senior management. To acknowledge this responsibility, these individuals will participate in the development of LDP Implementation Plans with their staff to obtain an organization-wide view of LDP compliance.

- ***Senior Information Resources Management Officials (SIRMOs)*** -- SIRMOs will serve as liaison between the programs and the IRM Steering Committee, providing assistance to program management and staff on meeting the requirements of the LDP and developing Program LDP Implementation Plans. All waiver applications for data collection efforts within the program must be reviewed and endorsed by the SIRMO. Thus, the SIRMO will work with program managers in reviewing data quality objectives and formulating Program LDP Implementation Plans and waiver requests, transmitting them to the IRM Steering Committee, and ensuring the committee's comments are addressed.

- ***Program and Environmental Initiatives Managers*** -- Managers of programs and environmental initiatives will be responsible for developing LDP Implementation Plans. These individuals will identify affected activities and collections, establish steps to be taken to ensure compliance, and implement the necessary procedures. Managers must convey requirements of the LDP to all participants, including the regulated community, and ensure that procedures are changed so that LDP compliance becomes a

Exhibit 4-6

Individual Responsibilities Under the Locational Data Policy

Individual/Function	Responsibilities
Assistant Administrators, Associate Administrators, Regional Administrators, Laboratory Directors, and the General Counsel	Establish procedures within their organizations to ensure that information collection and reporting systems are in compliance with the LDP; aggregate Program LDP Implementation Plans into organization-wide LDP Implementation Plan.
Senior Information Resources Management Officials (SIRMOs)	Serve as liaison with the IRM Steering Committee; provide implementation plan development to programs; review and sign-off on waiver requests
Managers of Programs and Environmental Initiatives	Implement LDP within their programs or initiatives, by identifying activities and relevant data collections; preparing Program LDP Imp. Plans and waiver requests; endorsing use of GPS
Data Base Managers	Ensure data bases are modified so as to allow inclusion of locational information and ensure documentation is updated
Permit Writers	Ensure that data required by LDP are collected and documented during permit issuance and renewal processes
Site Managers	Ensure that LDP requirements are addressed in any site planning and activities; ensure use of GPS
Field Investigators	Ensure that required locational data are available for locations under investigation
Compliance Monitors	Ensure that adequate locational data are available for all monitoring sites

part of their routine. If compliance with LDP requirements is infeasible, program managers must work with program SIRMOs to develop waiver requests based on DQOs.[2] Program and environmental initiative managers also must oversee the incorporation of GPS technology into their data collection processes.

- *Data Base Managers* -- Data base managers (also referred to as *system managers*) will be responsible for ensuring that data bases are modified to hold the information defined by the LDP. Activities include creating data elements to house the information and edit checks to assure that the data are complete, in the correct format, have documented accuracy, etc. Data base managers also will be responsible for updating data entry procedures and supporting documentation (such as users' manuals, data entry manuals, data element dictionaries, training curricula, etc.) to accommodate LDP requirements.

- *Permit Writers* -- Permit writers will be responsible for ensuring that data required by the LDP are collected and documented during permit issuance and renewal. This may be achieved by requiring that the regulated community supply appropriate information on application forms, that coordinate information be included on the permit itself, and/or that coordinates be determined as part of a compliance schedule. Involvement of permit staff should be spelled out clearly in the Program LDP Implementation Plan.

- *Site Managers* -- EPA personnel often will be responsible for managing site activities (e.g., cleanup, etc.) performed by contractors or state participants. Site managers should be certain that LDP requirements are addressed in all site planning documents, including contracts, cooperative agreements, QA project plans, site sampling plans, remedial investigation/ feasibility studies, and other appropriate vehicles. Site managers also will be responsible for ensuring that LDP requirements are met during site activities, including the use of GPS technology.

- *Field Personnel* -- Field personnel, including inspectors, technical consultants (e.g., hydrogeologists, engineers, etc.), mobile lab operators, and others will be responsible for assuring that the required locational identification data are available for the locations on which they are working. Activities may involve determining whether locational data already have been

[2]Waiver requests for environmental initiatives should be prepared with the assistance of the initiative team member with responsibility for data management.

collected, obtaining site access, and either verifying existing locational data in the field or creating new location identification data. Field investigators also will be responsible for transmitting locational data (either new or renewed) to the appropriate office personnel so that it can be recorded in a computerized data base, site paper file, or other appropriate format.

- *Compliance Monitors* -- Certain program personnel, who operate either out of EPA headquarters or regional locations, are responsible for receiving and reviewing compliance reports submitted by the regulated community. These include Maximum Contaminant Level violation reports under the SDWA, emissions monitoring reports, RCRA ground water monitoring well sampling reports, Superfund samples, ambient and effluent water monitoring samples, process monitoring samples, and other types of monitoring results. Compliance monitors will be responsible for assuring the existence of adequate location identification information for all monitoring sites. Compliance monitors also will be responsible for informing data generators of the requirements of the LDP and assuring that this information is submitted to them. The details of compliance monitoring should be a component of the LDP Implementation Plan.

4.4 State, Indian Tribe, and Trust Territory Delegates

As stated in Chapter 1, ***this Policy extends to all state delegates of Federal environmental laws.*** This section outlines the expectations for state adoption of mechanisms to ensure compliance with the LDP. Other organizations that are not states, *per se,* including ***Indian Tribes and Trust Territories have a similar political relationship to EPA, and will follow the same guidelines.***

4.4.1 Variability of State LDP Responsibilities under Different Types of Federal Environmental Laws[3]

States can be delegated authority by EPA for several major environmental laws, including RCRA, SDWA, CWA, and CAA. When authorized, a state will define its universe of regulated, monitored, and tracked entities under criteria that are at least as stringent as those established by EPA. Therefore, each authorized state will play a major role in determining those entities for which LDP requirements must be met and in collecting LDP-required data for them.

[3]EPA will defer to the states as the final authority in determining adequacy of location identification data when a state has been delegated responsibilities for a Federal environmental program.

There are other Federal environmental laws under which states may play a corresponding role to EPA but are not authorized to implement in lieu of EPA. Under these laws, such as "Superfund" (CERCLA/SARA), other types of opportunities exist for states to address the requirements and goals of the LDP.

In either case, ***the major responsibility for ensuring LDP compliance by states will rest with the EPA managers or state managers of EPA mandated programs and/or environmental initiatives.*** These individuals must provide direction, guidelines, and/or assistance to their state counterparts in meeting LDP requirements.[4] Data integration to support risk reduction, pollution prevention, and other components of EPA's strategic plan will be a responsibility that program and environmental initiative managers must uphold and share with their state partners. Definition of state roles will be a key component of Program LDP Implementation Plans (Chapter 6). ***Whenever states are responsible to report locationally-based data to EPA, those data must comply with LDP requirements.*** This section presents the different possible roles that states play in Federal environmental laws and environmental initiatives, and their significance to state responsibilities under the LDP.

State-Authorized Programs

For Federal environmental programs delegated to states, ***the LDP applies only to those regulated, monitored, and/or tracked entities that meet criteria of Federal law.*** For those entities, states must collect and document lat/long coordinates, method, accuracy, and description as part of their routine reporting requirements to EPA. In addition, pursuant to the goal of the LDP, the use of GPS and the attainment of better than 25m accuracy is encouraged. Excluded are entities regulated, monitored, and/or tracked by a state under more stringent state criteria. For example, if the pound-per-year threshold above which an entity must be regulated is lower in an authorized state than under the corresponding Federal law, LDP requirements must be adhered to only for those entities above the higher (i.e., Federal) threshold.

For those entities regulated, monitored, and/or tracked by a state but not meeting the criteria of Federal environmental laws, state adherence to LDP requirements is voluntary and encouraged. Adherence to the LDP for these entities will streamline data integration between other states and/or EPA. Such integration will simplify the information transmission process, especially if the requirements are changed and the entities later meet Federal criteria and have to be reported on to EPA.

[4]In some instances, states are host to EPA program pilots.

The criteria that determine whether an entity is within the scope of a Federal environmental law will be developed by EPA media program offices. Similarly, the state authorization criteria also will be developed by EPA media program offices. In their Program LDP Implementation Plan, ***EPA media program offices will include the specifications for reporting by authorized states to ensure LDP compliance in all environmental data collections.***

Environmental Programs Run by EPA

There are two cases for environmental programs being run by EPA and not a state: (1) the program is not normally one that gets delegated to states (e.g., CERCLA, TSCA, FIFRA, or an environmental initiative such as the Chesapeake Bay Cleanup Program), or (2) the state has not been granted authorization under a law that can be delegated (e.g., RCRA, SDWA, CWA, CAA). ***EPA lead on a Federal environmental program or initiative does not preclude states from meeting LDP requirements when reporting locational data to EPA.***

Collection and reporting of locational data by states to EPA under Federal environmental law or under an inter-agency environmental initiative can be arranged under a variety of mechanisms, including cooperative agreements and grant stipulations (Sec. 4.4.3). Again, the criteria that determine whether an entity is within the scope of a Federal environmental law will be developed by EPA media program offices. Similarly, the definition of state data reporting requirements, and the mechanisms and vehicles by which to meet those requirements, also will developed by EPA media program office or environmental initiative managers. In their Program LDP Implementation Plans, ***EPA media program office and environmental initiative managers will address the specifications for reporting by states for non-authorized programs or initiatives to ensure LDP compliance in all environmental data collections.***

Often, states will have their own environmental laws that correspond to the Federal laws. For example, many states have their own "Superfund" programs. Adoption of LDP requirements and guidelines by these state-only programs is encouraged by EPA. State adherence to LDP specifications will facilitate data integration among states and between states and EPA, and may reduce the amount of data collection that must be performed.

4.4.2 Reporting Requirements

The entities for which states must collect and document lat/long, method, accuracy, and description have been defined above based upon the relationship between the state and EPA for individual programs or

initiatives. There are several possible mechanisms by which LDP-generated data can be transferred to EPA, including state data entry into a major Federal system and use of other reporting tools (forms, surveys, reports, etc.). ***The mechanism(s) to be used under each environmental program or initiative must be identified in Program (or Initiative) LDP Implementation Plans. Guidelines must be given to states on how to use them.*** For example, redundant reporting of location identification data should be avoided. Guidelines should be provided to states on when and how to submit LDP data (e.g., lat/long may not have to be measured every time a DMR report is submitted). These guidelines should be a component of the implementation plan. The mechanisms for states to report LDP-generated data are discussed below.

Data Entry into Federal Data Bases

Some states have elected to have direct data-entry access to certain EPA data systems such as RCRIS, AIRS, STORET, and PCS. Data entry can be either on-line or via electronic submission (i.e., tape or diskette). EPA system managers must ensure that appropriate data elements are in place within the data systems so that states can enter LDP-required data. ***State data managers must ensure that the locational data elements are populated for the system records of Federal data bases that they are responsible to maintain.***

"Tool" Submission

As discussed above, there are other tools (e.g., reports, surveys, and forms) that are used to report state information to EPA. The aggregate of all reports, surveys, forms, etc. submitted to EPA by states is considered a "data collection" and, if locationally-based (i.e., about a *place*), falls under the scope of the LDP. ***For all such tools used by states to create a locationally-based EPA data collection, LDP requirements must be fulfilled.*** EPA program or initiative managers will ensure that the forms, surveys, reports, and other "tools" are modified to capture the data required by the LDP from the state. They also will provide guidance to states on how to meet LDP requirements if use of such tools by states is optional. ***State participants are responsible to make sure that the tools are filled out completely and accurately for LDP-required data.***

4.4.3 EPA Support to States for LDP Participation

States have the responsibility to comply with the LDP for data that they collect for EPA. There are several types of EPA resources available to states so that they may fulfill their responsibilities. These resources include:

- *Grants* -- Grants to states are available from program offices, regional offices, and the SEDM program. In order for grants to be directed towards providing assistance to states in meeting LDP requirements, stipulations to that effect must be part of the grant. ***States will be accountable to EPA for grants which include a stipulation for adherence to the LDP.***

- *Cooperative agreements* -- Cooperative agreements are commonly used to document the expectations of states and EPA in environmental projects jointly undertaken. For example, cooperative agreements are used in the "Superfund" program to identify actions for those sites that are highest priority to states, regardless of whether they meet Federal National Priority List (NPL) criteria. ***Cooperative agreements can be used to specify activities that will ensure LDP compliance by states,*** including support that will come from EPA and/or error resolution processes.

- *Loans* -- EPA often provides loans to states to fulfill certain mission requirements. For example, the State Revolving Fund loan program has replaced the Construction Grants program under CWA. ***Loan amounts and specifications could be created to include the support necessary in meet ing LDP requirements.*** Additionally, ***EPA regional offices can enter into a loan program with their states for equipment such as GPS devices.*** This arrangement may ensure that locational data submitted to EPA by states are created using the GPS technique and meet the 25m accuracy goal.

A description of the incentives for meeting LDP requirements by the LATF can be found in Appendix A.

4.5 Designated Representatives

Chapter 1 defined all designated representatives of EPA (i.e., ***contractors, grantees, and universities)*** as participants in the LDP. Those representatives will ***have the same responsibilities as those offices within EPA who collect pertinent data or contribute to relevant data bases.*** Language that requires adherence to Agency IRM policies and standards is often a standard part of Agency contracts and grants. Program LDP Implementation Plans should describe the means that programs will use to ensure LDP compliance by designated representatives.

Chapter 5

AGENCY-WIDE SCHEDULE

5. AGENCY-WIDE SCHEDULE

Successful and consistent implementation of the Locational Data Policy requires diligent coordination among EPA offices and outreach efforts to organizations external to the Agency. In order for implementation to be completed within an appropriate time frame, Agency-wide priorities and milestones for implementation have been established. These priorities and milestones are discussed in this chapter. Additionally, areas requiring coordination with other Federal agencies are identified.

5.1 Schedule and Priorities

This section defines the general schedule and milestones for LDP implementation and establishes priorities for certain activities. The milestones require action either on the part of OIRM or the programs. These milestones and their sequence are summarized in Exhibit 5-1.

5.1.1 Accuracy Goal

As discussed in Chapter 3, the Locational Accuracy Task Force recommended that an accuracy goal of 25 meters be established for new data. The Agency-wide goal is to have ***all locational data accurate to better than 25 meters by December 31, 1995.*** The LATF also recommended that, beginning in 1992, map interpolation be used to determine lat/longs with GPS being phased in gradually. After 1995, GPS should be the principal method of collecting locational data.

5.1.2 LDP Implementation Plans

The LDP calls for EPA managers to create implementation plans within their programs. Assistant Administrators will coordinate all the plans produced within their organization so that there is clear understanding of responsibilities. The components of the plan (Chapter 6) include such items as an estimate of the amount of information which must be collected and documented, the necessary steps to be taken to fully implement the LDP, estimates of time and resource requirements, and issues or constraints which must be confronted. OIRM will review LDP Implementation Plans and will pass their assessment of the plans to the IRM Steering Committee both to ensure that LDP requirements are adequately addressed and to coordinate LDP implementation Agency-wide. The plans ***should be completed and submitted to OIRM for review by the end of the FY92 (September 30, 1992).*** At that point, OIRM will review the implementation plans and confer with the submittors to resolve any outstanding issues. The IRM Steering Committee will give feedback to submitters of LDP Implementation Plans by December 1992.

5.1.3 Existing and Planned Systems

Many EPA data systems will require modification to house or use the locational coordinates and qualifiers required by the LDP. Some systems, such as STORET, can already accommodate some of the required information. Other systems currently have no fields to house any of the LDP-required information. It is critical that these data elements be included in computerized data bases to enable secondary use based on location. Among the first steps in LDP implementation is redesigning all existing and planned data systems to accommodate required locational data in the formats specified.

All modifications of existing and planned systems to accommodate lat/long coordinates, method, accuracy, and description should be designed by December 1992. See Chapter 2 for a discussion of EPA systems that fall under the scope of this policy. Implementation of the redesigns should then proceed as quickly as resources allow and be ***completed by December 31, 1995.*** Details of the necessary modifications can be coordinated between OIRM and the managers of individual program systems. In general, ***the types of changes that might be made to systems to make them compatible with the requirements of the LDP include:***

- ***Data element addition*** -- Creating new data elements for latitude, longitude, method (including datum, and map scale, accuracy, and description), accuracy, and description.

- ***Data element formatting*** -- Creating or modifying the data element fields to correctly house LDP data. Examples include formatting coordinate component values to the correct length; representing a decimal point in the correct place; allowing four digits to the right of the decimal point, or allowing the storage of a "+" or a "-" to indicate hemisphere.

- ***Edit check addition*** -- Developing edit checks in the data entry screens (or as part of the data base building process) to filter impossible values or to verify coordinates based on other locational information (e.g., validity of a lat/long pair verified by comparison to state or county boundaries), or to check an entire database in batch mode to flag "bad" data values for later correction.

- ***Functional capability development*** -- The development of capabilities for users of locational data; examples include searching within a geographic area, sorting by method, or screening by accuracy.

An explanation of needed/planned system modifications should be a component of each program's LDP Implementation Plan (Chapter 6). The computer system implications of the LDP are described more fully in Chapter 8.

5.1.4 Historical Data

The Locational Data Policy requires that historical data collections (existing data bases) eventually be populated with coordinate data for their locational entities. Chapter 7 of this document discusses updating options for historical data. General priorities for historical locational data were given in Chapter 2 and reiterated below:

- ***Data collections with geographic coordinates other than lat/long are of the highest priority for revision.***
- ***Data collections with lat/long coordinates defining the locations of their entities but incomplete or missing lat/long qualifiers are of second priority.***
- ***Data collections with no geographic coordinate data are of third priority but must comply with the LDP by the end of 1995.***

LDP Implementation Plans should address the schedule for updating historical data bases. In general, ***first priority data collections*** (those with location data in another coordinate system) ***will be made to comply with the information requirements of the LDP by the end of December 1992. Second-priority data collections*** (those in which lat/long metadata elements are missing or incomplete) ***will be updated to meet LDP data requirements by the end of December 1993. Third-priority data collections*** (those with no geographic coordinate data) ***will be brought into full compliance with the LDP by the end of December 1995.*** By the end of calendar year 1995, all office-created locational coordinate data (i.e., those produced by methods including address matching, zip code, centroid, etc.) should be replaced with more accurate field-collected coordinate data obtained using GPS.

5.1.5 Form and Contract (Tool) Modifications

There are many types of documents that program managers use as tools to collect the data needed to accomplish mission objectives. Many of these documents can be used to meet LDP requirements. These documents include:

- Notification forms (e.g., RCRA Form 8700-12).
- Application forms (e.g., NPDES Forms 1, 2c).
- Survey forms (e.g., RCRA biennial survey forms).

- Standard reporting forms (e.g., TRI Form R).
- Contracts (e.g., with RAC contractors for site cleanup).
- Grant documentation (e.g., for State/EPA grant monies).
- Cooperative agreements (e.g., for EPA-State Superfund cleanup prioritization).
- Memoranda of understanding.
- Pre-formatted diskette data collection tools (e.g., for the Water Body System).

New lat/long data will be collected in many programs to meet the requirements of the LDP. Where lat/long already are being collected, locational metadata (method, accuracy, and description) must become part of the data collection routine. Therefore, the tools for information collection must be modified to assure LDP requirements are met.

Modification of data collection tools must be coordinated with data system redesign efforts and should occur prior to collection of missing data. The tools used to collect data for high priority systems similarly should be given highest priority for redesign. ***All data collection tool redesign efforts should be completed by the end of March 1993.*** The plans, steps, and schedule for data collection tool redesign should be documented in each LDP Implementation Plan. In addition, LDP Implementation Plans should describe instructions to be given for providing locational data on the forms. Locational data may not need to be collected or submitted repeatedly with each form or report submission. The times and frequency of LDP data collection should be clearly identified on the forms and in the plan.

5.1.6 Role of FINDS

A final step in LDP implementation is populating coordinate fields in FINDS. This activity will be accomplished by a combination of address matching and capturing locational data from program systems. FINDS will serve as the information utility for providing facility locational information to the environmental community and the public. OIRM currently is undertaking an effort to modernize FINDS. One component of FINDS modernization is to include lat/long coordinates in each facility record. Steps to include locational data in FINDS will include:

- Creating lat/longs for the "facility in general" in the FINDS "master records" by address matching with corresponding records in the Dun & Bradstreet File, resulting in one set of coordinates per facility.

- Creating data element fields in the "alias records" of each facility to capture lat/longs from program systems if they are available, enabling a degree of tiering within FINDS.

The FINDS address matching project to populate the master record is scheduled for completion by June 1992. Population of lat/longs in the alias files will begin with synchronized updates. The specifications for performing synchronized updates for the highest priority national systems will be completed by April 1992[1].

5.2 Coordination with Other Federal Agencies

Implementing the Locational Data Policy requires coordination with other Federal agencies that have expertise and responsibilities for spatial data. Agencies that deal with information collection issues also will be included. Coordination between OMB, the FGDC, and the FGCC is especially important. Interaction between all involved agencies should occur through a centralized point, rather than through each office or program. Appropriate coordination points are described below.

5.2.1 Office of Management and Budget (OMB)

Under the Paperwork Reduction Act, OMB clearance is required when information is to be collected by the Federal government from more than nine non-Federal entities. The clearance process involves preparing and submitting a formal *Information Collection Request* (ICR), to justify the collection effort, describe the costs involved, and analyze the burden on the regulated community. The Office of Policy, Planning, and Evaluation (OPPE) within EPA provides the interface between OMB and the respective Agency offices by reviewing ICRs prepared by individual programs and environmental initiatives.

5.2.2 Federal Geographic Data Committee (FGDC)

The Federal Geographic Data Committee (FGDC) is an interagency committee chaired by the Department of the Interior. It is responsible for coordinating Federal digital cartographic and GIS activities and for establishing standards and specifications for digital cartographic data. EPA coordinates with the FGDC to ensure Agency compliance with these standards and to provide input into the Committee's decisions. Participation in this committee ensures EPA awareness of relevant activities in other agencies and to coordinate with these agencies where appropriate. OIRM's National GIS Program acts as EPA's contact with the FGDC. Under the LDP,

[1]As planned in the 11/91 Facility Identification Data Standard Implementation Plan (FIDSIP)

"EPA is committed to implementing its locational policy in accordance with the requirements specified by the Federal Interagency Coordinating Committee for Digital Cartography (FICCDC)."[2]

5.2.3 Federal Geodetic Control Committee (FGCC)

The FGCC, is a component of the Department of Commerce, and is responsible for national geodetic control surveys and datum. The transition from NAD27 to NAD83 requires coordination with the FGCC. OIRM's National GIS Program has appointed a representative to the FGCC for the Agency.

[2]Recently renamed the Federal Geographic Data Committee (FGDC) Circular A-16

Chapter 6

LDP IMPLEMENTATION PLANS

6. LDP IMPLEMENTATION PLANS FOR PROGRAMS AND ENVIRONMENTAL INITIATIVES

Implementation of the Locational Data Policy must be carefully planned so that it is effective. This document describes the overall Agency plan for implementing the LDP. EPA media programs and relevant data collection projects (such as the major environmental initiatives) must develop detailed plans for carrying out the LDP. For the purposes of LDP implementation, ***"media programs" are those responsible for implementing an environmental law.*** Therefore, media programs are "offices" that implement one or more of the laws such as RCRA, TSCA, CERCLA, CWA, CAA, FIFRA, SDWA, and other environmental laws. In some cases, media programs are divided into several sub-programs (e.g., Office of Water implements NPDES and oversees non-point source pollution projects). ***All sub-programs must be addressed in the implementation plan of each office.*** Similarly, ***environmental initiatives are joint-organizational, cross-media projects that do not fall under the jurisdiction of one media program.***

LDP Implementation Plans prepared for EPA media programs and environmental initiatives will accomplish the following:

- They will *give programs the opportunity to identify and anticipate all the steps* needed to meet the requirements of the LDP.
- They will give *the IRM Steering Committee the information needed to coordinate LDP implementation* Agency-wide.

One very useful tool for planning implementation of the LDP within a program involves preparing ***Data Quality Objectives*** (DQOs) for the data collection efforts. EPA Order 5360.1, *"Policy and Program Requirements to Implement the Quality Assurance Program,"* documents EPA's commitment to careful planning to ensure that data meet the needs for which they are collected. The resulting EPA Quality Assurance Program has developed a process built upon DQOs. Because locational data are an essential part of measurement data, DQOs should be prepared with locational data explicitly addressed.

DQOs currently require planners to define the level of data quality in terms of primary data use. Under the LDP, however, likely secondary uses also must be considered. These considerations may affect the definition of data quality.

After the quality objectives are defined, plans must be developed to assure attainment of them. These plans and the DQO process should be included in each program and initiative's LDP Implementation Plan. DQOs will help implementation in the following manner:

- DQOs will help data managers determine the exact level of accuracy needed for their locational data (the LDP strives toward a minimum accuracy of 25 meters).
- By knowing the level of accuracy needed, program managers can select appropriate methods for locational data collection.
- By determining the necessary locational data accuracy and appropriate methods, other plan components (e.g., steps-to-achieve and roles and responsibilities) may be defined.

Should it be necessary to apply for a waiver for any of the provisions of the Locational Data Policy, DQOs will be the tool that programs use to justify their waivers. The IRM Steering Committee will rely heavily on the DQOs when reaching a decision on the waiver application.

The Quality Assurance Management Staff within the Office of Research and Development can provide guidance and tools for the DQO process. ***DQOs should be a part of every media program's plans for locational data collection and documentation.***

The remainder of this chapter discusses the development of LDP Implementation Plans. It approaches these plans both in terms of their components and the process for their development. Components of LDP Implementation Plans are summarized in Exhibit 6-1.

6.1 Components of LDP Implementation Plans

Many decisions must be made by media program and data managers on how to implement the LDP in their organizations. Each of those decisions should be documented as a component of the LDP Implementation Plan. ***For each component, responsibilities must be assigned, a budget must be identified, and a schedule must be defined.*** If, upon developing a program LDP Implementation Plan, it is determined that some LDP requirement(s) cannot be met, then a waiver must be sought. This section discusses the components expected to be in each plan. An example of a table of contents for a LDP Implementation Plan is displayed in Exhibit 6-2.

6.1.1 Identification of Entities to which the LDP Applies

One of the first determinations program managers must make is ***which entities in that program require locational data under the LDP.*** Chapter 2 of this document identifies the types of entities for which location identification data should be collected and documented. ***Exact determination of the entities to which the LDP applies must be documented in the LDP Implementation***

Plans. This determination should be linked to DQOs based on the type of other non-locational data associated with the respective entities.

6.1.2 Definition of Activities During Which LDP Data Will Be Collected

There are many opportunities when location identification data can be collected. These opportunities can be divided into two categories:

- *Activities during which entity location data MUST be collected and documented.*
- *Activities during which entity location data CAN be collected and documented.*

Activities during which location identification data MUST be collected and documented should be identified in LDP Implementation Plans. These activities will usually occur while generating new entity identification data (i.e., data for that entity do not yet exist). Such activities might include notifications, applications, or sample plan development.

Activities during which location identification data CAN be collected also should be identified in the LDP Implementation Plan. These activities are usually those that offer the opportunity either to double-check the quality of existing locational data or to provide missing data, such as locational accuracy. Such activities could include inspections or routine emissions reporting. If gaining site entry may be a problem, an approach to overcome this obstacle should be described in this plan. If acquisition of location identification data is a component of a "technology transfer and data sharing" arrangement with another organization such as a state, this relationship and the requirements of the LDP should be explained and formalized.

6.1.3 Enhancements to Data Bases and Systems

Computerized data bases are usually the eventual destination of most data that EPA and its agents collect. Chapter 2 of this Guidance Document contains a discussion of the types of data bases affected by the LDP. ***In their LDP Implementation Plans, data managers must list explicitly the data bases maintained by or for their programs that will be affected by the LDP (i.e., those that need data element modification, the addition of coordinate data, or other changes). Further, data managers must state how the data bases will be manipulated/modified to comply with the LDP.*** When addressing data bases/systems in the LDP Implementation Plan, the following items should be included:

Exhibit 6-1

LDP Implementation Plan Summary

Component	Description
Entity identification	All entities regulated, tracked, monitored, etc., by program for which locational identification data must be collected and documented, along with the DQOs for the locational data
Data collection activities	Each activity for which LDP compliance (i.e., collection and documentation of location identification data) will occur
Data base/system enhancements	Modifications necessary to systems/data bases for compliance with the LDP (e.g., new data elements, edit checks, etc.)
Tool redevelopment	Modifications necessary to data collection instruments (e.g., application forms or reports used to capture LDP data)
Format of LDP-required data	Specifications for each type of entity location data (i.e., whether the entity is represented by a point, line, or area, etc.)
Methods	Methods for data collection/conversion to be used by program to comply with LDP, by entity or activity if appropriate
Roles and responsibilities	The roles and responsibilities of program implementors in complying with LDP, including EPA, states, etc.
Resources/Priorities	Estimates of expected costs of LDP compliance within program, and identification of potential funding sources and implementation priorities
Quality assurance	Quality assurance measures to be adopted to assure adequate quality of locational data, such as edit checks, double-checks, etc.
Constraints	Anticipated restrictions or limitations, and approaches for overcoming them to comply with LDP.
States	Reporting requirements and processes, resource availability

- ***The data bases/collections in which locational data will reside, including the following information:***
  - The purpose of the data base system.
  - The type of locational data already present in the data collection.
  - The numbers of entities requiring locational data (on an existing- and new-per-year basis) and the percentage of them already satisfying LDP requirements.
  - Changes (including an estimate of costs to make changes, and sources of funding) that are necessary to the data base/system to accommodate the required LDP data, such as:
    - Data system redesign (e.g., new data elements or edit checks).
    - Data acquisition (e.g., automated interface to other data bases or incremental field collection).
  - Centralized solutions (Chapter 7) to be used to meet LDP requirements, such as conversion software.
  - Incremental solutions (Chapter 7) to be used to meet LDP requirements.
  - Automated and/or manual mechanisms for error detection and correction.
- ***Priorities assigned to data bases/systems for conformance to the LDP.***
- ***Excluded data bases (i.e., those for which waivers will be requested) and reasoning for exclusion of data collections*** to which data managers feel that the LDP does not apply.

Computer system implications of the LDP are discussed in more detail in Chapter 8.

6.1.4 Tool Redevelopment for Collecting and Documenting Locational Data

A discussion of the many tools used by EPA programs to collect and/or require necessary information was presented in Chapter 2. Such tools include application forms and memoranda of understanding.

Exhibit 6-2

Example of Table of Contents for LDP Impelmentation Plan*

TABLE OF CONTENTS

1. Entity Identification
1.1 Documentation of the size of entity universe
1.2 Determination of the number of entities already in compliance with the LDP
1.3 Identification of all entities for which lat/long data must be collected/documented

2. Activity Identification
2.1 Definition of activities during which location identification data should routinely be collected
2.2 Definition of opportunities for incremental data collection (e.g., inspections)

3. System Modification and Software Acquisition Needs
3.1 Identification of systems to which the LDP applies
3.2 Determination of presence of needed data elements in key data systems
3.3 Comparison of program system data elements and formats to LDP requirements
3.4 Identification of other needed modifications and determination of plans, costs, etc.
3.5 Recommendations for edit checks or other error detection mechanisms
3.6 Identification of "functionality" of LDP data elements to ensure value to secondary users
3.7 Identification of systems to be excluded from the LDP, and waiver request plan
3.8 Determination of what level facilities and their sub-portions should be represented in FINDS
3.9 Determination of methods, and costs to acquire/convert data to LDP format

4. Tool Redevelopment Needs
4.1 Identification of forms, reports, surveys, etc. that apply to the LDP
4.2 Determination of adequacy of each tool to meet LDP requirements
4.3 Definition of steps needed to make changes to each applicable tool
4.4 Determination of implications of change
4.5 Recommendations for a process to notify regulated community of locational data requirements
4.6 Suggestions for linking tools, activities, or systems to reduce redundancy for capturing data

5. Standard Way of Locationally Representing Each Entity (e.g., point, line)
5.1 Examination of the range of entity types to recommend how each type should be represented
5.2 Determination of amount/type/audience for needed guidance/revision of locational definition
5.3 Definition of format for location data (e.g., to how many decimal places)
5.4 Assesment of burden hour requirements
5.5 Identification of document requirements
5.6 Definition of facility representation
5.7 Determination of whether date and source should be part of locational data documentation
5.8 Definition of datum used and whether conversion to NAD 83 is necessary
5.9 Determination of circumstances under which data might deviate from 25 m accuracy goal

6. Location Identification Determination Method
6.1 Recommended method(s) to collect coordinate data based on data, entity, activity, etc.
6.2 Documentation of status of/requirementss for use of selected method

Continued on next page

* -- May be modified upon development of LDP Implementation Plan Template

Exhibit 6-2

Example of Table of Contents for LDP Impelmentation Plan* (continued)

TABLE OF CONTENTS

7. Role/Responsibility Identification

7.1 Definition of planning responsibilities; identification of key participants
7.2 Identification of responsibilities for quality control
7.3 Documentation of budgetary commitments on the part of all participants
7.4 Definition of reporting requirements/responsibilities
7.5 Definition of a process for interfacing (e.g., between states and Federal program staff)
7.6 Identification of responsibilities for system redesign work
7.7 Determination of plans/guidance/budgets/standards which must be developed and implemented
7.8 Determination of responsibilities for collecting location idetification data within a program
7.9 Responsibilities for transmitting location identification data to EPA data bases

8. Costs and Funding Sources, Priority of Activities, Steps to Complete, Schedule

8.1 Determination of feasible incentives that could be put toward states
8.2 Identification of grants to assist states/agents implement the LDP
8.3 Estimation of the costs of system redesign(s)
8.4 Estimation of costs of data acquisition to comply with the LDP
8.5 Priority list of activities to be funded first in case of competing resources
8.6 Estimation of the amount of equipment needed and costs related to purchase, distribution, etc.

9. Quality Assurance

9.1 Process for QA/QC, including edit checks and double-checks against other references
9.2 Definition of responsibilities for QA, including error resolution
9.3 Review of DQOs to assure that capture of location idetification information is addressed
9.4 Review of QA plans to ensure LDP requirements are addressed

10. Constraints

10.1
Identification of possible impediments or dependencies to full-scale implementation of the LDP
10.2
Recommendations to address identified constraints

* -- May be modified upon development of LDP Implementation Plan Template

Other data collection and documentation tools also may apply. These tools may be in paper or electronic form, such as diskette or tape. ***The LDP Implementation Plan must identify the tools used to collect data, whether any changes will be needed to comply with the LDP, and how those changes will be made.*** Changes to these tools may require notification in the Federal Register or coordination with OPPE or OMB. Sometimes these tools may be part of an information exchange program with state or local governments. All modifications and steps to accomplish them should be made clear in the LDP Implementation Plans.

OIRM briefly studied the forms listed in the April 1989 EPA Forms Catalog to determine a partial listing of candidate forms which might require changes to comply with the Locational Data Policy. The methodology of the study was as follows:

- Forms that might be relevant to the Locational Data Policy were identified by reviewing the titles of the 637 EPA forms found in the EPA Forms Catalog.
- Each form was examined to determine those that already contained relevant locational information, such as address, lat/long, etc.
- The information on each form was reviewed to determine whether LDP information is currently missing.

The percentage breakdown of the forms is presented in Exhibit 6-3. Of the 637 forms listed in the EPA Forms Catalog, 59% were eliminated from review based on their titles. The remaining 262 forms were examined to determine if they contain relevant locational information. Fifty-eight (9.3%) of the 262 forms were found to have a significant spatial data component. These forms were then examined to determine the extent to which they already comply with the LDP requirements.

The list of these relevant forms is presented by program in Appendix C. These forms may need modification to ensure that sufficient/appropriate information is collected in compliance with the LDP. It may be possible, however, for a program to link computer systems to obtain the necessary locational information, thereby reducing the number of forms on which, or the number of times when, locational data must be collected. The implementation plan of each program or environmental initiative must describe an approach to ensuring that the proper locational information is included on the appropriate forms.

6.1.5 The Format of LDP-Required Data

Data managers must define specifications for each type of locational data. This definition will include a description of:

- ***Whether the entity location will be represented by a point, line or polygon, including whether a specific point (e.g., the "front door," the "driveway," or the "facility centroid") will be used to represent what is actually an area.***
- ***To how many decimal places the locational coordinates will be documented.***
- ***Whether source and date will be collected and documented.***
- ***How the NAD83 requirement will be fulfilled.***
- ***Whether the level of locational accuracy will be a deviation from the 25 meter goal, and thus, require a waiver request.***

Specification of the exact form of locational data will allow data managers to evaluate whether those data will adequately support their missions and/or secondary use.

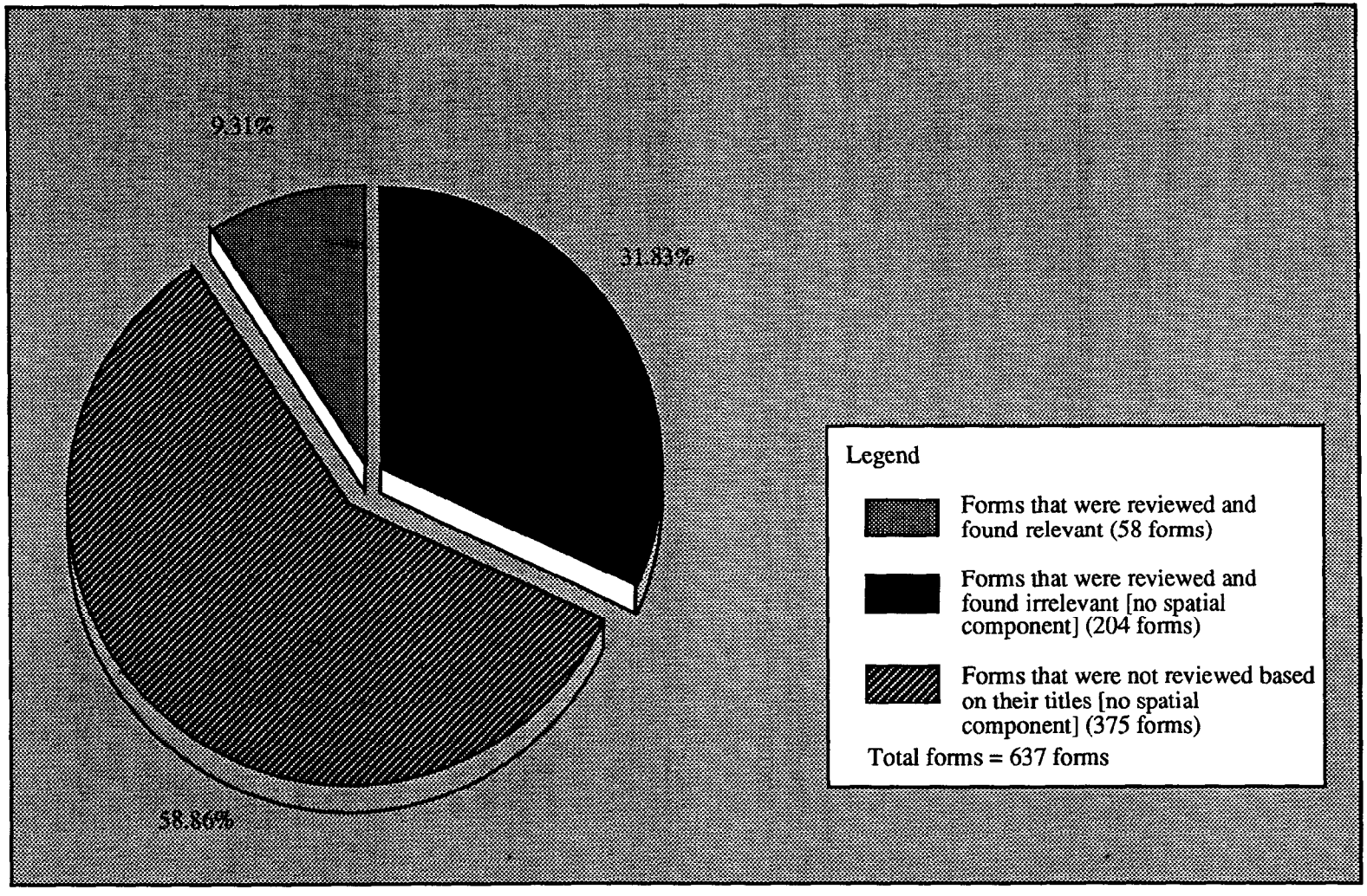
6.1.6 Identification of Methods to be Employed

LDP Implementation Plans should include a discussion of locational data collection methods being considered by the program to meet LDP requirements. In developing these plans, program managers may decide that certain methods are appropriate for certain situations. For example, conversion software may be adequate to provide missing locational data in existing data records but Global Positioning Systems may be used in the field for new data collection. Eventually, all locational data should have similar or better accuracy than that which can be obtained by using GPS.

Chapter 7 of this LDP Guidance and the *Guide to Selecting Latitude/Longitude Collection Methods* present an introduction to some of the methods available for locational data collection and conversion. Each is addressed in terms of the expected costs and ranges of accuracy. Several EPA technical resources, including EMSL-Las Vegas, the National GIS Program, and the Regional GPS Work Group can provide more assistance in method selection. Beginning in 1992, use of GPS should be phased in and so indicated in the LDP implementation Plan. A primer on GPS use accompanies this guidance package.

Exhibit 6-3

Review of the 637 Forms in the EPA Forms Catalog



6.1.7 Definition of Roles and Responsibilities

Implementing the LDP within a program requires coordination of many involved parties. Thus, it is important that roles and responsibilities be clearly assigned in the LDP Implementation Plans. ***Roles and responsibilities should be defined for:***

- ***EPA/HQ and Regional program personnel.***
- ***State participants.***
- ***Other government organizations (e.g., DOE, COE, USGS).***
- ***The SIRMO.***
- ***EPA/HQ and Regional personnel other than those in the immediate program.***
- ***Contractors.***
- ***Grantees.***
- ***Universities.***
- ***The regulated community.***

Any participant not listed above, but who could play a key role in program implementation of the LDP, should be identified as well.

6.1.8 Resources/Priorities

Adequate commitment of resources is one of the most critical factors upon which successful implementation of the LDP depends. Resource availability must be clearly planned because it affects implementation schedule and scope. Because resources are limited, implementation priorities and associated resource use should be defined in the LDP Implementation Plans. ***The LDP Implementation Plans must address resources in terms of a <u>cost estimate</u> of LDP implementation including both for initial and recurring costs. Collaboration with other programs or environmental initiatives for collecting LDP-required data might be considered during plan development as a way to reduce costs.***

The *Guide to Selecting Latitude/Longitude Collection Methods* describes a process for estimating a budget for locational data collection. Chapter 4 and Appendix A of this document identify resources that can be made available to media programs.

6.1.9 Quality Assurance Measures

Another key component of LDP Implementation Plans is the presentation of an approach to quality assurance of locational data. Quality assurance can be implemented in several ways, ranging from repeated collection of locational data for the same entity, to edit checks on computer system data entry screens. ***Managers of programs and environmental initiatives must design and describe a plan to address quality assurance of locational data, including identification of measures to be taken and responsible personnel.*** Adherence to quality assurance plans will ensure that locational data achieve the level of quality needed to meet their intended application.

6.1.10 Constraints

Managers of programs and environmental initiatives will face obstacles in implementing the LDP. These obstacles may be in the form of delays, resource limitations, regulatory authority, or dependence on other organizations or events. In their LDP Implementation Plans ***managers of programs and environmental initiatives must identify the expected constraints that will be confronted, and present an approach for addressing them.***

6.1.11 State Roles and Responsibilities

EPA will rely on states for the generation, collection, documentation, and submission of much of its locational data. States will rely on EPA for guidance and assistance in meeting this responsibility. Section 4.4 identifies the responsibilities states will have under the LDP. The ***specific activities to ensure state compliance with the LDP must be documented in the LDP Implementation Plans of all media programs and environmental initiatives.*** Topics to be addressed include:

- State reporting requirements under the LDP, including activities during which states must meet LDP requirements (e.g., report submission)
- Coordination of data collection and reporting activities to avoid redundant locational data collection by states
- Establishment of a process for resolution of errors in locational data submitted to EPA, and other procedural guidelines that might be necessary to enable state adherence in the LDP
- Resources (grants, equipment loans, etc.) to be applied towards state adoption of the LDP.

Assignment of state responsibilities under the LDP may depend on whether a state has been delegated authority for a particular Federal law. LDP Implementation Plans may have to address separate circumstances for delegated and non-delegated states.

6.2 Process for Developing LDP Implementation Plans

Chapter 4 of this LDPG specified that ***LDP Implementation Plans are to be completed and submitted to OIRM for review by June 1992. OIRM will work with the IRM Steering Committee and the submitters to review and finalize their plans by the end of September 1992.*** OIRM offers the following suggestions for the process for developing their LDP Implementation Plans.

6.2.1 Coordination of Steps to Achieve

By compiling the components of the LDP Implementation Plan as described previously, managers will identify essential steps. Execution of these steps must be coordinated so that LDP implementation progresses as smoothly as possible within each program or environmental initiative. A proposed table of contents for a LDP Implementation Plan is presented in Exhibit 6-4. OIRM also offers workshops on creating LDP Implementation Plans.

6.2.2 Phasing

Managers of programs or environmental initiatives may find it beneficial to phase-in implementation of the LDP. For example, certain entities, data collections, or processes may be targeted for conforming to the LDP first, others second, and still others third. ***The defined sequence of events and phases should be part of the LDP Implementation Plan.*** If the steps to be taken or the phases to be implemented do not conform to the Agency-wide plan in this document (for example because of the time needed to coordinate state activities) a waiver should be sought.

Chapter 7

DATA CONVERSION AND COLLECTION METHODS

7. DATA CONVERSION AND COLLECTION METHODS (GEOCODING)

In complying with EPA's Locational Data Policy, program and initiative managers will have to decide whether to convert old locational data to new formats, collect entirely new locational data, or use a combination of conversion and collection methods. The guidance presented in this chapter and the accompanying *Guide to Selecting Latitude/Longitude Collection Methods*, are intended to introduce the array of available locational data conversion and collection methods and to assist in selecting an appropriate method based on needs, accuracy, and available resources. More detailed information on data conversion and collection methods can be obtained from EPA's GIS Program and from the Office of Research and Development (ORD). Guidance on the use of GPS technology can be found in the *GPS Primer* which also accompanies this document.

A key concept to be introduced in this chapter is the idea of *geocoding*. The *Guide to Selecting Latitude/Longitude Collection Methods* defines geocoding as "the application of procedures, techniques, and technologies for the purpose of identifying and documenting the geographic location of an entity." Therefore, collecting new data is an integral part of the geocoding process. In contrast, converting old data is not considered to be "geocoding." Data collection (i.e., geocoding) must precede conversion.

7.1 Centralized vs. Incremental Data Conversion and Collection

The *Guide to Selecting Latitude/Longitude Collection Methods* discusses two distinct approaches to data conversion and collection: centralized and incremental. These "approaches" are the "institutional" ways of fulfilling LDP requirements.

The *centralized* approach is the process of converting or collecting all the needed locational data at once. An example of centralized data collection is the creation of lat/long coordinates for an entire data base of facilities by using address matching software to convert facility addresses to lat/long coordinates. This approach is being used for conversion and creation of locational coordinate data for facilities in EPA's FINDS data base.

The *incremental* approach involves collection of new, or replacement of existing locational data over time, by measuring lat/long coordinates during routine activities such as inspections or surveys. Incremental collection of locational data becomes an additional function of field personnel. Routine data collection or update activities become opportunities for these personnel to create new, or improve existing, location identification data.

The principle differences between the two approaches lie in the time and dollars required to collect all the necessary data (a relatively short, intensive effort for the centralized approach vs a long term, gradual effort for the incremental approach). The differences between the approaches are summarized in Exhibit 7-1. Note that there is probably a continuum of tradeoffs between incremental and centralized conversion and collection techniques.

7.2 Conversion Methods

Data conversion is a method of creating lat/long coordinates from location identification data in another format. Although it may avoid the cost of new data collection, it may incur the possible "cost" of accuracy loss. This section describes methods available for converting from:

- Non-coordinate locational information, such as address or ZIP code, to latitude/longitude.
- A different geodetic coordinate system to latitude/longitude.
- One datum to another.

Of these, only the conversion of non-coordinate data is a true "geocoding" method. The others, instead of creating geodetic coordinates, convert them form one format to another. Under each of the three categories this section describes, appropriate process, levels of expertise, limitations, expected accuracies, and relative costs. Additionally, general guidance is provided on the assistance available for each conversion method and on method selection. A process for determining a suitable geocoding method is presented in the *Guide to Selecting Latitude/Longitude Collection Methods*. A comparison of locational data conversion methods is provided in Exhibit 7-2.

7.2.1 Conversion from Non-Coordinate Geopolitical Locational Information to Latitude/Longitude

Two primary methods exist for converting from geopolitical locational data to latitude and longitude. These methods are *address matching* and *ZIP code centroid*.

Address Matching

Automated conversion from address information to latitude and longitude is commonly referred to as "address matching." Address matching uses digitized street data to match an entity's address to the

Exhibit 7-1

Incremental vs. Centralized Data Conversion and Collection

CONVERSION	Converting existing location identification data in records of data base one-by-one, over time	Conducting an effort specifically to convert existing locational data of an entire data base to LDP-appropriate format all at once
COLLECTION	Collecting new (or double-checking existing) locational data by using the field opportunity of other routine activities (e.g., inspections)	Gathering new locational data for an entire data base or network of entities all at once

Exhibit 7-2

Comparison of Locational Data Conversion Methods

Method/ Characteristic	Address Matching	Zip Code Centroid	ARC/INFO	GCTP	NADCON/ CDATUM
Description	Converts from address to latitude/longitude through matching of address to nearest street section and interpolating	Assigns coordinates based on ZIP Code centroid to entity within ZIP code	Converts from one coordinate system to another through the use of algorithms	Converts from one coordinate system to another through the use of algorithms	Converts from one datumto another through modeling datum shift information and interpolating. CDATUM used for ARC/INFO, NADCON used for ASCII
Entity Type	Any entity with an address	Any entity with a ZIP Code	Any entity which has a geodetic coordinate assignment	Any entity which has a datum assignment	Any entity which has a datum assignment
Accuracy	50-100 meters	Ranges from 100-500 m in urban areas to many km in rural areas	Accuracy of original coordinates is maintained	Accuracy of original coordinates is maintained	Conversion accuracy of 15cm for conterminous US and 1.0 M for other areas
Cost/Point	$7-$30/Point	$.05-$5.00/Point	n/a	n/a	n/a
Required Expertise	Requires no special expertise - process automated	Requires no special expertise - process automated	Requires no special expertise - process automated	Requires no special expertise - process automated	Requires no special expertise - process automated
Manual Effort	Requires little manual effort - process is a computerized batch operation	Requires little manual effort - process is a computerized batch operation	Requires little manual effort - process is a computerized batch operation	Requires little manual effort - process is a computerized batch operation	Requires little manual effort - process is a computerized batch operation
Limitations	Street data avail for only 350 urban areas Errors result from non-standard spellings, abbreviations Point sources not accurately represented by address	Accuracy varies with geographical size of ZIP Code Point sources not accurately represented by ZIP Code	Does not perform datum conversions	Limited to batch conversions on PC data files	Degradation of accuracy in remote areas with sparse or nonexistent geodetic documentation
Approach	Centralized	Centralized	Centralized	Centralized	Centralized

coordinates of the nearest street intersection (approximating the latitude/longitude coordinates of the actual entity). The software first identifies the street segment that includes the entity's address in its address range. Assuming that addresses along the street are proportionally spaced by address number, the software then interpolates to identify a point along the street that theoretically corresponds to the entity's address. For example, if a street file has a segment labelled 1000 to 5000, the software will place an address of 1530 at a point 1530 - 1000/(5000 - 1000) = 0.1325 times the length of the segment from the end labeled 1000 and then calculate the latitude and longitude coordinates of that point based on the coordinates at the beginning of the street segment.

This method requires specialized software (a geographic information system) and data (street segments digitized to a standard coordinate system). Specialized GIS software is required because the addresses must be rectified to an x/y coordinate framework. These capabilities are offered by ARC/INFO at the ARC level. The digitized street information, which is usually the most expensive capital component, can be obtained from the Bureau of the Census Dual Independent Map Encoding (DIME) files, covering urbanized areas, and the Topologically Integrated Geographic Encoding and Referencing (TIGER) files currently available in a pre-census format for all 50 states and outlying areas. TIGER/Line, which is the line network of the TIGER system, utilizes cartographic information from Geographic Base Files (GBF)/DIME and from USGS 1:100,000 scale national map series. In addition to the line segments, these files have census geographic codes and address ranges for the right and left side of each segment in metropolitan areas. These detailed files are currently available only for 350 of the largest metropolitan areas.

Additionally, private vendors provide address matching services from proprietary street file data bases. These files are usually enhanced versions of the DIME files, corrected for address range attributes or street trace information. These private vendors often maintain large numbers of existing business entity location coordinates. The private vendors usually can provide a range of solutions for addresses falling outside DIME coverage areas through matching facilities to ZIP code centroid or postal routes. Accuracies vary from vendor to vendor, however, and are not well documented.

Very little manual effort is required to perform address matching because it normally is a computerized batch operation. Assuring quality *input* information, however, could involve significant expenditures depending upon the data base used. Direct costs associated with address-matching consist of acquiring the address data base and the address-matching software. For a typical highly urbanized

area of approximately 180 square miles, the data base cost ranges from no charge for Bureau of the Census DIME files, to approximately $10,000 for commercial conversion software. Of the two options just presented, the commercial data would be of higher quality, because it reflects digitization of street networks from USGS quadrangle sheets and incorporation of DIME file information. DIME files alone do not represent accurate topological structures. Other costs for EPA would result from the need to use ARC/INFO, particularly the NETWORK module. Depending upon the platform used, this software costs anywhere from $10,000 to $80,000, plus the cost of the necessary hardware and training. ***In the studies performed in support of the LATF[1], per point costs to obtain latitude/longitude coordinates through address matching ranged from $7 to $10 with the services of a vendor and between $20 - $30 when performed "in house."***

The accuracy associated with address matching depends upon how the proportionality of a city's addressing scheme, the particular software used, and the quality of the initial data (address information and the baseline street file). Another limitation of address matching is that digitized street data are available only for approximately 350 U.S. urban areas. Additionally, errors during address matching commonly result from nonstandard spellings and abbreviations and from missing or incomplete address data in facility and street files. For example, one company's address may be recorded in a data base as 7562 East 12th Street, S. while the same address is entered as an address matching input as 7562 East 12th Street, South. This inconsistency may cause a "match reject" which would then be placed in a special file for the user to review manually and correct. Front-end data quality efforts or sophisticated address matching software can compensate for abbreviations and other systematic anomalies. Spelling differences and incomplete address information are difficult to rectify, however. Additionally, the data files are not always topologically correct. The line work often has gaps or the address ranges are incorrect or missing.[2] In studies conducted to date, accuracies of latitude/longitude coordinates created from conversion of street addresses have ranged from 5 to 100 meters. This variance is explained by the proportional estimation algorithm used in address matching.

Another major limitation of address matching arises from the fact that addresses do not accurately represent the locations of the actual pipes, stacks, or similar entities which are of environmental concern. The

[1]Four sites (San Gabriel, CA, Chattanooga, TN, Nashua, NH, and Old Southington, CT) with a total of 27 points were analyzed against a control or a known location in a geocoding study performed for the LATF. In the case of the San Gabriel Basin, GPS was the control and other methods were compared to it for accuracy.

[2]Even the highest quality data vendors require the user to sign an agreement acknowledging no guarantee of the integrity of the data.

suitability of using an on-street location to approximate the location of an off-street entity must be determined based upon Data Quality Objectives.

ZIP Code Centroid

Location coordinates also can be assigned based on the centroid of the ZIP code in which the entity is located. ***Using this approach, all entities within a ZIP code district are assigned the same centroid coordinates.*** The sizes of these districts are inversely proportional to population density and may vary from a single building in Manhattan to over 10,000 square miles in Alaska. ***Due to this variation based on population density, the most accurate ZIP code centroid locational coordinate determination can be expected in urban areas.***

Accuracy for this method can range between 50 meters to many kilometers depending upon the size of the ZIP code district. In an EPA test of 191 locations derived from aerial photography in the Chattanooga area, 88% of the facilities had actual locations between one and five kilometers from their respective ZIP code centroid. The median distance was three kilometers. Only one facility was located within 100 meters of the ZIP code centroid.

This method requires little manual effort because calculations can be performed in batch mode on a computer. The cost associated with the ZIP code centroid method is usually between $0.05 and $0.10 per point. This method can only be performed for an entity for which a ZIP code exists and for which DQOs clearly warrant very crude locational accuracy.

7.2.2 Conversion from One Coordinate System to Another

Conversion between coordinate systems occurs through the use of various algorithms that project from one coordinate system to another (e.g., Universal Transverse Mercator [UTM], state grid, etc. to latitude/longitude). Two options are described for converting data between coordinate systems. These are ARC/INFO and General Coordinate Transformation Package (GCTP).

ARC/INFO

ARC/INFO, developed by the Environmental Systems Research Institute (ESRI), is a GIS that can perform coordinate transformation. This GIS, which can be accessed through a micro, mini, or mainframe computer will inter-convert UTM and State Plane coordinates to latitude/longitude coordinates. ARC/INFO is the EPA standard for GIS

software and currently is available on the PRIME computers located in each region.

General Coordinate Transformation Package

The United States Geological Survey (USGS) has recommended the General Coordinate Transformation Package (GCTP) for use by Federal agencies to convert data between geodetic coordinate systems. Use of this PC-based package is endorsed by EPA. GCTP can be used to convert from virtually any coordinate system, including UTM and State Plane, into latitude/longitude. Conversion through GCTP requires very little manual effort as it operates in batch mode. During conversion, the accuracy associated with the original coordinates is maintained. No new error is introduced by the conversion.

A limitation of the GCTP is that it will only do batch conversions on PC data files. For many EPA users, this limitation may mean transferring data from a mainframe system to a PC and back again.

7.2.3 Conversion from One Datum to Another

Two options are described for converting from NAD27 to NAD83. The software that implements the conversion are NADCON and CDATUM.

NADCON

The National Geodetic Survey has developed NADCON software to transform large quantities of existing coordinate information between NAD27 and NAD83. ***NADCON interpolates locational data to convert from one datum to another.*** The interpolation is performed through minimizing total curvature of shift values, where curvature is defined as the rate at which the shifts vary from place to place. Shift information is obtained for more than 100,000 first- and second-order control stations. This conversion method has been programmed in FORTRAN 77 for use on many computing platforms, including PCs.

NADCON preserves accuracy of existing coordinates to approximately 15 centimeters for the conterminous U.S. where geodetic control is good. In remote areas, with sparse or nonexistent geodetic control, accuracies are not as high, but are rarely worse than 1.0 meter. This level of accuracy is adequate for cartographic products with mapping

scales of 1:200 and smaller.[3] The final accuracy of converted coordinates is dependent on the accuracy if the input coordinates.

Data to be converted from NAD27 to NAD83 using NADCON must first be converted to ASCII. NADCON requires very little manual effort as the conversion can be done via batch processing. For further information on this product, contact the Geographic Information Systems (GIS) Program within OIRM or the National Geodetic Information Branch in Rockville, Maryland, at (301) 443-8631.

CDATUM

The designation of ARC/INFO as the EPA standard for GIS software has necessitated the need for software which converts locational data between NAD27 and NAD83 in ARC/INFO format. Accordingly, EPA's Environmental Monitoring Systems Laboratory (EMSL) - Las Vegas developed CDATUM, which uses NADCON algorithms and control files to perform coordinate transformation between datums on ARC/INFO coverages. CDATUM is FORTRAN VII-based and takes advantage of the ARC/INFO subroutine libraries available on the LVGIS VAX which maintains an ARC/INFO object code license.

7.2.4 Guidelines on How to Choose a Locational Data Conversion Method

A detailed process for selection of a geocoding method is given in the *Guide to Selecting Latitude/Longitude Collection Methods* accompanying this document. Suggestions for selection of a conversion method are given here. As discussed in the preceding text, many methods are available for converting existing non-coordinate and coordinate data into latitude/longitude points. ***In all cases, the conversion method used should be identified in the "METHOD" data element.*** Detailed guidance on documentation of the method used to convert lat/long coordinates can be found in Section 2.2.4 of this document. When choosing a conversion method, the following factors should be considered:

- ***Data Quality Objectives (DQOs)*** -- DQOs should be determined before proceeding with the conversion method choice. The DQO process is valuable for identifying the level of required accuracy for a particular entity or application. Once DQOs are established, the decision of whether to convert existing location data or collect new data can be made in accordance with EPA's 25-meter accuracy goal.

[3] "The Impact of the North American Datum of 1983 on Cartographic Products," Federal Digital Cartography Newsletter, Summer 1990, p.6.

- *Type of conversion needed* -- Does the existing data base have address information, UTM, or State Plane, or are latitude and longitude coordinates available but need to be converted between datums? The answer will determine what type of software is needed.

- *Type of entity for which the information has been collected* -- Is the location for a facility, an outfall, a smokestack, etc.? This answer will assist in determining the feasibility of certain conversion methods, such as address conversion software.

- *Location of the entity* -- Is it in a major metropolitan area or a rural area? Again, this criterion will determine the applicability and expected accuracy of certain methods, such as address conversion software.

- *Available resources* -- What resources (personnel, time, and funds) can be allocated to addressing LDP requirements? This consideration will help decide whether to collect new data or to apply a conversion method.

- *Hardware environment* -- Does the organization have access to the mainframe or only PCs? This answer will help in deciding whether certain software will be usable by a particular organization.

Program managers should take these factors into account when deciding which conversion method to use for non-LDP-conforming data. In certain cases, these factors may indicate that it would be more efficient and effective for an office simply to collect new location data.

7.3 Collection Methods

A wide variety of methods exists for collecting new latitude/longitude coordinates. The following section provides a brief description of many of these methods as well as a discussion of the level of expertise required, limitations, accuracy, time, and cost associated with each.[4] Additionally, general guidance is given on how to choose a collection method. The methods discussed in this section are summarized in Exhibit 7-3. Cost and accuracy associated with each method are illustrated in Exhibit 7-4.

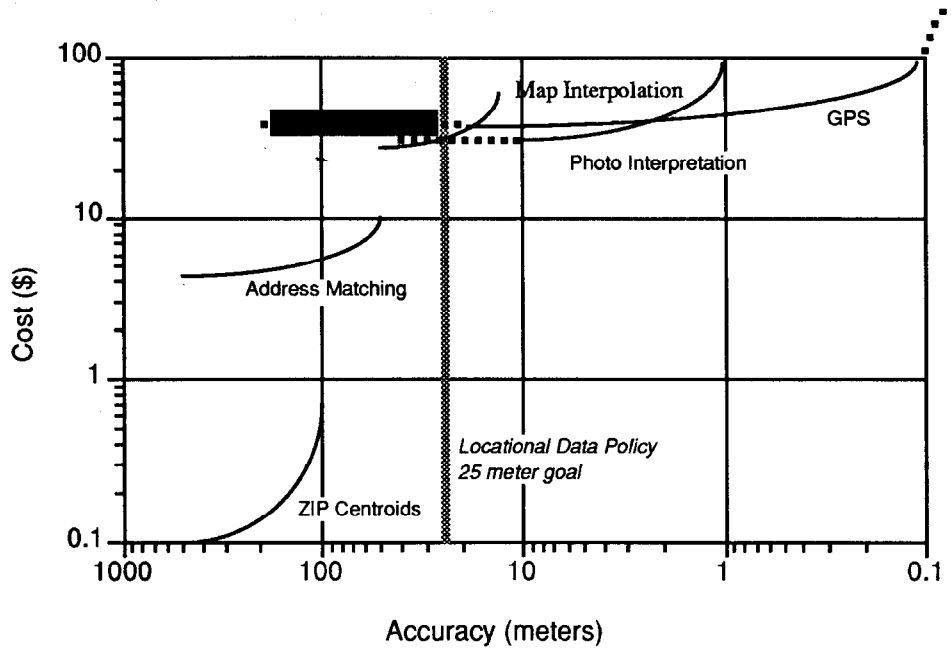
[4]Much of this information was obtained from: Fitzsimmons, Charles K. *"Evaluation of Selected Methods for Determining Geographic Coordinates."* and *"Description of Location Data Collection Techniques"*, Environmental Research Center, University of Nevada, Las Vegas.

Exhibit 7-3

Comparison of Locational Data Collection Methods

Method/ Characteristic	Conventional Surveying	Map Interpolation	Photo Interpolation	LORAN-C	GPS
Description	Traditional method of measuring horizontal distances, elevations, directions, and angles for the purpose of gathering locational data	Geopositioning by means of maps through either manual efforts or digitization	Interpolation performed either manually or through digitization of aerial photos/ satellite images	Location system using broadcasting stations and receivers to determine latitude/ longitude	System of satellites, ground stations, and receivers which provide global position infromation
Entity Type	Any entity within terrain limitations	Any entity which can be located on a map	Any entity which can be located by remote sensing/ aerial photography	Any entity within a LORAN- C coverage area	Any entity within the GPS satellite coverage area
Accuracy	Very high - to the centimeter- level	25-1000 M depending on map scale and skill of interpolator	5-300M depending on resolution and scale	50-150M with correction factors applied	100-200M for autonomous selective availability; 15-100M for differential mode
Cost/Point	High	$20-$60/Point	$40-$60/Point	Receiver cost of approx $1500 $100/Point	Receiver costs range from $3000-$150000 $40-$170/Point
Required Expertise	Requires trained, licensed surveyor	Requires basic map knowledge	Requires trained photointerpreter	Requires knowledge of LORAN receiver, publicationss, and correction methods	Requires technician with knowledge of GPS receiver.
Manual Effort	High. Each site must be visited to obtain a reading.	High.	Moderate to high.	High. Each site must be visited to obtain a reading.	High. Each site must be visited to obtain a reading.
Limitations	Limited use in rough/ inaccessible terrain Requires high level of expertise	Limited in accuracy by map scale, accuracy, orientation, and experience of map interpolator	Limited in accuracy by photo, optical, and perspective distortion, relief displacement, and experience of photointerpreter	Limited coverage in most inland regions Susceptible to bad weather/ atmospheric disturbances	Satellite availability limited Accuracy degraded under selective availability
Organizational Considerations	Incremental/ Centralized	Incremental/ Centralized	Incremental/ Centralized	Incremental/ Centralized	Incremental/ Centralized

Exhibit 7-4
Cost vs. Accuracy Curves of Lat/Long Collection Methods



Goal:
25 m or better

7.3.1 Conventional Surveying

Conventional surveying is a traditional geocoding method that requires precise measurements by a licensed surveyor. There are many types of conventional surveys, such as cad astral surveys which are performed to establish legal and political boundaries for land ownership and taxation purposes, and geodetic surveys, which are global surveys performed to establish control networks used in land mapping. ***Surveying can achieve very high accuracy - to the centimeter level.*** If surveying is used in concert with other more sophisticated techniques, such as photogrammetry, very high accuracies can be virtually assured. There are some limitations to surveying, however. ***For example, surveying cannot be performed easily in very rough terrain and is, therefore, most appropriate for accessible areas.*** **This method also** ***necessitates a trained, sometimes licensed person to perform the survey. Additionally, due to the labor intensive nature of surveying, costs per point tend to be quite high.***[5]

7.3.2 Map Interpolation

Map interpolation is a method of geocoding measuring relative distances using maps, usually USGS topographic quadrangle sheets. Coordinates of a point on a map are determined either through *manual* or *automatic* interpolation. Topographic maps are printed with 16 map tics per quadrangle: four corner tics and 12 interior tics. On a 7.5 minute map, tics are placed every 2.5 minutes. To determine coordinates for points of interest manually, a scale (such as a bar scale, an engineer's scale, or a graduated ruler) that matches some increment of latitude and longitude on the map is used to measure the x and y distances from the relevant point to the nearest map tic. The scale values are then multiplied by the appropriate factors to obtain the degrees, minutes, and seconds for latitude and longitude and are added to the coordinates of the map tic.

Coordinates also may be obtained automatically through *digitization*. Digitization involves the transformation of paper maps into a computer-readable format through marking or plotting on a map and then using a digitizing table to input relative positions of points, lines, or polygons with unknown coordinates. The digitization software extrapolates the location of the points marked on the map by comparing it to the known coordinates of the map corners and calculating the distance and direction of the point from the corners. The coordinates of each point are then stored in the computer.

The accuracy associated with map interpolation can be no greater than that associated with the map itself and is usually less due to the potential error introduced by the interpolation process. National Map Accuracy Standards

[5]American Management Systems, Inc., "Geopositioning: Techniques, Technology, & Services (Draft)" October 24, 1988, p.4.

maintained by USGS specify a maximum horizontal deviation from the true location of not greater than 0.02 inches on a map. The relative accuracies of 7.5 minute quadrangle maps, therefore, are as follows:

- A 1:24,000 scale map - 40 feet.
- A 1:63,000 scale map - 105 feet.
- A 1:100,000 scale map - 167 feet.
- A 1:250,000 scale map - 417 feet.

For the highest accuracy, maps made from stable materials, such as Mylar, should be used instead of paper, as this removes potential changes in map dimensions due to humidity and temperature. The largest source of error in map interpolation occurs in the field, when difficulties can be encountered while orienting the map to observed entities and then visually interpolating the position on the map. If the interpolation is not performed in the field (office-based), the point or facility must be located on the map using an address, aerial photograph, or other source of information. In the *Guide to Selecting Latitude/Longitude Collection Methods*, the achievable accuracy for map interpolation using USGS 1:24000-scale maps ranges from 12-50 meters for easily identifiable features. Larger scale (i.e., higher resolution) maps will yield accuracies in the 5-25 meter range.

Map interpolation can be utilized for any entity that can be clearly located on a map. The USGS 1:24,000 scale, however, is usually the smallest reasonable scale for locating sub-facility points, such as stacks and pipes. ***One potential drawback associated with this method is the inability of the user to locate a site accurately due to the scale or date of map being utilized.*** For example, an old or outdated map may not depict the entity or show other easily recognizable features.

According to the Guide to Selecting Latitude/Longitude Collection Methods, costs for office-based map interpolation ranged from $40 to $60 per point[6] while field-based efforts range from $28 to $40 per point. The major component of this cost is labor and varies depending upon the interpolator's familiarity with map products and scales. ***In general, map interpolation requires a moderate to high level of manual effort.*** Although a trained specialist is not required to locate items on a map, the interpolation of points from USGS topographic maps requires at least a basic knowledge of map concepts, including symbology and measurement. Another consideration of map interpolation cost is the expense of buying the maps themselves. Paper maps from USGS usually cost less than $5.00 each while Mylar maps cost $56.00 each.

[6]This cost figure includes an estimate for research time.

7.3.3 Photo Interpretation

Photo interpretation may be performed either on aerial photographs or satellite images in a manner similar to that used with maps. Entities are identified on the photographs or images, the geographic positions are rectified to a map, and coordinates are either manually or digitally interpolated using one of several techniques. One technique involves locating an entity through the use of ground-based reference points, such as roads and streams. Additionally, a zoom transfer scope (ZTS), which optically superimposes a view of the photograph on a map, can be used to interpolate. It is important to note that photo interpretation must be used in conjunction with a map or some other device that enables the user to rectify points on the photo to coordinate planes.

The main advantage of photography and imagery over maps is that they are usually more current than maps. Aerial photographs and images can usually be acquired within 24 hours, limited by weather conditions and the source utilized. Maps usually are several years old at best.

Color infrared film at 1:24,000 scale is preferred for aerial photography. The National High Altitude Photography Program maintains an extensive U.S. coverage at approximately 1:80,000 scale, and new photographs at a 1:40,000 scale are becoming available. The process through which satellite images are obtained is referred to as *remote sensing*. Satellite imaging technology records the reflection of both visible and invisible light from the earth's surface and then converts these values to computer-readable format.

In spite of the apparent "realism" in photogrammetry, this method also has its limitations. Photographs vary in vertical scale across their extent whereas maps do not. They exhibit *"relief displacement"* in which closer features have larger images than farther features. The tilt angle and focal length of the camera also introduces a *"perspective distortion."* Such distortions increase from zero at the center of the image to a maximum at the edges of the image. Special equipment, such as an analytical stereo plotter, is used to produce maps from aerial photographs. This equipment removes the relief and tilt distortion from the source photographs.

In addition, there are no grids or tics on the photos by which to reference the coordinates of a plane. Therefore, coordinates must be calculated from photos with additional information including known reference points, current maps, addresses, elevations of the points in question, elevations of the camera, and the focal length of the camera. Points of interest on the photos can also be digitized as in map interpolation.

The accuracy achieved from photo interpretation is dependent upon the resolution and scale of the photograph as well as the techniques used to remove relief displacement and perspective distortion. This method can be

extremely accurate with an average accuracy under 5 meters for most common photo scales. To further improve the accuracy associated with photo interpolation, aerial photography can be merged with GPS as a "hybrid system." Hybrid systems are discussed in more detail in Section 7.3.6 of this document.

Photo interpretation requires a moderate to high amount of manual effort. For example, EPIC[7] has outlined the planning steps for an analysis of aerial photography. These steps are geared toward hazardous waste sites and may be more detailed than necessary for other programs. The following basic steps would be required for most locational data collection projects using photographs or remote satellite images:

- Obtain appropriate maps
- Prepare flight plan
- Search for archive photos and order the relevant photos
- Contract flights
- Check photos for quality
- Search for points or areas of interest
- Use maps and address information to identify entities
- Prepare data for transfer to other data systems as required.

Photo interpretation can be used for any entity that can be located on an aerial photograph for a cost of between $40 and $100 per point. Aerial photographs, not being labeled and not using the easily recognized symbology of maps, require a trained photointerpreter to perform the interpolation, adding a significant cost to this method. In addition, photo interpretation can be used only for entities that can be located on aerial photographs.

7.3.4 LORAN-C

LORAN-C (Long Range Navigation) is a location system which is usually used to aid ships and planes in navigating. LORAN employs broadcasting stations and portable receivers to determine latitude and longitude. These broadcasting stations are arranged in triads where one is the *"master station,"* and the other two are *"slaves."* Each station broadcasts signals at specified time intervals. Coordinates are then determined by calculating the difference in the amount of time required for a signal to reach a receiver from the master and slave stations. This calculation allows a distance hyperbola to be determined, thus giving the distance of the receiver from the slave station. Calculating the distance hyperbolas for both slave stations provides the relative location of the receivers through triangulation. Because the LORAN master station's absolute location is known, the receiver's absolute location can be computed. LORAN coverage is available with at least limited

[7]See Appendix B - Contacts for LDP Implementation Assistance

reception for most of the northern hemisphere. Locations may be calculated at distances up to 1500 km over land and up to 300 km over water.

According to the Federal Radionavigation Plan, accuracy for a LORAN at sea is 436 meters with a repeatable accuracy of between 19 and 90 meters. In U.S. Coast Guard studies, 150 meter accuracies have been observed 600 km from the master station. EPA Region IV and EMSL-LV tested LORAN in December of 1987 in the Chattanooga study. The researchers found an uncorrected accuracy of over 600 meters from the true locations. When correction factors (calculated from 59 absolute locations) were applied, accuracy improved to between 5 and 153 meters from the true locations. A discussion of LORAN-C was not included in the *Guide to Selecting Latitude/Longitude Collection Methods*.

LORAN can be used to obtain coordinates for any location within a coverage area and by any individual with a LORAN receiver. However, LORAN is most effective in coastal regions and at sea, and is relatively ineffective in inland regions. LORAN was designed for maritime usage, therefore most LORAN transmitters are located in coastal regions. Also LORAN frequencies are assigned to take full advantage of propagation over water. LORAN ***performs best on clear and dry days and worst on cloudy and wet days.*** In fact, the major source of error in LORAN measurement is caused by atmospheric disturbances, such as electrical discharges, or by power-generating equipment. These disturbances, which are most likely in urban areas, warp distance hyperbolas and cause incorrect calculations in the receivers' relative locations. These errors occur up to an order of +927 meters (0.5 minutes) and remain relatively constant over several hours. ***To greatly reduce error, correction factors can be calculated from measurements obtained at sites where the absolute location is known.*** Errors in LORAN measurement can also be introduced by weak signals and by the relative position of the receiver to the LORAN stations. If a receiver is close to high elevations or other landforms which block signals, accuracy will decrease.

Coordinates can be obtained for approximately $100 per point with labor being the greatest component of this cost. ***Manual effort required for LORAN is very high*** as a technician must participate in survey planning activities, survey area reconnaissance, individual site visits and measurements, and post-processing. In Chattanooga, 70 locations were surveyed in 220 hours with an average rate of three hours per point (including the time required to calibrate the receiver and to perform repeated measurements at one point). ***Another component of cost is the expense of a LORAN receiver. A single LORAN receiver typically costs around $1000.***

7.3.5 Global Positioning Systems (GPS)

A detailed primer on GPS accompanies this document. A brief introduction to GPS is given here.

The NAVSTAR GPS (Navigation System with Timing and Ranging Global Positioning System), a system of satellites, ground stations, and receivers, was designed to provide 24 hour global, three-dimensional position information in real-time to the Department of Defense (expected to be fully deployed in June 1992). When fully operational it will be possible to obtain nearly instantaneous positions of aircraft, land vehicles, and ships 24 hours a day, any place in the world, from a constellation of satellites using GPS technology. This method, which avoids many of the problems inherent to land-based methods will also support civilian applications with a slightly lower accuracy.

Portable GPS receivers compare coded time and position data broadcasts from the satellites to calculate locations. Each of the satellites in the network broadcasts two types of radio signals, C/A-code and P-code, on preassigned frequencies. Signals must be received from at least four satellites simultaneously to unambiguously calculate location in three dimensions (latitude, longitude, and elevation). If the elevation of a point is already known, then only three satellites are needed. Currently, 19 out of the planned 27 satellites are operational. This limited deployment restricts the acquisition of satellite signals and affects the accuracy and availability of GPS data. Studies have shown that there is a high degree of variance introduced if a receiver shifts its "lock on" from one set of satellites to another. Therefore, due to the limited satellite availability, significant time must be spent planning field work around satellite visibility. This problem will diminish in 1992, with scheduled deployment of the full constellation of satellites.

Two types of GPS receivers exist: navigation units (C/A code receivers) and the more precise geodetic units (P-code receivers). Geodetic receivers are more precise because they receive a much more complex signal and are capable of assessing signals from more than four satellites at a time (centimeter accuracies are expected with this type of receiver when the system is fully deployed).

GPS can be used to locate positions on the surface of the earth in two modes: *autonomous* or *differential/relative*. Autonomous positioning utilizes only

one receiver at the point of interest and depends upon the ephemeris[8] broadcast by satellites. ***Higher accuracy is possible through the use of the differential mode, where two or more receivers are used simultaneously - one at a point of known coordinates and the other at the point of interest.*** The receiver at the known benchmark corrects errors in the satellite's transmitted signal, including those purposely introduced by the military to degrade the effectiveness of GPS technology in the civilian community. The coordinates of the point of interest are determined through simultaneously post-processing the receivers' measurements by computer. The accuracy is then a function of the baseline distance between the two receivers and ranges between two and five million parts per million of the baseline. For example, if the two receivers are 160 kilometers apart, the expected deviations in accuracy are less than 11 meters. Tests have found that centimeter level accuracy is possible through the use of two receivers taking 45 to 60 minute static measurements. Millimeter accuracy is achievable if post-processing utilizes ephemeris data from the National Geodetic Survey (these data are available approximately two weeks after a GPS reading is obtained).

The military reserves the right to selectively degrade or even to deactivate the NAVSTAR system to protect national security interests. Degradation, which is referred to as *"selective availability,"* implies that the satellite signals are programmed to provide erroneous locations to Earth receivers. The Department of Defense guarantees, however, that 100 meter accuracies will still be achievable under selective availability. This condition may be overcome by the use of differential correction and post-processing software.

The accuracy associated with GPS is quite high, usually within +/- 25 meters. Under optimal conditions and with the use of a geodetic-quality GPS instrument, accuracies can be obtained to better than a centimeter. The accuracy depends upon the type of GPS receiver being used. For example, a high-end geodetic quality instrument (P-code receiver) can achieve extremely high levels of accuracy while a lower end navigation quality instrument achieves lower levels of accuracy. Accuracy can be improved for civilian receivers by performing *static* observations. When all satellites are deployed in mid-1992, GPS (in differential mode) should regularly be able to achieve better than 10 meter accuracies.

Although accuracy is not affected by radio frequency interference or weather, a clear line of site (LOS) from the receiver to the satellite must be maintained (i.e., satellite must be a minimum of 20 degrees above the horizon). Another limitation of GPS devices currently is the ***limited satellite availability or "window"*** due to the incomplete GPS constellation.

[8]Ephemeris is the prediction of current satellite position, given in a table or transmitted to the user in a data message.

The average collection cost per point, as estimated in the *Guide to Selecting Latitude/Longitude Collection Methods*, ***was between $50 and $100 for centrally planned, dedicated GPS surveys. For decentralized GPS surveys (i.e., performed by field staff already on-site conducting other duties) the cost per point was between $35 and $70.*** These cost figure include a large amount of labor because a ***high degree of manual effort is needed for GPS.*** A technician must be sent to each site.

It must be emphasized that, in its present state, GPS is not a "push button" operation. ***Technicians must be adequately trained in pre-survey, survey, and post-survey procedures. The technician in the field must be constantly aware of signal strengths, satellite availability, noise sources, and topography. When a survey is completed, a tremendous amount of data must be sorted and screened before an optimal location solution is determined*** . Additionally, GPS is quite expensive with C/A-code navigation receivers ranging from $3000 to $15,000, and P-code geodetic quality receivers starting at $30,000.

7.3.6 Hybrid Systems

Hybrid systems are defined in this context as the use of a combination of technologies. LORAN, GPS, and/or inertial technologies have been incorporated into several devices that seek to capitalize on the strengths of each system. An example of a hybrid system (i.e., map interpolation/photo interpretation) is discussed in Section 7.3.3 of this document.

Some hybrid systems offer the advantage of using the GPS signal during satellite availability and of using a back-up technology when the window closes. These hybrid systems are subject to the same types of errors as LORAN and GPS. Additionally a significant post-processing problem occurs if both technologies were used for a single mapping session. Error sources are different for each technology and result in significant differences in locational accuracy which must be resolved. There is little data available on these systems, but accuracies and variable costs are thought to be similar to those experienced with GPS and LORAN. Additional costs will be incurred if significant post processing is required due to a shift in technologies during a mapping session. Costs for hybrid units range from $3,000 to $10,000.

7.3.7 Guidelines on How to Choose a Locational Data Collection Method

As discussed in Chapter 3, GPS is the preferred technology for acquiring location identification data. Guidance for employing GPS is given in the GPS Primer accompanying this document. Technical assistance for GPS can be found within the EPA Regional GIS/GPS Work Groups.

GPS may not be suitable for all circumstances. A process for determining the most suitable locational data collection method is given in the accompanying *Guide to Selecting Latitude/Longitude Collection Methods*. Another method may be found to be more affordable, or the circumstances may warrant less stringent accuracy requirements. ***In all cases, the collection method used should be identified in the "METHOD" data element***. Detailed guidance on documentation of the method used to collect lat/long coordinates can be found in Section 2.2.4 of this document. If a method other than GPS is deemed necessary and supported by clearly defined DQOs, the following factors should be taken into consideration:

- ***Type of entity for which the information is being collected*** -- Is the entity in question a facility, an outfall, a smokestack, etc.? This information will help managers decide which type of method can be used. For example, only certain scales of maps can be used for locating sub-facility points (e.g., stacks or pipes). Photo interpolation can be used only for entities that can be located on aerial photographs (e.g., not under ground or under tree-cover).

- ***Location of the entity*** -- Is the entity located in a major metropolitan area, in a coastal region, or surrounded by rough terrain? This consideration will help determine which method is most appropriate. For example, LORAN is effective in coastal regions. Surveying cannot easily be performed in rough terrain.

- ***Program requirements for accuracy*** -- What level of accuracy is the program attempting to achieve? The level of accuracy required in a program may dictate use of a certain type of method, especially if the accuracy requirements are stringent. EPA's locational policy goal is to attain better than 25 meter accuracy for all locational data by 1995. The point of reference for accuracy should be the most visible and accessible point of the entity for which lat/long is being measured. This point should be documented in the "Description" data field.

- ***Available resources*** -- What resources (personnel, time, and funds) can be allocated to the collection? Resource availability may determine selection of certain methods. GPS requires a high degree of manual effort and also is quite expensive. It may not be a viable option where resources are limited.

- ***Necessary expertise*** -- Does the program have individuals who possess spatial data collection expertise, will they have to train existing personnel, or will they rely on contractor expertise? Decisions relating to personnel expertise and training affect overall costs and, therefore, desirability of certain methods.

- *Provider of the information* -- Who provides the information (i.e., the facility operator, EPA staff, etc.)? The response to this question has implications for the type of method used because certain types of equipment may not be accessible to all parties and novice data providers may require EPA to draft extensive, clear instructions on use of the method.

Through responding to these questions, referring back to the text and the summary tables in this document, and consulting EPA technicians in the GIS Program, at EMSL-LV, and at EPIC, project managers should be able to determine which of these collection methods best suits their needs and Data Quality Objectives. A process to follow for choosing a location identification method is provided in the *Guide to Selecting Latitude/Longitude Collection Methods* accompanying this document.

Chapter 8
SYSTEM IMPLICATIONS

8. SYSTEM IMPLICATIONS

The Locational Data Policy sets forth requirements for collecting and documenting lat/long coordinates for entities regulated and/or tracked under Federal environmental law. The policy does not specifically address the automation of locational data after they have been collected and documented. Because the entities, processes, and systems vary so greatly among environmental programs it is difficult to generalize how data should be computerized. Some guidelines, however, are provided for managing locational data in computer systems. These guidelines promote uniformity of locational data and expedite their access by primary and secondary users. Therefore, the formats for locational data are suggested in this chapter, and are recommended for both computer (and paper) records.[1]

8.1 Location Identification Data Management

There are many issues relating to data automation that should be addressed even before lat/longs are collected and documented. With the ultimate objective being high quality, accessible locational data, uniformity in the management of computerized resources is necessary.

8.1.1 Relationship Between LDP Data Elements

The Locational Data Policy requires that four elements be collected: latitude and longitude, method, accuracy, and description. Definition for these elements are as follows:

- ***"Lat/long coordinates" are a repeating <u>set</u> of numeric fields*** to allow a point, line, or area to be defined.

- ***"Method" is a set of coded alphanumeric fields with three components*** (see Chapter 2) ***and is a descriptor of the technology used to collect lat/long coordinates, the datum, and the map scale.***

- ***"Accuracy" is a free-format, numeric field with prescribed components that describe the accuracy range of the locational data*** (Chapter 2); the accuracy field may be linked in a computer system to the method field so that accuracy values (ranges) can be derived based upon the type of method used.

- ***"Description" is a free-format, textual field containing certain types of information describing the exact place where lat/long was determined and whether a point, line, or area is being defined.***

[1]Names for LDP data elements are specified in Section 5(d) of the LDP.

8.1.2 Structure and Format of LDP Data Elements

The following recommendations promote consistency in the structure and format of the Agency's locational data.

Latitude and Longitude

Lat/Long data elements are used to describe the geographic location of the point being measured. The following statements apply to these fields:

- ***Latitude and Longitude are to be <u>two separate data elements</u>***
- ***Latitude is to be presented, in all system outputs, before longitude***
- ***There are to be no embedded blanks*** between values of degrees, minutes, seconds, and/or fractions of seconds (e.g., +DDMMSS.SSSS)
- ***Each measurement (i.e., latitude AND longitude) should have a symbolic hemisphere qualifier preceding it ("+" or "-")***; those values MUST NOT be letters (N, S, E, or W); if no qualifier precedes a coordinate, that coordinate should be assumed to be a positive number (in accordance with common mathematical convention)
- ***Values for degrees, minutes, and seconds that are less-than-ten should be padded with leading zeros***; thus a value of six minutes should be given as "06." Values for degrees longitude that are less than one hundred should similarly be preceded by zeros. *Ninety-seven degrees, 11 minutes, and 6 seconds west longitude* should be encoded as -0971106.
- ***The decimal point need not be stored as a character*** if the system internally justifies the coordinate values appropriately; ***the decimal point must appear, however, in all outputs.***

Method

Method is a set of three fields describing how the lat/long coordinates were determined. Exhibit 2-2 summarizes the codes that should be used to document *method*.

- ***The method data element should allow up to 10 alphanumeric characters representing the specific code for method and/or tool used (Exhibit 2-2).***

- ***The data element for datum should be structured as two numeric digits to communicate the year of the datum used (e.g., 83).***

- ***The data element for scale should be the "x" value of the 1:x representation of the scale (e.g., "100,000" for a 1:100,000 scale map).***

- ***The method fields are left-justified (one reads from the left) and should have no embedded blanks to separate components.***

Agency-wide standardization of method codes is imperative because this field will often be used to select locational data for secondary use. EPA GIS program managers will maintain the complete list of codes to be used to document *method*.

Accuracy

The *accuracy* data element is a numeric value that conveys the range within which the true location is expected to fall with 95% confidence. It is expressed in the next lower unit of measurement than that in which the coordinates are given (e.g., if the coordinates are to the *second*, the accuracy should be in *tenths-of-a-second*). Accuracy determined in meters must be converted to degrees of latitude/longitude. The accuracy estimate will help secondary users determine whether certain locational data can meet their needs.

The components of the ***accuracy*** data element are:

- Numerical value of the range

- Units, documented as:

  - deg for degrees
  - min for minutes
  - sec for seconds.

When values of accuracy are displayed, they should be preceded by a "+/-" to indicate a range. The accuracy field should have no embedded blanks. Examples of accuracy may be "+/- 1 sec" for a coordinate set of "+292210,-1090720" or "+/-0.05 sec" for a coordinate set of "+292210.1,-1090720.6." ***In cases where the accuracy of the latitude is known to differ from the accuracy of the longitude, a single accuracy estimate***

should be provided (the least accurate of the two[2]). For example, if the accuracy of the latitude value is +/- 0.5 second but the accuracy of the longitude value is +/- 1 second, the accuracy for the whole set of coordinates is "+/-1 sec," the least stringent of the two accuracy values.

Description

The *description* data element is a free-format, textual (alphanumeric) field of unspecified length that should communicate clearly two pieces of information:

- ***Exactly what the lat/long coordinates are of,*** NOT what they represent (e.g., the coordinates may be *of* the front door of a facility but are being used to *represent* the location of the whole facility)

- ***Whether the coordinates define a point, line, or area;*** that is to say, whether there is one or more-than-one pair of coordinates in the locational definition. This clarification allows the data reviewer to know that there may be several sets of coordinates defining an area, and not just one set.

Description need not be longer than a phrase. For example, a description might read as "discharge point at end of pipe #1" or "1st of 7 coordinate pairs starting at the northwest corner defining site boundary."

8.1.3 Position and Function of Location Identification Data

Locational data must adequately serve two functions within a data base: (1) they must define clearly the location of the entity they represent, and (2) they must allow data users to select among them to secure only data of suitable quality for their use. The position and function of the locational data in the computer system must support these two needs.

Position of Location Identification Data

Location identification data must be clearly associated with the entity they represent. For data bases about places, the locational data should be a direct attribute of the entity that they represent. This association might be accomplished by storing location identification data in the header record or in the "C0" level, or wherever they will be clearly linked to the entity whose location they define.

[2]In all cases, the goal for accuracy should be better than 25 meters.

In some situations, it might be appropriate to store multiple sets of location identification data if several tiers of entities are tracked in the same data base. For example, if a data base identifies smokestacks at a facility, the facility header-record should contain the location information of the facility-in-general, but the sub-records in the data base about each stack should contain locational data for each stack. Another possibility is that the locational data might be associated with a sub-unit of a complex entity and not with the entity-in-general at all, especially if the sub-unit is the critical unit for the program.

Function of Location Identification Data

Users must be able to use location identification data in several ways:

- ***To search through and select entities*** in the data base based upon their location
- ***To screen out unwanted entity records*** based on the characteristics of imbedded metadata (e.g., too inaccurate or unknown method).

System managers may want to review the current functional capabilities offered to users of locational data in their system. If at all possible, they should enhance their systems to offer selection and screening capabilities. Merely incorporating storage of new locational data elements is not enough to promote secondary use of the data. Reports and queries must be modified because these enhanced system capabilities will allow users to reap the benefits of the effort invested in locational data collection.

8.1.4 Relationship to Other Locational Data

There are many other spatial and locational identifiers used in environmental protection, such as address, county, hydrologic unit, and UTM coordinates. The Locational Data Policy does not preclude these alternative location definitions. Rather, ***the information required by the LDP is in addition to, NOT instead of, other locational data normally collected.*** Other types of location identification data may be useful as a double-check. For example, lat/long coordinates can be screened for validity by comparing against county boundary when "county" location is identified for an entity.

8.1.5 Edit Checks

Installation of edit checks into system data entry screens or software will help ensure that locational data are realistic and complete. Edit checks can be useful in double-checking data accuracy, and in verifying the format and quality of data submitted by the regulated community.

Various types of edit checks already are available in many EPA systems. For example, STORET compares lat/long coordinates to county boundaries and flags improbable coordinates. Other systems compare lat/long coordinates to distance from the ZIP code centroid. Some systems do not allow addition of a record to the data base if lat/long coordinates are missing. System managers may want to devise one or more edit checks on their system depending on their data update process and the types of other location identification data available for each entity.

8.2 Locational Data Sharing

To promote secondary use of data and reduce costs for new data collection, OIRM currently is examining options for sharing locational data in computerized systems. Although some of the options are still in the conceptual phase, a brief introduction is given below.

8.2.1 State Entry in FINDS

FINDS is EPA's inventory of facilities regulated under Federal environmental law. On a pilot basis, EPA is allowing states to enter data directly into FINDS. Some states are entering information into FINDS records for facilities that are regulated only by the states.

FINDS contains basic identification data about regulated facilities and pointers to other sources of more detailed program data. Information in FINDS includes EPA facility ID code, name, address, and locational data. At present, FINDS has only one set of lat/long coordinates per facility in its master record file, but additional lat/longs (copied from National Program Systems) can be stored in its alias file.

8.2.2 GRIDS

The Geographic Resources Information and Data System (GRIDS) is designed to help users share baseline spatial data and to provide access to ARC Macro Language (AML) for use in ARC/INFO Geographic Information System applications. The main goal of GRIDS is to reduce costs associated with development of GIS applications within EPA by providing access to commonly needed spatial data sources and tools. GRIDS will provide the following data sources:

- USGS Hydrography Digital Line Graphs - 1:100K.
- USGS Transportation Digital Line Graphs - 1:100K and 1:2M.
- USGS Geographic Names Information System (GNIS) data.

- Defense Mapping Agency Digital Elevation Model (30 second data).
- USGS Land Use and Land Cover Data.
- USGS Hydrologic Basin Boundaries.
- USGS Federal Land Boundaries.
- US Bureau of the Census Pre-Census TIGER Line Files.
- US Bureau of the Census Master Area Reference File (MARF) 2.
- US Bureau of the Census Summary Tape Files (STF).
- US Fish and Wildlife Resources Area Geographic Database.
- US EPA Reach File, Version 3.
- US EPA Ecoregion Boundaries.

The AML Library module allows users to extract AML routines created and used by other EPA program offices. GRIDS is available in prototype on EPA's IBM 3090 mainframe in the line mode (TTY). For further information, contact Bob Pease at 703-557-3018.

8.2.3 National Spatial Data Transfer Standard

The Spatial Data Transfer Standard (SDTS), developed by the USGS National Mapping Division, 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 developing the standard is to promote and facilitate transfer of digital spatial data between dissimilar computer systems. Section 4.1.3.5 of the SDTS provides the specifications for latitude/longitude coordinates.

Appendix A

LDP AND LATF RECOMMENDATIONS

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

APR 8 1991

OFFICE OF
THE ADMINISTRATOR

MEMORANDUM

SUBJECT: Agency Locational Data Policy

FROM: F. Henry Habicht II
Deputy Administrator

TO: Assistant Administrators
Regional Administrators
Associate Administrators
Inspector General
General Counsel

Through an extensive consensus-building process involving all Agency offices and Regions, and many of our State and federal partners, the Agency has developed a Locational Data Policy. The policy is designed, both literally and figuratively, to put EPA data on the map.

The policy requires consistent use of latitude/longitude coordinates throughout the Agency to identify the location of entities that concern us. It also sets a goal, which is now technically achievable, that all such coordinates be accurate to within 25 meters.

I believe that your compliance with this policy, and with the Facility ID Data Standard, will do more to enable data integration than almost any other initiative undertaken during my tenure at the Agency. I urge you to read the attached policy and implementation guidance and ensure that your organization aggressively implements the policy.

You may direct questions on this policy and guidance to Al Pesachowitz, Director of the Office of Information Resources Management, at FTS-382-4465.

We are making historic progress in data & knowledge integration — Thanks for your ideas, attention and support

cc: Senior IRM Officials
National Program System Managers

Attachments
1) Locational Data Policy
2) Locational Data Policy Implementation Guidance (draft)

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 latitudes north of the equator and longitudes east of the prime meridian

- **-** specifies latitudes south of the equator and longitudes west 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.

Locational Accuracy Task Force

Findings and Recommendations

Prepared for:
IRM Steering Committee
U.S. Environmental Protection Agency

Prepared by:
Locational Accuracy Task Force
U.S. Environmental Protection Agency
Jack McGraw, Chairman

December 13, 1990

Executive Summary

The Locational Accuracy Task Force (LATF), over a period of six months, collected and weighed a considerable amount of information on geocoding technologies and programmatic requirements in order to reach the following conclusions:

- More analysis needs to be done to determine the accuracy and costs of geocoding methods, yet it is clear that the future holds the greatest promise for GPS technology;
- To achieve cross-media integration, some EPA programs, regions, states, and other Federal agencies have already taken significant steps to develop a locational accuracy policy;
- Unless a clear goal is stated, the data in the Agency's data bases will continue to be a "mixed bag";
- A strategy is needed to help individual programs comply with the proposed recommendations;
- No matter which geocoding method is chosen, the high costs of equipment, training, and data collection will make the implementation of a standard expensive for the Agency;
- Acquiring or updating locational data for old data in order to meet a standard will be too costly and difficult;
- Incentives, such as technical transfer and data sharing with state and local governments, will be more effective than enforcement measures in attaining accuracy goals;
- Steps must be taken immediately to populate FINDS with locational coordinate data using the most reasonable method available; and
- The consensus of the LATF is that accurate locational data is essential to risk management and multi-media decision making.

These conclusions led the LATF to formulate five recommendations for the IRM Steering Committee:

1. Establish a 25 meter goal— The Agency should set a 25 meter or better (±1.0 second) level of accuracy. This goal would apply to new data only. New data is considered that which is collected after the policy is enacted and would

EXECUTIVE SUMMARY

include data collected by site visits, supplied by facilities, or collected to support special programmatic needs.

2. Set GPS as the standard— The Agency should set Global Positioning Systems (GPS) technology as its standard. The Agency should concentrate on the large scale acquisition of equipment and on providing training. In the short term, the Agency should support map interpolation.

3. Define a deviation process— The Agency should define a process for programs to request exemptions from the policy. Using the Data Quality Objectives (DQO) process, programs obtain a waiver from either the latitude/longitude policy itself, the accuracy target, the method of data collection, and/or the time schedule by applying to the IRM Steering Committee or some other decision-making official or body.

4. Pursue incentives— The Agency should pursue incentives rather than enforcement actions to achieve the target accuracy. The LATF has identified 8 financial incentives and 3 information incentives for this purpose. The financial incentives are: resource "reserve pools," "tapping" resources, new grant conditions, fiscal year carryover funds, supplemental funds, fees for data use, State/EPA Data Management grants, and state grants. The information incentives are: State/EPA data sharing, public/private partnerships, and common ground with regulated facilities.

5. Upgrade FINDS— The Agency should upgrade FINDS, using address matching to populate the data base with locational coordinate data. The cost of the effort is estimated at this time to be $2 million. The accuracy of the locational data in FINDS would improve as individual programs update the data base. EPA should also pursue the "gateway" concept to allow FINDS to link with program data bases.

The next steps for the Agency are to:

- Develop a strategy to communicate the changed policy to the environmental community, especially states;
- Complete work begun by OIRM on guidance for implementing the policy;
- Pursue issues of: funding, security, quality assurance, incentives, and upgrades to EPA data bases; and
- Develop a plan to revisit the locational accuracy issue in 18-24 months.

Chapter 1: Recommendation #1: 25 Meter Goal

The first recommendation of the Locational Accuracy Task Force is that the Agency establish a 25 meter goal for locational accuracy. This means that latitude/longitude coordinate data for a site should be accurate to within 25 meters (±1.0 second).

EPA programs are expected to develop a strategy for meeting the 25-meter goal as expeditiously as practical, but in no case later than January 1, 1995, for all new and routinely updated computerized records.

In order to better understand the LATF's recommendation, this first chapter describes the following:

- Why 25 meters?
- Why a goal and not a standard?
- What is *new* data?
- What does 1995 mean?
- The benefits of the goal
- Issues and obstacles to implementation.

Introduction

The LATF, over a period of six months, collected and weighed considerable information on geocoding technologies and programmatic requirements for locational data. With this information, the LATF formulated the five recommendations reported in this document. This next section provides an overview of some of the analysis that went into those final recommendations and provides further detail where deemed necessary.

Why 25 meters?

The LATF, when deliberating on a target accuracy for the Agency, considered the following:

- What is a *reasonable* level of accuracy, given the status of current program requirements, current technologies, and the costs of implementation?

1: RECOMMENDATION #1: 25 METER GOAL

- What level of accuracy is or will be needed to support critical Agency decision-making functions, such as risk management?

In response to the first question, the LATF surveyed 25 EPA programs to determine their current or planned policies for the collection of locational data. [See LATF: *Acceptability Report: Program Offices* for more information.] Eleven of these programs already require 25 meters (±1.0 second) or better accuracy. A second survey which evaluated six geocoding methods (explained later) found 25 meters to be within the range of accuracies for field-based (i.e., conducted in the field) geocoding methods. The costs of implementing the most accurate of these methods will be high, since the Agency regulates or is interested in nearly 1,900,000 sites/ facilities (assuming no cross-regulation).

In response to the second question, the LATF agreed that achieving a base of accurate locational data (e.g., 25 meters) is as important as achieving a base of locational data *at the same level of accuracy* (e.g, all data at 100 meter accuracy). The latter condition is necessary to develop a common basis for the comparison and evaluation of locational-based data and cross-media support environmental analyses.

The LATF arrived at its final recommendation by balancing the need for a feasible target or standard against the need for data as accurate as possible in order to support the Agency's mission.

Why a Goal and not a Standard?

The preliminary results of a LATF study of geocoding methods produced some interesting results. In this study, the accuracies and associated costs of six methods were analyzed: address matching, map interpolation (office-based, field-based, and rectified photograph). [See *Current Methods for Determining Locational Coordinates* (Version 2.0) for a description of these methods.]

The study showed that all six methods exhibited wide fluctuations in levels of accuracies under certain conditions. Even Global Positioning Systems (GPS) may produce results outside of the target accuracy if satellites shift or if the Department of Defense downgrades the signals for reasons of national security. The truth is that there is no realistic way to ensure that reported coordinates are within the targeted range of accuracy unless they are compared to known reference points.

FINDINGS AND RECOMMENDATIONS OF THE LATF

Due to this fluctuation in the levels of accuracy using various geocoding methods, a standard measurement such as ±25 meters would be unenforceable. At best, the Agency can only choose a standard technology which produces accurate readings under typical conditions. The Agency must then implement that technology into its operations in such a way as to ensure the best results.

It is for these reasons that the Agency has chosen a *goal* of attaining high-quality, accurate locational data, rather than a standard. The LATF has chosen GPS technology, proven to be the most feasible yet accurate means today of determining latitude/longitude coordinates, to be the standard (see Chapter 2). It is the intent of the LATF that the Agency's programs employ this technology to move towards the stated goal.

What is *New* Data?

The LATF recommends that the locational accuracy goal apply to all data collected after the policy becomes effective. The most recent data will have the greatest potential for immediate impact on decisions; therefore, new data should have the highest priority for compliance with this goal.

Old data need not be brought up to the new accuracy standard until a specific objective exists to justify the effort. If there is a need for applying the new accuracy goal to old data, it can be done in several ways:

1. Collect the new data at the accuracy target during regular site visits;

2. Require compliance with the new target upon normal re-permitting; or

3. Collect the data for specific needs by visiting sites or hiring a contractor to provide the information.

What Does 1995 Mean?

An objective of the LATF recommendations is to allow program offices to begin implementation of the Locational Data Policy in a manner that it considers technically and fiscally responsible, within certain boundaries. To minimize boundary excursions, a "deviation process" will need to be established and implemented (see Chapter 3). This process requires a clear statement of what locational data standard the program office intends to meet and the timeframe and

1: RECOMMENDATION #1: 25 METER GOAL

procedure (eg.; address matching, map interpolation, or GPS) that will be employed to meet that standard.

Approval of deviation applications will initially be responsive to practical near-term needs of program offices, but approval will become markedly more difficult as **January 1, 1995** approaches, and virtually impossible thereafter.[1]

Program offices will be encouraged to begin filling the "location" data element within their data base by appropriate means. Later, as a facility's permit is up for renewal OR as it is undergoing a periodic inspection, an approved geocoding method such as GPS should be used to acquire new location data and replace old data in the applicable data bases.

[1] After January 1, 1995 the deviation application process will be much more rigorous. It is believed that by then all locational data entered into a system should meet the <=25 meter standard. The satellite system supporting GPS will have reached its full complement of 24 satellites by June 1992 or earlier. In addition, this policy will have impacted the Agency's budget process in a significant manner in both the FY93 and '94 budget request cycles (along with the modest beginnings in FY91 and reprogramming in FY92).

Benefits of the Goal

The most important benefit of implementing a 25-meter goal for locational accuracy is that it will result in a common basis for the comparison and evaluation of all of EPA's program data. With this basis the Agency can achieve truly integrated environmental analysis, planning, and management.

Presently, all EPA programs except OSWER have regulatory requirements for the use of latitude and longitude designators. (OSWER has specific, non-regulatory program directives.) However, there is no common required accuracy for latitude/ longitude specification amongst the various programs. Hence, the least accurate data become the common denominator for all evaluations in which those data are used. The limits of this least common denominator decreases the reliability of any evaluations.

There will be many other benefits to implementing the LATF recommendations. Some of these benefits we are realizing already; others will come later. For example, the data from all programs will be able to be presented on maps, allowing the visual association of pollution sources to their potential impacts on environmentally

vulnerable sites such as water supplies, wetlands, spawning and nesting locations, recreation areas, and population centers. That kind of information will become an increasingly valuable tool for defining environmental priorities, targeting monitoring and compliance programs, and fashioning future strategic plans for protecting environmentally threatened areas.

Besides those uses, the Agency's Geographic Information System (GIS) capability has proven to be a masterfully effective way to explain to the public, legislators, and the media the extent and location of environmental problems. GIS tools may also be used to evaluate the extent of environmental improvement under varying management scenarios.

Issues and Obstacles

The path to implementing the LATF recommendations is not without barriers— programmatic needs differ, costs are large, some existing data bases need modification, definitions must be uniform, and training must be accomplished. In addition, preliminary estimates put the cost of implementing a standard at well over $50 million.

Another major obstacle to implementing the LATF recommendations is the need to consider the reporting burden of the Agency's regulated entities, both in terms of gaining Office of Management and Budget approval, as well as in terms of the feasibility and fairness to reporters.

The LATF recognized these factors and has thoroughly weighed them in formulating its recommendations.

Chapter 2: Recommendation #2: GPS Standard

The Locational Accuracy Task Force is recommending the adoption of a technology-based standard for locational accuracy. This implies that global positioning systems (GPS) will be the technology of choice for achieving accuracy goals. However, there are a number of issues associated with a GPS-based standard that need to be addressed in order to undertake an Agency implementation.

Present and Future State of GPS Technology

The NAVSTAR Global Positioning System is owned, operated and maintained by the U.S. Department of Defense. In its fully deployed status, it will consist of 28 satellites in independent earth orbit. Each satellite broadcasts two kinds of radio signals on pre-assigned frequencies: the C/A-code and P-code. Earth-bound receivers may be used to acquire at least four signals to determine a location in latitude, longitude and elevation coordinates using mathematical and doppler solutions. When fully deployed, the system will be able to provide 25m minimum accuracies for navigation units (C/A-code receivers), 24 hours per day. There is reason to believe that actual accuracies will be <10m in most situations.

A much more precise GPS unit is the geodetic device (P-code receivers). Since these units receive a much more complex signal and are capable of receiving more than four satellites at a time, much higher accuracies are obtainable. When fully deployed, the system will be able to provide centimeter-level accuracies.

Currently, 18 satellites are deployed. This limits the acquisition of satellite signals and has implications on the accuracy and availability of GPS data. Our studies have shown that there is a high degree of variance introduced if a receiver shifts "lock" from one set of satellites to another. Because of limited availability, significant time must be spent in pre-survey activities, planning location acquisitions around satellite visibility.

The military reserves the right to selectively degrade or even deactivate the system at any time to protect national security interests. Degradation (termed selective availability) implies that the

FINDINGS AND RECOMMENDATIONS OF THE LATF

satellites signals are programmed to provide erroneous locations to earth receivers. However, DOD "guarantees" that 100m accuracies will still be achievable, even under selective availability. Selective availability may be overcome by the use of two receivers for differential correction and the use of post-processing software.

Conclusion

The GPS satellite system is not fully deployed at this time, and this has implications on the current use of the technology. However, based on our limited studies, it shows potential to reasonably achieve 25 - 100m accuracies with adequate precision and repeatability under optimal conditions in its present deployment status. There is every reason to believe that, with final deployment of the constellation in mid-1992, GPS may achieve <10m accuracies with high precision and repeatability.

Assumptions

The above conclusion is based on the following assumptions:

- The use of multiple GPS receivers for differential correction is necessary to overcome the effects of inadequate satellite coverage and military degradation of the signal.
- Optimal conditions implies that the GPS unit is receiving four satellite signals with good satellite-to-receiver angles and little-to-no influence from weather conditions, topography, or external radio noise.
- GPS is not a stand-alone technology. More often than not, a field person is using other information sources such as maps and aerial photos to locate themselves for GPS data acquisition.
- An experienced user implies that a trained user is operating the receiver in the field and performing the necessary post-processing steps in the office.

Concerns

The following concerns will need to be addressed to fully exploit the utility of GPS:

1: RECOMMENDATION #2: GPS STANDARD

- What is a facility? The issue of repeatability becomes a major concern without a clear definition of a facility. Most often programs are interested in point data, but what point? Front door? Centroid? Outfall? There will be a number of occurrences where a facility will have multiple regulated entities on the property, with multiple reporting requirements.

- Real-time correction— Under selective availability, it is possible to offset the introduced error by using two receivers for differential correction. While the majority of our data collection may be served by post-survey support, there will be times when an accurate location is necessary in the present timeframe. Examples such as returning to a pre-determined, unmarked sampling location or emergency response situations come to mind. In these scenarios, it may be necessary to establish a radio link with a base station, allowing for real-time corrections to be passed via the radio link.

- What types of receivers and how many? There are several vendors entering the market with a wide range of prices for their units. C/A-code navigation units range from $3- 15 K, while P-code geodetic units start at $30K. Will low cost units be adequate for Agency use, given our accuracy targets? How many of the geodetic and navigation units are necessary per Region?

- Effective use of community networks— A community network implies that there is a centralized approach to managing GPS correctional data. Regions I and VII have advocated the deployment of central geodetic units for continuous logging of GPS data. These data sets could then be used to differentially correct navigation unit data and thereby offset the effects of selective availability and highly variant data. While this sounds like a solid solution, very little is known about the effective range of community networks. The effects of topography, microclimate and radio noise sources may need to be studied at a regional level before implementing community networks.

- Training— In its present state, GPS is not a "push button" operation. Users need to be adequately trained in pre-survey, survey, and post-survey operations. The operator in the field must be constantly aware of signal strengths, satellite horizons, noise sources, and topography. When a survey is completed, there is a tremendous amount of data to be sorted and screened before an optimal locational solution is determined.

- Utility— Currently, there is a lack of software tools for the sorting/screening of GPS data during post-processing. The vendor community is responding, b ut the user community may be the best source of GPS data utilities.

Implementation Issues

The above assumptions and concerns have direct bearing on a national implementation of GPS within the EPA. We have identified several steps that need to be addressed:

- Policy determinations— A clear definition of a facility is necessary.

- Requirements analysis— Define the following:

  - How many units are needed per Region? Several of the EPA Regions have identified the need for 5-7 navigation units and 1-2 geodetic units per Region. There is every reason to believe that the states will interpret the Location Data Policy (LDP) as an EPA requirement that will necessitate an increase in grant funding to states to meet the requirement. Dennis Kirk surveyed three states and found that they foresee a need for 20-30 navigation units per state.

  - What types of units are needed for each Region? What is the mix of C/A-code and P-code receivers necessary for a Regional implementation?

  - There is a standardized data exchange format that would promote compatibility between manufacturers' equipment. However, some manufacturers may not subscribe to the data elements necessary to promote exchange. In addition, the exchange program has not been ported to UNIX platforms and has not been tested by EPA.

  - What is the role of community networks? Could the EPA Region be the central repository of GPS data for use by states? This might help offset the state concerns about the LDP requirement without adequate support from the Agency. What is the minimum equipment suite necessary to support a central GPS repository? What is

1: RECOMMENDATION #2: GPS STANDARD

the effective range of community networks under various conditions?

How do we obtain GPS training? Do we fill the training need internally or via vendor? What are the necessary fundamentals for any curriculum?

- What is the need for real-time locational accuracy? How often and under what circumstances will we need to know where we are, right now, with no induced inaccuracies? The definition of this element will determine the number of radio transceivers necessary for real-time correction.

- How soon can we define and document standard procedures for pre-survey, survey, and post-survey GPS activities? EMSL-LV has begun this process with the drafting of the GIS primer (now in review). However, with advances in the receiver technology and absorption of case studies, this documentation will have to be revised and updated.

Equipment Costs

I have estimated the cost of a typical Regional GPS suite of equipment with the following assumptions. A Region wishes to:

- Maintain a central repository of GPS correctional data;
- Maintain a community network;
- Have adequately trained personnel; and
- Maintain 5 navigation units for field work.

COST OF A TYPICAL REGIONAL GPS SUITE

Cost of central unit to support community network:	
- 1 geodetic P-code GPS receiver:	$50K
- 1 dedicated Sun Sparc 1+ workstation with adequate storage and data archival capability:	$45K
10 days of training (including travel, per diem, salary and fees):	$10K
Purchase of 5 field units:	$70K
TOTAL ESTIMATED COST:	$170K

FINDINGS AND RECOMMENDATIONS OF THE LATF

NOTE: This estimate does not account for any purchases by the EPA for use by the states.

Final Cautions

We would caution the Agency against total dependence upon GPS technology. As promising as it appears, DOD Assistant Secretary of defense for Command, Control, Communications, and Intelligence has continued to state the need to retain the ability to degrade the signal or completely deactivate it for national security reasons. The Agency must not loose sight of the need to maintain or increase proficiencies in the use of other technologies for locational determination. Our studies have shown that simple map interpolation has the potential for achieving a 25m accuracy when in the hands of an experienced map user. additionally, address-matching capabilities will continue to improve as vendors respond to the need for highly accurate address data.

Chapter 3: Recommendation #3: Deviation Process

The third recommendation of the Locational Accuracy Task Force is that the Agency define a deviation process for the Locational Accuracy Requirement.

Waiver Policy

In the locational accuracy policy, the Agency needs to:

- Establish criteria by which to evaluate, and to grant or deny, waiver requests for exclusion of each individual data collection activity from (1) the locational policy in general, and (2) the locational accuracy requirements in particular.
- Be uniform in the execution of the waiver process.
- Assure that agency-wide locational coordinate objectives are being accommodated and considered in data collection.

Purpose of Having a Waiver Provision

The purpose of having a waiver provision is to:

- Give the Locational Accuracy oversight group an opportunity for reasonableness about the implementation of the locational data accuracy requirement;
- Demonstrate a sensitivity to the challenge faced by program managers to adapt to the accuracy requirement;
- Establish an approach and mechanisms for enforcement of the locational data accuracy requirement;
- Provide an orderly, structured process to ensure consistency with the 25 meter locational accuracy goal by the 1995 target date;
- Ensure that progress toward the locational accuracy goal is attentive to, and driven by, data quality objectives;

- Maximize effective use of existing technology while the Agency builds Global Positioning Systems (GPS) capability; and
- Have an effective process for identifying Agency data collection efforts which reasonably should not be subject to the locational data policy and minimum locational accuracy standard.

Waiver Process

To incorporate into Data Quality Objectives (DQO) process a demonstration of necessity of a waiver from the locational data policy and/or accuracy requirement that:

- Describes the objective of the data collection effort and contrasts it to Agency objectives for locational data accuracy.
- Defines an approach to meeting the locational data accuracy objectives for the data collection effort (and how it deviates from the Agency-wide requirement.
- Provides a justification as to why the Agency-wide locational data (and accuracy) is infeasible for this particular data collection effort including cost.
- The program requesting a variance from the locational policy or accuracy target would prepare an individual petition for a particular data collection project.

Submitting/Approving Waiver Requests

All waiver petitions are to be reviewed/approved by a subcommittee appointed by the IRM Steering Committee. The subcommittee would be composed of 3 to 5 members of the Locational Accuracy Task Force, and represent policy, technical, and media program interests.

The final waiver provision should specify the composition of the subcommittee, and it is felt that at least three of the subcommittee members should be from the program media.

The Program commits to preparing a written request for waiver containing appropriate arguments and documentation. The written waiver request with supporting documentation is submitted to the

3: RECOMMENDATION #3: DEVIATION PROCESS

subcommittee, signed by the Program's Senior Information Resources Management Office (SIRMO).

Decisions of the subcommittee are subject to review by the IRM Steering Committee. An annual report to the Steering Committee is due to review program policy effectiveness.

Requirements of the Waivers and Procedures for Granting Them

Waiver requirements should encompass all the criteria for granting a waiver. For example, has the Program documented cost of implementing the locational data policy and accuracy standard?

(1) Develop Procedures for Granting or Denying Waivers, including:

- A definition of roles and responsibilities (who must prepare the waiver request, who reviews it, who has final authority, etc.) of processors.
- A definition of enforcement mechanisms, and the roles of the "enforcers," both at the HQ level and in the field.

(2) Definition of roles and responsibilities of applicants, including:

Timing of waiver request submissions and approvals/disapprovals.

(3) Program implementation plans:

- All waiver requests must be specifically tailored for individual program element data collection efforts, and must contain both the rationale behind the requested variance and specific Data Quality Objectives (DQO's) for the data for which the variance is being requested.
- The waiver should include reasons why the locational policy can't be met, such as resources, program regulations, applicability of lat/long data, and alternate sources of lat/long data (i.e., Construction Grants can obtain lat/long data from PCS if data linkages between data bases are in place).

- All waiver requests should address the overall Agency plan for achieving consistency with the locational goals.

Criteria for Granting a Waiver

The criteria for granting a waiver should be:

- Has a program thoroughly analyzed and documented its needs for locational data? Are all primary uses of the data well understood?

- Has the program solicited input from secondary users of their data? (This will help to ensure that Agency-wide/multimedia requirements are incorporated into the locational data requirements of individual programs.) Are the data intended for single purpose objectives not relevant to more general or multimedia applications?

- Is the program mainly concerned with facility level data as opposed to entity level data (i.e., pipes, stacks, wells, etc.)? A singular locational reference for a regulated facility should not be required to support 2.5 meter locational accuracy. On the other hand, entity level data generally should support 25 meter accuracies for locational data.

- Has the program documented the expected cost of implementing the locational data policy and accuracy standard? Has the program identified where this burden lies (i.e., Headquarters, Regional Offices, States, the regulated community)?

Note: Burden— Cost should not be a deciding factor in a waiver decision. However, it should be taken into account when developing a phased implementation program.

Chapter 4: Recommendation #4: Incentives

The fourth recommendation of the Locational Accuracy Task Force is that the Agency pursue incentives for locational accuracy.

Short Term Resource Policy Recommendations

As a portion of the recommendations, task force representatives agreed unanimously to include financial and informational incentives. The attachment, Resource Policy Options for Location Standards Implementation, lists and describes various options which could be used separately or collectively depending on the particular data base affected.

Because of the magnitude of accumulating and maintaining highly accurate locational data, we need to look at creative ways of leveraging resources as well as investing additional EPA resources. Several of the options listed provide incentives to State and local governments and the private sector to assist in defraying the costs of meeting more exacting EPA locational standards.

In addition to incentives to external entities and overall implementation, the task force discussed providing internal incentives to offices which accelerate implementation of locational data standards in their data bases. One way to do this would be to hold funds either in the Administrator's Office or with the OIRM to reward offices during the year as they brought their system into compliance. Such an approach would require the OIRM Steering Committee to establish and monitor the standards upon which these incentive funds would be distributed. The Steering Committee should consider pursuing making such incentives available.

The task force also felt that a Reilly/Habicht endorsement of this policy was essential for its success. To initiate planning for meeting the goal in FY 1995, we recommend that locational data be included as an item for discussion in the Administrator's February kick-off planning meeting for FY 1993 with the Assistant Administrators (AAs) and Regional Administrators. The inclusion would not only highlight the issue of locational accuracy in data bases, but it would also focus all of EPA's senior management on assessing the

resource needs and impacts on their information providers and users for better locational accuracy. The AAs could be asked to include in their plans how they will address the issue in FY 1992 through the target fiscal year of 1995. This would better define the Agency's ability and commitment to implementing this policy.

Among the options on the attachment the task force recommends as the most realistic to consider, though all require either internal or external approval and/or development, are creating resource reserve pools, establishing new grant conditions, using State/EPA data management grants, providing better information products and support infrastructure for locational data, and promoting Public/Private Partnerships.

"Early" Incentives for Achieving Better Quality Locational Accuracy

There are a range of financial and other incentives available for achieving better quality locational data. The financial incentives which appears most viable for "early" implementation include:

- "Tapping" resources, including Agency-wide, media or program-specific taps;
- Requesting special supplemental funds;
- Creating resource "reserve pools" from existing state grant program funds;
- Imposing new grant conditions in existing state grants;
- Initiating state grant "givebacks;"
- Earmarking fiscal year carryover funds;
- Establishing a "fee" system for data access and use; and
- Funding through State/EPA Data Management Grants.

The "information" incentives are based upon the increasing value of locational and other environmental information to public and private sectors. These incentives reflect the reality that no one organization by itself can achieve complete and accurate information. Cooperation and stewardship in managing information are essential.

4: RECOMMENDATION #4: INCENTIVES

The most promising information based incentives for early implementation are:

- Information products to foster State/EPA data cleanup, sharing and decision making;
- Business opportunities for third-party, value-added environmental information providers (public- private-partnerships);
- Mutual benefits to all parties (EPA, State, regulated facilities and the public) of using good information; and
- An EPA infrastructure to support the Nationwide collection, processing and dissemination of environmental locational data.

Financial Incentives

The following paragraphs provide a brief description of each policy option and are not listed in a recommended priority of use:

- Tapping Resources — This option is common Agency practice for funding special interest, unfunded programs. Funds would be drawn off affected EPA offices on a to-be-determined basis and channeled into areas necessary to implement the policy.
- Supplemental Funds — The Agency could request additional special funds for this activity from OMB if the need for funds did not coincide with regular annual budget submissions.
- Resources Reserve Pools — The Agency could look into the option to "reserve" a portion of each state grant program's funds, and dedicate the resulting "pool" of funds specifically to this activity. Interested states could opt to perform the activity, increasing their state grant funds from the pool.
- Grant Conditions — The Agency already funds the states to implement information management activities in existing state grant programs. The Agency could insert a new condition into all of its grants requiring states to perform this activity with existing funds.
- Grant Givebacks — The states could fund this activity from existing grant funds by "giving back" a portion of their

FINDINGS AND RECOMMENDATIONS OF THE LATF

grants to the Agency to perform this activity for them. The states would request the Agency to deobligate grant funds to carry out specific activities for them.

- Fee System — The Agency could establish a fee to be paid by potential users of the data base, if the information has value to private industry. The fees collected could be used to fully or partially fund states' implementation of the system, depending on the amounts collected.

- State/EPA Data Management (SEDM) Grants — This new grant program begins in FY 1991. It could be used to partially fund some initiatives (e.g., staff, equipment, contractor support, central services, data, etc.) to improve the collection, quality, use and dissemination of locational data.

The financial options above could be executed in combination, using existing Agency program authorities; however, there may be special considerations for each option or combination of options, depending upon specific dollar amounts and other operating requirements. Some of these considerations might be:

- Notification to the Appropriations Committees for reprogramming actions that exceed ceilings;

- Grant deviations for exceptions to formula-driven state grant programs;

- Legal determination of an appropriate use of Trust Fund accounts;

- Federal Register notices or other public notification processes; and

- Office of Management and Budget (OMB)/Treasury approval of new accounts.

Depending upon the financial options selected, elements of Office of Administration and Resources Management (OARM), CLA, Office of General Counsel (OGC), and the media National Program Managers (NPMs) would need to fully explore and develop a comprehensive resource strategy, assuring that practical, legal, and other necessary approvals are obtained.

4: RECOMMENDATION #4: INCENTIVES

Information Incentives

The following paragraphs provide a brief description of each of the information incentives listed above:

- Information Products for State/EPA Staff — The State/EPA Data Management Program demonstrates that information in itself can be a strong incentive for cooperative data management efforts. Both EPA and the States have a common need for accurate locational data for organizing and analyzing environmental data.

  Locational data is best collected and verified as close as possible to the local level (e.g., State field office personnel, regulated facility), but incentives are needed to secure local support. Aside from using regulatory and fiscal incentives, providing environmental information for free that is useful at the local level can be a good incentive. This will necessitate some rethinking of the near term priorities and information products of EPA's State/EPA Data Management (SEDM) and GIS Programs.

  Effort should be placed on developing relatively simple, PC based, "downloaded" GIS and EPA data base products (e.g., Facility INDex System (FINDS) or Hazardous Waste Data Management System (HWDMS) Inventory Report, a GIS-like, 8 1/2 x 10 USGS Quad Map overlay of locations of regulated facilities) that are useful at the local level. These products should support locational data cleanup as well as local decision-making. CD ROM is a promising technology that could be used in the future to disseminate large quantities of information, particularly GIS data. The same PC diskette and CD ROM based products could also enable EPA to provide better public access to environmental data.

  Business Opportunities For Third-Party Information Providers — EPA should consider public/private partnerships to foster growth of third-party, value-added, environmental information providers. There is growing interest in business opportunities in this area. For example, law firms conducting property title searches, investors considering land acquisitions and the financial industry considering mortgage loans are asking EPA through the Freedom-of-Information Act (FOIA) for information and locational data for properties that may have environmental

FINDINGS AND RECOMMENDATIONS OF THE LATF

liabilities. Therefore, EPA and State information has commercial, economic value.

EPA has a near term need for locational data in urban areas based upon street address matching. Rather than pay a company such as Dun & Bradstreet or Donnelley Marketing to do address matching, EPA should explore partnerships that "trade" address matching services for "FOIA" services (e.g., environmental information).

EPA is limited in its ability to respond well to many of these types of information requests under FOIA, yet EPA is a sole source of some key information. Companies may see a new market for their information services by offering added value to what EPA would normally provide in filling a FOIA request. There is opportunity for EPA (and States) to save resources that will be needed to meet the ever growing FOIA workload.

- Mutual Benefits To All Parties Of Using Good Information — Our recent experience with the Title III Toxics Release Inventory (TRI) demonstrated the benefits and problems of public dissemination of environmental data. There is clearly a mutual benefit to EPA and the regulated community of providing the public and private sectors with good information. The EPA regulated community may be reluctant to provide information in general, but it is in their overall self interests to be sure accurate information is being used. The locational data for their facility is no exception.

  The regulated facilities could provide locational data in several ways. They can mark-up U.S. Geological Survey (USGS) topographic maps provided by EPA and mail them to a central point for determination of coordinates. EPA's Office of Waste Programs Enforcement (OWPE) is using this approach for 7200 National Pollutant Discharge Elimination System (NPDES) major dischargers. Another possible option is to submit an official land survey for the facility with key points (e.g., center of primary facility, discharge pipe, etc.) identified. More homework is needed to explore this approach. In either case, EPA would need to provide an infrastructure to determine coordinates, process them into a data base and provide GIS output for coordinate verification.

  Provide Infrastructure For Processing & Disseminating Locational Data — Achieving good quality locational data requires an infrastructure to provide data conversion, maintenance, integration and dissemination. Conversion is

4: RECOMMENDATION #4: INCENTIVES

the process of taking an annotated map, property survey, etc. and determining latitude and longitude coordinates. Maintenance is the process of input, verification and revision of locational data in the numerous EPA data bases. Integration is the process of pulling together and comparing data to save redundant data collection and to also crosscheck data quality. Dissemination involves both the process of returning the data to the originators for verification as well as making the accurate data widely available.

Some of this infrastructure may best be handled by a central Headquarters or Regional staff with proper equipment (e.g., digitizer, GIS workstations) and specialized training (e.g., to input to EPA systems, to crosscheck facility data via FINDS). An example would be the CSC contractor support teams for FINDS data maintenance and the OWPE contractor support for converting Permit Compliance System (PCS) "Majors" coordinate data.

Aside from the "processes" part of the infrastructure, there is also a need for technical assistance, equipment and information products for managers and staff involved in acquiring accurate locational data. The information products include not only GIS like products but also current policy and guidance documents, reference material and reports from EPA data bases (e.g., TRI, PCS, HWDMS, Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS), FINDS/FACTS).

EPA should consider providing as much of this infrastructure as possible as a way of encouraging and supporting locational data collection and verification at the local level.

Chapter 5: Recommendation #5: FINDS

The fifth recommendation of the Locational Accuracy Task Force is that the Agency upgrade FINDS.

Introduction

With the significant investments being made in locational information across EPA's many programs we need to insure that FINDS is poised to serve as *the* information utility for providing the environmental community and the public access to this very important data resource. Access to EPA's environmental location and basic identification data must be viewed as an incentive to get the non-EPA community to share their valuable spatial information resources with us.

To complete the daunting effort needed to develop the refined spatial data to conduct GIS projects, every group must do and then share its part. EPA must pay its spatial data dues by creating and maintaining a vigorous, viable, nationally supported system able to act as a gateway to all our facility, discharge and other environmental entity data. This is not a trivial task. However, failing to address this need will result in many separate, disconnected, likely incompatible systems being built. Extremely poor data access, and great difficulty in doing the data cross checking needed to insure continued data integrity would be the end result.

We cannot on one hand ignore the need to seriously deal with FINDS, and on the other hand make the significant resource investment needed to collect and quality control our spatial data as proposed by the LATF. Nor can we continue to "squirrel away" these data in separate unsupported systems.

Vital EPA and state integrated applications above and beyond GIS uses are at stake here including: tracking multi-media inspection activity, implementing multi-program enforcement, conducting risk assessments, and providing simple access for our states to application and permitting information.

Two key issues regarding FINDS and the agency's direction on locational accuracy need to be resolved: 1) how should FINDS be

5: RECOMMENDATION #5: FINDS

enhanced to support the planned EPA and state efforts collecting and using locational data? and 2) are there ways that the existing data in FINDS can be used with techniques like address matching to provide a quick start on making good, comprehensive locational data available at least at the EPA regulated facility level?

Retooling the FINDS System

There is an OIRM initiative underway now to modernize FINDS. It must be expanded to consider and then tackle these issues in more than a superficial way. FINDS needs to be treated as a mainstream, "serious" system like Aerometric Information Retrieval System (AIRS). The GIS community needs to be actively drawn into the requirements analysis.

The current FINDS system was not designed to address our existing location data needs — only a single latitude longitude can be tracked at the facility level. For example the capability does not exist to handle locational data at the discharge or operable unit level although tracking these entities have become critical to many GIS applications. FINDS also needs added system edit facilities to help users spot changes in program system information to identify major facility and discharge status changes and possible data errors. It is recommended that the role of FINDS be analyzed in the context of the locational data requirements perceived by the GIS community and other locational data users.

Opportunities for drawing in related efforts that are logical parts of the overall solution include fusing into a single project the FINDS efforts along with important work now underway in the Software Development Center's "Gateway Project". This project seeks to provide access and sorely needed data browsing and data integration capabilities, through a common user interface, to the major national systems. The gateway would in part rely on FINDS to provide the facility reference information for linking to the major program systems like PCS and AIRS.

Another possible opportunity that needs to be reviewed is the use of Structured Query Language (SQL) database technology for delivering the sophisticated data model and performance likely to be needed to meet added FINDS requirements.

We also need to consider the impact of collecting refined locational data on the individual national program systems. Most (not all) systems can store facility locational data. Some systems

(not more than a few) handle pipe/stack/operable unit level location data. Work needs to be begun, with the national system program managers, to insure that each system has the capabilities needed to enter, store and deliver to the user the full range of locational data being collected as we expand our location data efforts. This is a *major* undertaking and must not be trivialized. Serious marketing, by our most senior managers, must be done to convince the individual Headquarters program managers that their best efforts are needed here to provide these system changes. Media program managers must get that message that they should commit the resources to modify their systems not only to benefit their current media data clients but also to allow the environmental community at large to tackle data integration and risk assessments.

In some cases, FINDS may, as a last resort, have to house some program data until the national system can be modified. In any case, establishing ownership of collected location data by the appropriate delegated group is key. Success in EPA developing and maintaining solid locational information will only come when there is a well understood partnership between the national system managers, OIRM (in its FINDS information utility role) and the state and regional locational data consumers.

FINDS and Facility Address Matching

The LATF discussed the potential of using address matching of FINDS data to quickly generate quality facility data that would, until replaced by more refined data, be available for the many spatial analyses underway in the agency that require regulated facility location data. The thought was to have OIRM explore the viability of quickly filling the locational "holes" in FINDS, using a contract service to tackle this significant effort on a national basis.

The contractor would work the FINDS address information through a process that would check, and if needed, refine the address syntax, then attempt an automated address match. Reprocessing of those match rejects that looked feasible would also be attempted.

We need to keep in mind that address matching is no panacea for poor locational information. This technique would generally be successful only for non-rural facilities. It would not be recommended for facilities where good GIS or program system locational information could be drawn upon. Also, there may be a need for reviewing the address information before conducting the address matching for a given facility. The bottom line is that while it

5: RECOMMENDATION #5: FINDS

appears that address matching is worth pursing it will not by itself be the total "cure" for the generally non-existent state of facility location data. This technique needs to be thought of as one element of a larger effort aimed at generating and maintaining quality geocoded information for our inventory of regulated entities.

Overall Recommendations

The overall recommendations for FINDS are:

1. Retooling the FINDS System:

Recommendations:

- Develop a project plan and begin requirements analysis to address locational data needs (this would be carried out in parallel with the on-going FINDS redesign);
- Draw together FINDS and Gateway project efforts;
- Define follow on steps for refocusing the FINDS/Gateway system efforts using agency system life cycle methodology to detail implementation options, cost-benefits etc.; and
- Begin collaboration with program offices to address media system changes required so they could support locational data.

Cost Implications:

- Added funding is required in FY 1991 and FY 1992 to conduct the FINDS requirements analysis;
- FY 1992 funding will be required to support accelerated FINDS and Gateway development efforts; and
- Program offices may need added financial support in FY 1991 and FY 1992, above their current base, to modify their systems to handle refined latitude/longitude data and begin related data cleanup and collection efforts.

FINDINGS AND RECOMMENDATIONS OF THE LATF

2. Facility Address Matching:

Recommendations:

- Develop a detailed methodology for address matching, taking into account identifying and using existing latitude/longitude data from some program systems and selectively excluding FINDS records that are not true facilities, e.g. RCRA Transporters;
- Perform cost analysis and review external contract sources for conducting address matching of FINDS addresses, based on FINDS statistics; and
- Develop a project plan and begin national facility address matching efforts.

Cost Implications:

Funding to plan and then implement selective facility address matching of FINDS addresses needs to be set earmarked for both FY 1991 and FY 1992 .

Appendix B

CONTACTS FOR LDP IMPLEMENTATION ASSISTANCE

Appendix B

Contacts for LDP Implementation Assistance

(January 1992)

ORGANIZATION	TOPICS	NAME	TELEPHONE NO*
OARM/NDPD	GPS Procurement	John Shirey	(919) 541-5730
OIRM/IMSD	Data Administration	Jeff Sabol	(202) 260-8974
OIRM/IMSD	State/EPA Data Mgmt	Michele Zenon	(202) 260-5913
OIRM/PSD	National GIS Program	Tommy Dewald	(703) 557-3083
OIRM/PSD	FINDS	Ingrid Meyer	(703) 557-3145
OIRM/PSD	GRIDS	Bob Pease	(703) 557-3018
OPPE	Information Collection Requests	David Schwarz	(202) 260-2706
ORD	Data Quality Objectives	Nancy Wentworth	(202) 260-5763
ORD/EMSL-LV	GIS Center of Excellence	Mason Hewitt	(702) 798-2377
EPIC	Information Collection Requests	Terry Slonecker	(703) 349-8970

* To confirm phone numbers, call the *EPA Locator* at (202) 260-2090.

Appendix C

PARTIAL LIST OF FORMS RELEVANT TO LDP

Appendix C

List of Forms by Program that Possibly Need to Be Modified for LDP Compliance

Superfund Program

Form #	Date	Form Title
2070-11	(7/81)	Potential Hazardous Waste Site-Site Inspection Report *(Site Identification)*
2070-12	(7/81)	Potential Hazardous Waste Site Preliminary Assessment
2070-13	(7/81)	Potential Hazardous Waste Site-Site Inspection Report
2070-14	(7-81)	Potential Hazardous Waste Site-Current Disposition
9100-1	(2/88)	Technical Enforcement Support at Hazardous Waste Sites
9110-2	(8/88)	Organic Traffic Report (for CLP use only)
9350-1	(1/90)	Toxic Chemical Release Inventory Reporting Form (Form R)
9510-1	(7/88)	Substantiation to Accompany Claims of Trade Secrecy: Emergency Planning/Community Right-to Know Act

Water Program

Form #	Date	Form Title
3320-1	(ND)	NPDES Discharge Monitoring Report
3510-1	(8/90)	Application Form 1-General Information-Consolidated Permits Program
3510-2B	(6/80)	Application for Permit to Discharge Information-Consolidated Permits Program (Form 2B)
3510-2C	(6/80)	Application Form 2C-Wastewater Discharge Information-Consolidated Permits Program
3510-2D	(9/86)	Application Form 2D-New Sources and Dischargers: Application for Permit to Discharge Process Wastewater
3510-2E	(7/86)	Facilities Which Do Not Process Wastewater
3560-3	(3/85)	NPDES Compliance Inspection Report
3560-4	(2/80)	Deficiency Notice-NPDES
7500-52	(6/80)	SPCC Inspection Summary Sheet
7500-53	(9/80)	SPCC Inspection Field Sheet

Appendix C (continued)

Water Program (cont)

Form #	Date	Form Title
7500-55	(1/89)	NPDES (DMR) Laboratory Performance Evaluation
7500-60	(3/87)	Civil Litigation Review
7520-6	(10/85)	UIC Permit Application
7520-7	(9/90)	Application to Transfer Permit
7520-8	(9/90)	Injection Well Monitoring Report
7520-9	(2/84)	Completion Form for Injection Wells
7520-10	(9/90)	Completion Report for Brine Disposal, Hydrocarbon Storage or Enhanced Recovery Well
7520-11	(9/90)	Annual Disposal/Injection Well Monitoring Report
7520-12	(9/90)	Well Rework Record
7520-14	(9/90)	Plugging and Abandonment Plan
7530-1	(9/88)	Notification for USTs
7550-6	(1/73)	NPDES Application for Permit to Discharge-Short Form A
7550-22	(7/73)	NPDES Application for Permit to Discharge-Standard Form A-Municipal
7550-23	(7/73)	NPDES Application for Permit to Discharge Wastewater--Standard Form C

Toxics Program

Form #	Date	Form Title
7710-3C	(9/90)	Chemical Substance Inventory Report
7740-5	(3/83)	TSCA Investigation Summary
7710-35	(5/82)	Manufacturer's Report Preliminary Assessment Information
7710-52	(ND)	Comprehensive Assessment Information Rule-Reporting Form
7710-53	(12/89)	Notification of PCB Activity
7740-21	(3/90)	Mercury Reporting

Appendix C (continued)

Solid Waste Program

Form #	Date	Form Title
3510-3A	(5/80)	Acknowledgement of Application for a Hazardous Waste Permit
3510-3B	(5/80)	Acknowledgement of Application for a Hazardous Waste Permit (Verification)
8700-12	(6/85)	Notification of Hazardous Waste Activity
8700-12B	(2/80)	Acknowledgement of Notification of Hazardous Waste Activity
8700-13A	(5/80)	Generator's Annual Report
8700-22	(9/88)	Uniform Hazardous Waste Manifest
8700-22A	(9/88)	Uniform Hazardous Waste Manifest (Continuation Sheet)

Air Program

Form #	Date	Form Title
3520-2	(5/85)	Lead Additive Report for Refinery
3520-3	(5/76)	Lead Additive Report for Manufacturing Facility or Site
3520-7	(9/87)	Notice of Violation of Section 211 of the Clean Air Act

Pesticides Program

Form #	Date	Form Title
3540-2	(3/77)	Notice of Inspection
3540-5	(5/76)	Report of Analysis
3540-8	(11/88)	Application for Registration of Pesticide Producing Establishments
3540-8A	(11/88)	Application for Registration of Pesticide Producing Establishments
3540-16	(10/81)	Pesticides Report for Registration of Pesticide Producing Establishments
3540-20	(4/75)	Use Investigation Report
3540-22	(4/75)	Corrective Action Report

Appendix C (continued)

Pesticides Program

3540-25	(3/77)	Notice of Pesticide Use/Misuse Inspection
3540-26	(3/77)	Receipt for Pesticide Use/Misuse Samples
8500-1	(6/83)	TSCA, FIFRA, SARA Title III Investigation Report
8580-7	(3/83)	FIFRA Investigation Summary

Enforcement Program

Form #	Date	Form Title
3500-5	(5/90)	Fuels Field Inspection

Appendix D

FACILITY DEFINITION GUIDANCE FROM THE FACILITY IDENTIFICATION DATA STANDARD IMPLEMENTATION PLAN (FIDSIP)

1.4.1 FACILITIES TO WHICH EPA FACILITY ID CODES MUST BE ASSIGNED

The word "facility" represents a wide range of entities as defined by each of the environmental media programs within EPA. Inconsistency in the definition of a "facility" among the environmental programs poses a challenge in ensuring the uniform assignment of facility identification codes. It is difficult to have a single, all-encompassing facility definition that satisfies the specifications of all the programs. Therefore, the FIDS defines a facility as "*...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, a hospital, a national park, an office building, or a prison.*" [1] This definition, while seemingly general, allows program managers (who help determine whether an entity should have an EPA facility ID code) to apply ***a "common-sense" approach to uniquely defining a facility***. This approach is taken because, among the various environmental laws, a facility:

- Might be a ***discrete location*** (e.g., with well-defined property boundaries) at which there is environmental regulatory activity (e.g., a permit has been issued to this location, or monitoring at this location is required)
- Might be separate areas linked by ***a common environmental concern*** (e.g., a spill spanning several properties), and might represent a "site" rather than a single plant
- Might have several geographically separate portions which are ***linked by common ownership*** (e.g., a Federal facility with non-contiguous sections).

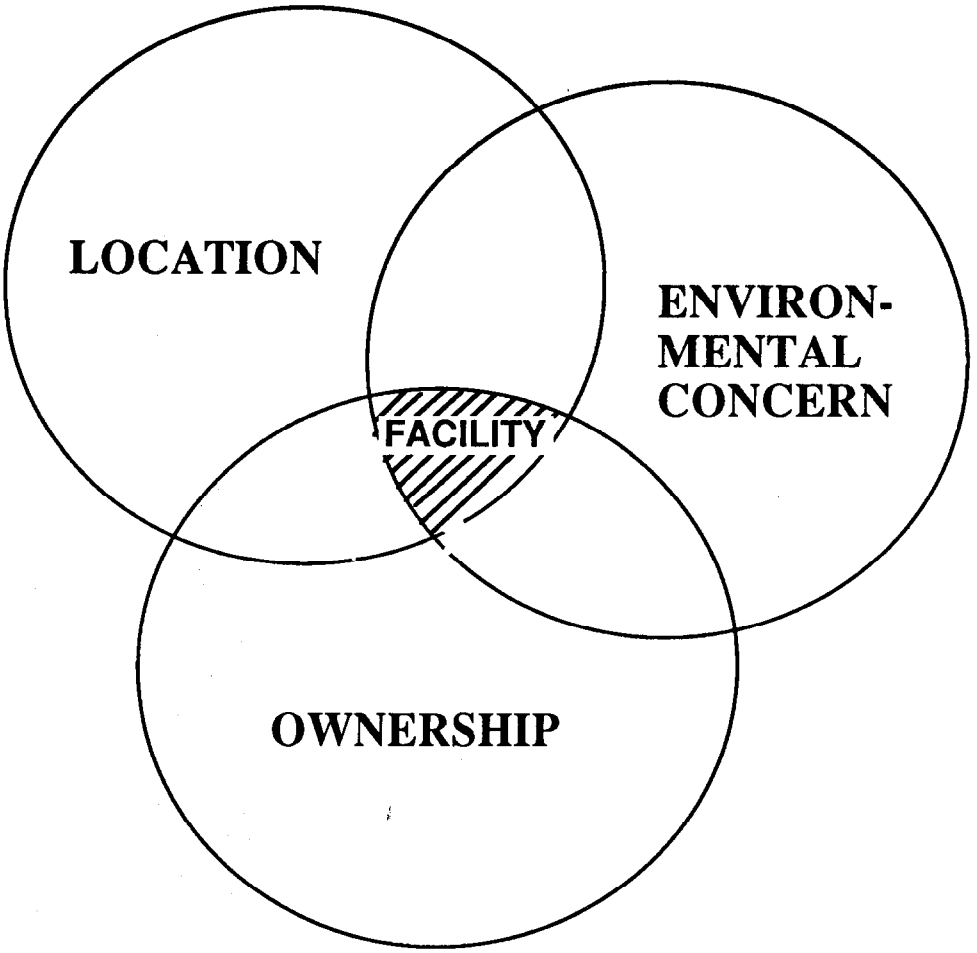
Exhibit 1-1 illustrates these three ways that an entity could be conceptualized as a "facility." Program managers, in developing and implementing their FIDS implementation plans (Chapter 4), will ultimately identify which of their regulated entities are within the scope of the FIDS. They might apply the following conditions when deciding whether an entity of environmental concern is a "facility":

- Is the entity the most all-encompassing level, defined by commonality of ownership or similar environmental circumstances?

[1] -- Appendix A, EPA Order 2180.3, 4/9/90, p. 5, Section 8(a)

Exhibit 1-1

Facility Concept



Combination of spatial extent, regulatory concern, and ownership that characterizes a "facility".

- Do activities occur that are regulated under Federal environmental law?
- Do or could those activities cause environmental concern at that location?

If an entity meets all of these considerations, whether it is regulated by EPA or a delegated state, it is within the scope of the FIDS. One rule is always true: every facility will be assigned only one EPA facility identification code, although a single facility may be regulated by many programs and thus have many program IDs.

Exhibit 1-2 summarizes the guidelines for assigning EPA facility identification codes to entities of environmental concern. Appendix B presents suggestions for what could be considered a "facility" within each national environmental program. The sections below offer more detailed guidelines for determining whether an entity of environmental concern is indeed a "facility."

1.4.1.1 REGULATED FACILITIES

The ultimate source identifying which "entities" are considered "facilities" is found within the statutes of Federal environmental laws. If a facility comes under the jurisdiction of any Federal environmental law administered by EPA, regardless of whether that law is actually implemented by EPA, a delegated state, or a local agency, then it must be included in FINDS and issued an EPA facility ID code. The responsibility for implementing the many environmental programs varies, with EPA managing some programs (Superfund, TSCA[2], FIFRA[3]) and states managing others (RCRA[4], NPDES[5], CAA[6], SDWA[7]). This fact has implications in delegating responsibilities for assigning EPA facility ID codes. Chapter 3 discusses responsibilities among the various participants of facility ID code assignment in detail.

1.4.1.2 NON-REGULATED FACILITIES

In certain cases it may be appropriate to assign EPA facility ID codes to entities that are not regulated. Such cases might include:

---

[2] -- Toxic Substances Control Act (TSCA)
[3] -- Federal Insecticide, Fungicide and Rodenticide Act (FIFRA)
[4] -- Resource Conservation and Recovery Act (RCRA)
[5] -- National Pollutant Discharge Elimination System (NPDES) under the Clean Water Act
[6] -- Clean Air Act (CAA)
[7] -- Safe Drinking Water Act (SDWA)

- Facilities that have been exempted from environmental regulation (but might otherwise be regulated if not for the exemption, for example, if they were "grandfathered out" when a law was enacted[8])

- Facilities regulated under state environmental laws that are not regulated by any Federal environmental laws

- Facilities that are *tracked* because of possible environmental effects, but are not yet regulated *per se*.

This last case shows the value of the "common sense approach" to assignment of facility ID codes. A facility which is currently tracked but not regulated, or which is regulated under state, but not Federal, law is not required to have an EPA facility ID code under the FIDS. Tracking information on that facility by assigning an EPA facility ID code to it before it actually attains "facility status" will ensure continuity in information tracking about the facility should its status change. Therefore, ***assignment of EPA facility ID codes to non-regulated facilities is encouraged (but not required).***

1.4.1.3 COMPLEX FACILITIES

The FIDS requires that the facility ID code be assigned to the most "all-encompassing" interpretation of a facility[9], meaning that what could be considered subportions of a facility (such as individual point sources, non-contiguous portions, or operable units) are not facilities themselves, but are "parts of the whole" which must be linked by the same EPA facility ID code. In other words, EPA facility ID codes are to be assigned to all environmentally-regulated facilities, and all subportions of a facility such as its outfalls and waste disposal areas are to have the same EPA facility ID code as the most all-encompassing level of what can be considered the facility[10]. This is particularly relevant to:

- ***Facilities regulated under several laws*** -- Many facilities are regulated under several environmental laws because they have subportions that meet the criteria for "regulated entity" under all those laws, such under both EPCRA[11] and NPDES.

---

8 -- For example, there have been occurrences of this situation in the reauthorization of the Clean Air Act

9 -- "*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."* FIDS, 4/9/90, Sec. 8.b., p. 6.

10-- Sub-portions have to be linked to the all-encompasing facility level in the data base so that FINDS can recognize the facility-level data within the system.

11 -- Emergency Planning and Community Right-to-Know Act (EPCRA)

Exhibit 1-2

Guidelines for EPA Facility ID Code Assignment

If a facility is...	*Then the facility ID code is assigned..*
Composed of non-contiguous or off-site portions	• To the facility in its entirety (each of the non-contiguous portions has the same facility ID code)
Part of a complex of different businesses or entities (e.g, an industrial park)	• To all portions of complex, if they are linked to each other by a common environmental concern; • To each portion, if independent concerns
Composed of sub-portions	• The facility in its entirety (each of the sub-portions has the same facility ID code)
Not regulated (e.g., tracked, exempt, regulated only by state law)	• To the facility at the discretion of the program (i.e., optionally) but recommended in case of eventual regulation

- ***Facilities with non-contiguous portions*** -- Many facilities are composed of several portions that are separated by other things (such as roads or another property) that are not part of the facility. Each of the non-contiguous portions of the facility will carry the same EPA facility ID code, indicating that a relationship exists between these different locations.

- ***Off-site portions*** -- Occasionally, adverse environmental impacts not within property boundaries of a facility can be directly attributable to activities of the facility. Certain of these off-facility sites may have cleanup activities. Or, a facility might be required to have off-site monitoring stations. Although these off-site locations may get their own program-specific ID codes, they all should carry the same EPA facility ID code as the facility with which they are associated.[12]. This is particularly true for RCRA ground water monitoring wells which are stored individually, not by facility, in STORET.

If, in a particular case, there is ambiguity about the level to which the facility ID code should be assigned (i.e., which level is the "most comprehensive"), the decision should be based upon the environmental circumstances of the site.[13] For example:

- ***If all the facilities within an industrial park are linked by a common environmental concern***, such as one that results in designating the whole park as a single Superfund site, or there is common ownership and one Form R submitted under EPCRA (Title III) is appropriate, ***then the site is assigned a single EPA facility ID code***

- ***If the entities are unique and independent*** from others with which they happen to be in close proximity (or share a location) and have no common linkage other than location, they ***could be considered separate facilities and could thus each be assigned separate EPA facility ID codes.***

Facility situations are so unique and varied that it is impossible to cover all possible scenarios or to develop a generalized rule satisfactory for all situations. ***It is therefore crucial to have professional staff involved in EPA facility ID code assignment and a mechanism in place to resolve discrepancies.*** Particularly troubling definitional cases should be brought before the Facility Identification Advisory Task Force (FIAT) for resolution. A brief description of the FIAT is given in Chapters 2 and 3.

---

12 -- Off-site portions have to be linked to the all-encompassing facility level in the data base so that FINDS can recognize the facility-level data within the system.

13 -- *"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."* EPA Order 2180.3, 4/9/90, Sec. 8(a), p. 6

1.4.2 FACILITIES THAT ALREADY HAVE "EPA ID CODES" FROM FINDS

FINDS was already operational at the time that the FIDS was established (April, 1990). Staff from the hazardous waste programs, both RCRA and Superfund, relied heavily upon FINDS to assign facility identification codes for creating records in their own program data bases and tracking other sources of data for their facilities. Therefore, identification codes had already been assigned through FINDS to many facilities, particularly hazardous waste facilities, by April of 1990. The FINDS identification codes were 12 digits long and had either a DUNS[14] number or a GSA[15] code embedded in them. Some of the programs needing these codes, in particular RCRA and Office of Federal Activities, used the codes embedded in the FINDS identification codes to sort through and select facility records, or relate those facilities to other data bases (such as Dun & Bradstreet).

These FINDS ID codes are "grandfathered" in the FIDS and do not need to be replaced with the new EPA facility identification codes. ***Facilities to which "EPA ID codes" had been assigned prior to adoption of the FIDS do not have to have them replaced with new EPA facility identification codes (the data element name, however, should be changed to "EPA Facility Identification Code").*** The appendix to the FIDS clearly states that "*...These codes do not have to be replaced, i.e., new standardized ID codes for these facilities need not be assigned.*"[16]

By "grandfathering in" the IDs assigned prior to adoption of the FIDS, as well as by assigning new IDs to new facilities according to the standard, the population of identification codes assigned by FINDS will eventually become a mixture of old and new codes. Because of this heterogeneity, program systems will no longer be able to rely upon embedded meaning in these codes. This may mean developing an independence to any meaning embedded in the ID codes (such as software that assumes that part of the ID code is a DUNS number). This aspect should be considered in the program implementation plans (Chapter 4).

---

14 -- Dun & Bradstreet DUNS Universal Numbering System
15 -- General Services Administration (GSA)
16 -- Appendix to EPA Order 2180.3, 4/9/90, Sec. 4.4, pp. A-4 to A-5

1.4.3 ENTITIES THAT ARE NOT FACILITIES

There are many entities of environmental concern that are *not* "facilities." These entities are not locationally based and/or permanent. Such entities include:

- ***Places that can be locationally identified by their latitude/longitude coordinates and about which data are collected,*** such as ambient monitoring stations, river reaches, protected habitats, or ecoregions, which are tracked or monitored for their environmental significance ***but are not (nor ever will be) regulated***

- ***Businesses that are regulated but cannot be identified by their "environmental" location because they are mobile***, such as transporters of wastes or water haulers

- ***Temporary entities***, such as highway spills that are quickly cleaned up, or portable operators, that operate at a particular location (with a permit) for a short period of time and then move to another location to operate (e.g., barges, mobile air pollution sources)

- ***Corporate locations*** that may be identifiable by their Dun & Bradstreet numbers, but at which ***no activity which could cause pollution at that location occurs,*** such as corporate headquarters offices or broker locations (these companies may be permit holders for regulated activities at other sites, and those sites would be "facilities")

- ***Places*** for which permits to build have been applied, but ***which have not yet been constructed.***

There are, and will be, entities that do not fit existing definitions or guidelines. For example, uncontrolled hazardous waste sites are almost all unique in their spatial, ecological and corporate circumstances. Therefore, careful control and experienced judgement must be used in determining whether an entity is a "facility" that requires assignment of an EPA facility ID code. The FIAT, as introduced in Chapter 2, will act as the oversight organization for FIDS implementation and will be responsible for determining whether an entity type is within the scope of the FIDS. Additional responsibilities of the FIAT will be discussed in Chapter 3.

★U.S. GOVERNMENT PRINTING OFFICE 1993-715-003/87011