ISSUES IN DIGITAL CARTOGRAPHIC DATA STANDARDS

Report #4

Digital Cartographic Data Standards: Examining the Alternatives

Harold Moellering, Editor

87-304

January, 1984

National Committee for Digital Cartographic Data Standards
Numerical Cartography Laboratory
158 Derby Hall
Ohio State University
Columbus, Ohio
U.S.A. 43210

The Committee operates under the auspices of the American Congress on Surveying and Mapping. This work of the Committee is being funded in part by U.S. Geological Survey Grant #14-08-0001-G-807. Production assistance has been provided by the Ohio State University Research Foundation (RF Project #715367).

PREFACE

This report is the fourth in the series which discusses the work of the National Committee for Digital Cartographic Data Standards. It contains five papers which discuss the alternatives available to establish digital cartographic data standards for the United States. The first paper by Moellering provides an introduction and background to the issues while the remaining papers by Edson and Moellering, Schmidt, Chrisman, and Nyerges provide detailed discussions of the alternatives for each Working Group. The Committee has organized a special session of public hearings on these alternatives at the Spring Annual meetings of the American Congress on Surveying and Mapping in Washington, D.C. on Tuesday, March 13, 1984, 8:30 A.M.

This report represents the work of the Committee for the second year of operation, that of examining the alternatives. We now invite public comment on the alternatives as presented and discussed herein. Please note that there are five sheets in the back of this report where one can provide comments and opinions for the consideration of the Committee. Please note that only written comments can be processed by the Committee due to limited staff and resources. Please send all written comments to the DCDS headquarters at the following address:

National Committee for Digital Cartographic Data Standards
Numerical Cartography Laboratory
158 Derby Hall
Ohio State University
Columbus, Ohio
U.S.A. 43210

Harold Moellering
Series Editor

TABLE OF CONTENTS

ALTERNATIVES AVAILABLE FOR THE DEVELOPMENT OF DIGITAL CARTOGRAPHIC DATA STANDARDS: BACKGROUND AND INTRODUCTION

by

Prof. Harold Moellering
Numerical Cartography Laboratory
158 Derby Hall
Ohio State University
Columbus, Ohio
U.S.A. 43210

INTRODUCTION

In recent years there has been an increasing recognition of the need to develop standards for digital cartography. To that end the National Committee for Digital Cartographic Data Standards was founded and organized in January of 1982 by the author. In order to solve any scientific problem in an efficient manner, one usually proceeds by specifying the general form of the problem and then proceeds by becoming progressively more specific. Digital cartographic data standards are no exception. The general goals of the Committee are (Moellering, 1982):

> To provide a professional forum for all involved, federal, state and local public agencies, private industry, and professional individuals to express their opinions, assessments, and proposals concerning digital cartographic data standards. After sufficient time for the formulation, circulation, discussion, reformulation and comment, these proposed standards will be submitted to the U.S. Bureau of Standards to become national digital cartographic data standards.

The primary tasks of the Committee are as follows:

1. To examine and define the scope of these standards efforts in more detail;

2. To define the number, scope, and goals of Working Groups and to appoint the groups;

3. To define general policy for the orderly examination, discussion, and adoption of the standards proposed by Working Groups;

4. To establish liaison with all interested Government agencies, private companies, academic institutions, professional societies, and groups responsible for standards in the major neighboring technical areas;

5. To issue periodic reports from Working Groups and the Committee in general; and

6. To submit to the U.S. Bureau of Standards the final proposed standards.

COMMITTEE ORGANIZATION

The Committee is organized as shown in Figure 1 into a Steering Committee,

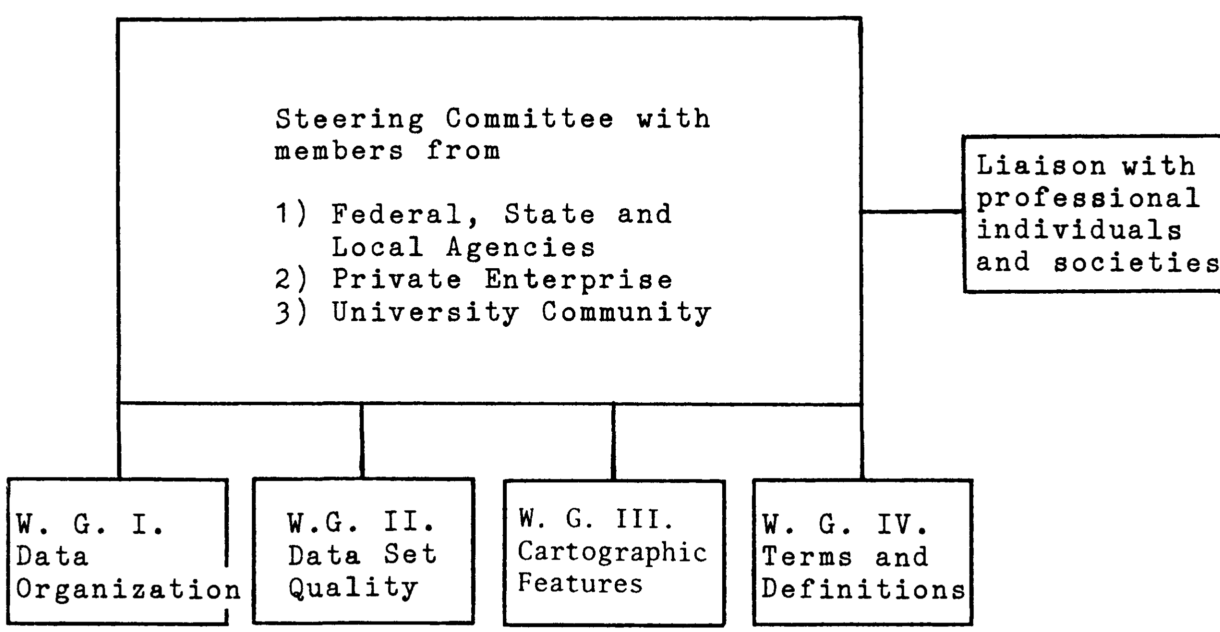


FIGURE 1. ORGANIZATION OF THE NATIONAL COMMITTEE FOR DIGITAL CARTOGRAPHIC DATA STANDARDS

four Working Groups (WG) and an Executive Committee. The membership of the Steering Committee is composed of the Chairman and eleven well known figures from the field of cartography. One should note that these individuals come from all three major segments of the profession: Federal, State and local agencies, the private sector, and academe. This mix of individuals was chosen very carefully in order to produce a reasonable balance of members from all of these three major areas of cartography.

The duties of the Steering Committee are to act as a policy review and formulation body. This body originally defined the Working Groups, their general scope, and goals. The Steering Committee continues to monitor the work of the WGs in order to insure that the efforts of the WGs systematically cover the area of their responsibility and to assure that no unnecessary overlaps or underlaps of effort occurs between WGs.

The Working Groups are the units where the bulk of the actual effort of the Committee takes place. The four WGs have been in operation for a year and a half: WG I, Data Organization; WG II, Data Set Quality; WG III Cartographic Features; WG IV, Terms and Definitions. The members of the WGs have been chosen with equal care for their expertness in the area in which the WG operates, while at the same time an effort has been made to maintain the overall balance between the three major constituencies of the profession. The WGs are examining the topics within their purview in great depth and are developing a great deal of insight into these topics.

The Executive Committee is a group composed of the Chairman and Vice Chairman of the Committee and the WG Chairs and Vice Chairs. This group manages the day to day operations of the Committee and its Working Groups and provides the organizational leadership for this standards effort.

The membership of the Committee is as follows:

<ins>Chairman</ins>: Prof. Harold Moellering, Ohio State University

<ins>Vice Chairman</ins>: Mr. Lawrence Fritz, National Ocean Service

<ins>Members of the Steering Committee</ins>:

Mr. Lawrence Fritz, National Ocean Service
Mr. Dennis Franklin, Defense Mapping Agency
Mr. Robert Edwards, Oak Ridge National Laboratory
Dr. Tim Nyerges, GeoSystems Software Inc.
Mr. Jack Dangermond, Environmental Systems Research Institute
Dr. John Davis, Kansas Geological Survey
Dr. Paula Hagan, Wang Laboratories
Prof. A.R. Boyle, University of Saskatchewan
Prof. Waldo Tobler, University of California
Prof. Dean Merchant, Ohio State University
Prof. Hugh Calkins, SUNY Buffalo

Working Group I, Data Organization

Dr. Tim Nyerges, GeoSystems Software Inc., Chair
Dr. Donna Peuquet, Univ. of California, Vice Chair
Mr. Fred Billingsley, Jet Propoulsion Laboratory
Mr. William Liles, Technology Service Corp.
Mr. Robin Fegeas, Geological Survey
Mr. Davie Pendleton, National Ocean Service
Mr. Dan Rusco, Defense Mapping Agency
Prof. Ray Boyle, Univ. of Saskatchewan
Mr. Robert Edwards, Oak Ridge National Labs.

Working Group II, Data Set Quality

Prof. Nick Chrisman, Univ. of Wisconsin, Chair
Mr. Charles Poeppelmeier, Defense Mapping Agency, Vice Chair
Dr. John Davis, Kansas Geological Survey
Prof. Dean Merchant, Ohio State University
Mr. Fred Broome, Census Bureau
Mr. George Rosenfield, Geological Survey
Mr. George Johnson, National Ocean Service
Mr. Wallace Crisco, Bureau of Land Management
Mr. John Stout, Petroleum Information Inc.

Working Group III, Cartographic Features

Mr. Warren Schmidt, Rand McNally & Co., Chair
Prof. Robert Rugg, Virginia Commonwealth Univ., Vice Chair
Dr. Joel Morrison, Geological Survey
Mr. Robert Jacober, U.S. Air Force
Mr. Richard Hogan, National Ocean Service
Dr. Beth Driver, Technology Service Corp.
Mr. Fred Tamm-Daniels, Tennessee Valley Authority
Ms. Mary Clawson, ITT Research Institute

Working Group IV, Terms and Definitions

Mr. Dean Edson, E-Quad Systems, Co-chairman
Prof. Harold Moellering, Ohio State University, Co-chairman
Mr. Erich Frey, National Ocean Service, Vice Chair
Prof. Hugh Calkins, SUNY Buffalo, link to WG I
Mr. Frank Beck, Geological Survey, link to WG II
Prof. Mark Monmonier, Syracuse University, Link to WG III

Observers:

Mr. Ben Ramey, Geological Survey
Mr. Lowell Starr, Geological Survey
Mr. Henry Tom, Bureau of Standards
Mr. Roy Saltman, Bureau of Standards

Ex Officio:

Mr. Walter Robillard, President, American Congress on Surveying and Mapping
Mr. R. Anthony Novotny Jr., President, American Cartographic Assn.
Mr. John Uehlinger, Executive Director, American Congress on Surveying and Mapping

CONCEPTUAL BACKGROUND AND GENERAL TASKS OF THE COMMITTEE

The conceptual background for this effort has been defined in detail in an article in NCDCDS Committee Report No. 1 (Moellering, 1982) but shall be briefly summarized here. The conceptual milieu in which the committee is operating has also expanded dramatically in recent years as noted in the earlier paper. The concepts of real and virtual maps greatly clarify the situation of the new digital cartographic products and how they relate to the more conventional products (Moellering, 1980). Transformations between real virtual maps define most important operations in cartography and have been an interesting concept for the design of modern cartographic systems. Nyerges (1980) has devised the notions of deep and surface structure as they apply to cartographic information and has shown that surface structure representations of cartographic information are real and virtual type I maps while cartographic deep structure is usually represented in the digital domain by type III virtual maps. It is also possible to look at these standards efforts in terms of deep and surface structure. Surface structure is the graphic representation of cartographic information such as a conventional map or CRT display. Over the years many principles have been defined for cartographic design which must be followed if one is to have an effective map. However, the deep structure, that area of spatial relationships between cartographic elements of cartographic information which are not graphic, is where much of the digital information resides which is stored in modern cartographic data bases. In essence, the primary task of this committee is to bring conceptual order to the area of deep structure in digital cartography.

The question now becomes one of how a Committee such as this can help to bring conceptual order to the area of deep structure in digital cartography. Early meetings of the Steering Committee in the Spring of 1982 revealed that one of the most pressing problems in digital cartography is that which arises when one endeavors to use a data base that was developed, compiled and built by an organization other than one's own. The problem is that to a very large extent such data bases are incompatible with one another. Such incompatibilities arise for several reasons:

1. the inherent nature of the information being captured is different (e.g. topological data vs. geometry only),

2. the data models, data structure and data organizations being used are different,

3. the quality of the data varies widely and in many instances is not even assessed,

4. many definitions for cartographic features conflict with each other which means that the feature codes do it,

5. the terms and definitions being used in all of the instances above are used in widely varying ways which are usually inconsistent

Although one cannot directly change situation 1), it is feasible to do something about the next four. It is because these are major problems facing the field that the Steering Committee defined the four Working Groups as they now stand.

At this point, it is useful to point out some things that the Committee is not doing. The Committee is not doing any work in cartographic communication. A lot of research has been conducted on this topic in the past, and the general principles of map design and communication are reasonably well understood, although a comprehensive work beinging these findings together in a systematic manner has yet to be written. A second area in which the Committee is not becoming involved is that of cartographic hardware. Although at the outset it might appear that this area should be examined by the Committee, the nature of proprietary rights precludes it.

THE WORKING GROUPS

The Working Groups as originally defined have been following a set of basic goals:

1. To assess the state of current knowledge and understanding in the technical area,

2. Define any gaps in such knowledge and understanding necessary to specify digital cartographic standards in that area,

3. To invite presentations and opinions from all interested parties relating to the standards area,

4. To prepare technical working papers of their deliberations and discussions.

These first two stages of the work have concerned the specifying of the issues and gaps in our knowledge, and specifying possible alternative solutions for standards. The general tasks for the WGs are as follows:

I. Working Group on Data Organization

   1. Examine cartographic data models
   2. Examine cartographic data structure
   3. Examine cartographic data interchange

II. Working Group on Data Set Quality

1. Fidelity of graphical data, metric and topological
2. Coding reliability
3. Update and other temporal information
4. Lineage of a data set
5. Checking procedures used by the producer to verify quality

III. Working Group on Cartographic Features

1. Define feature classes
2. Define structure and levels of classes
3. Define feature codes

IV. Working Group on Terms and Definitions

1. Collect new terms defined by working groups
2. Define other new terms

More recent efforts by the WGs have refined the original statement and the following is a brief summary of the current direction of progress being made by each WG and tasks for the coming year.

Working Group on Data Organization

The scope and goals of WG I are to identify problems in cartographic data interchange and their consequences at the operational and conceptual levels. The work has concentrated on existing data bases and data models with an emphasis on high speed transfer of, and the possibility of homeomorphisms between, large data bases. The WG has been identifying terminology and definitions of terms currently being used in the area.

The current work of this group has revealed that it is not feasible to try to specify a single format for data exchange because the cartographic data world is far more complex than that, especially of one considers both vector and raster structures. Rather the movements has been in the direction of some kind of "family of formats" approach. Two major alternatives are 1) developing a superstructure which contains a small number of defined formats which can handle most kinds of cartographic data structures, or 2) to pursue a Data Description Language (DDL) which is standardized and in its turn defines the data format in hand. In the next year the major challenge will be to make some difficult choices from these two different approaches in order to define an interim standard.

Working Group on Data Set Quality

When one receives a data set from some source other than one's own organization, in most cases, there are a lot of questions about data set quality which are not easily answered. For example, it is not usually known what the original data source(s) was and what scale(s) the data were gathered. It is usually not known what the original coordinate system was and to what ellipsoid they were associated. The error rates for the coding of substantive data are usually not specified, nor does one know if this data set has ever been updated. There are many attributes of a data set which should be made known to the prospective user of that data set which seem to fall into five basic categories: fidelity of graphical data, metric and topological; coding reliability; update and other temporal information; lineage of a data set; and checking procedures used by the producer to verify quality. This sort of information would be very informative to the user and indeed be very helpful in deciding whether a particular data set could successfully be used for a particular purpose.

The Working Groups has not specified the basic categories of information which should be provided to the potential user and has also examined alternative methods for specifying these mathematically and logically. In the next year the task is now to choose the most effective approach and to write them up in a tight and logically understandable statement of interim standards.

Working Group on Cartographic Features

The fundamental challenge of the Group is to harmonize the feature definitions and coding schemes used by the major agencies in cartography. The current work of the group indicates that such a scheme should probably be scale independent, not directly tied to any fixed data model, but rather to a more flexible schema, attributes should be allowed to be multiple and accurately describe the feature characteristics. The group has also been collecting substantive definitions of the features themselves. The work for the coming year involves gathering together a comprehensive list of such definitions and subsequently making choices of the preferable definition that produces the best coverage for the entire set of definitions chosen for an interim standard.

Working Group on Terms and Definitions

These efforts have unearthed terms and definitions which have not been defined in a way which is universally acceptable. Although a fair amount of work had already been expended in producing the International Cartographic Association glossaries of terms and definitions, there are many terms in numerical and analytical cartography which are being used in this effort which have not been previously concisely defined. A system has now been devised to effectively handle the terms generated by the other Working Groups and a method for processing comments concerning the definition of these terms is in operation.

A second task has involved an attempt to bring order to the terms used for cartographic objects. An analysis of the alternative strategies has been conducted and is currently being extended. The task for the coming year is to make the difficult choices that best reflect the meanings of the terms used and of the cartographic objects recognized in such a way that they harmonize with those used in other disciplines.

OTHER STANDARDS EFFORTS IN CARTOGRAPHY

It turns out that there are several other efforts underway in various parts of the world to develop digital cartographic data standards. The motivation for these efforts is essentially the same as for the NCDCDS, that of reducing the complications produced when utilizing data bases obtained from other organizations. They will be discussed in the rough order that they were founded, Australia, Canada, United States, United Kingdom, and the International Hydrographic Organization. It can be anticipated that more such groups will be founded elsewhere in the world in the future.

Australia

The Standards Association of Australia (1981) published a standard for the interchange of feature coded cartographic data. It was developed by the Institute in cooperation with more than 20 organization as participants. As such the standard specified coding methods and data structures for features on maps and charts.

Canada

In April of 1982 the Canadian Council on Surveying and Mapping (1982) issued a three volume draft report which presents proposed standards for topographic features, quality evaluation of topographic data, and EDP standards for that data. The drafts are under discussion at the present time.

United States

As the reader may be aware, a Federal Inter-Agency Coordinating Committee on Digital Cartography has been formed during the past year. It was formed as a result of a report by the General Accounting Office (1982) on duplicative efforts in the Federal agencies in digital cartography, and a mandate from the Federal Office of Management and Budget (1983). Although the Committee is still in the early stages of formation, five subgroups have been formed, one of which is concerned with standards. The remaining groups cover other topics.

Although the Federal Committee is primarily interested in cartographic activities in the Federal sector while the NCDCDS is concerned with cartographic activities in the entire profession at a more general scale, there are several areas of common interest. These areas are currently being explored and methods of cooperation and coordination between the two committees are being examined. Fruitful results are anticipated.

STANDARDS WORK IN COGNATE AREAS

Although almost all of the activities of the Committee have been focussed on the cartographic body of knowledge, there are standards efforts going on in a number of cognate areas that are related to cartography. From that point of view it is important that the Committee remains informed of the standards activities in such areas. Figure 2 shows a schematic diagram

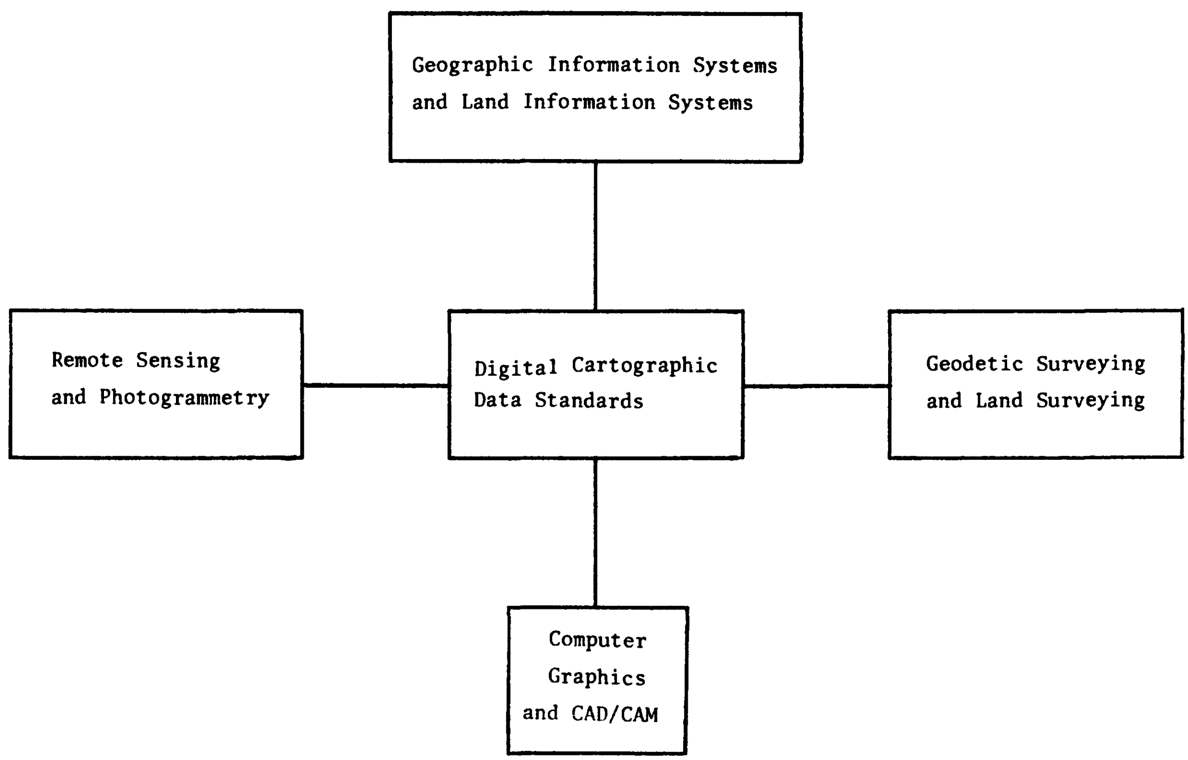


FIGURE 2. COGNATE AREAS WHICH INTERFACE TO DIGITAL CARTOGRAPHIC DATA STANDARDS

of the cognate areas for cartography. In each area standards efforts are either going on now, or have taken place in the past. It is clear that the development of digital cartographic data should not take place in isolation. Therefore efforts have been going forth to establish liaison relationships with other professional organizations which play a cognate role in relation to digital cartography. Naturally, it is also important that such areas are kept informed of recent developments occurring in digital cartography.

FUTURE WORK

The DCDS Committee is now wrapping up the second cycle of work on examining the alternatives. This report presents the work of the Committee. The final step for this cycle is to obtain comments from the cartographic profession on this work and the alternatives. All members of the cartographic community are invited to participate in this process by sending their written comments to the Committee. Standardized comment forms are found at the back of this report. Please send all written comments to:

National Committee for Digital Cartographic Data Standards
Numerical Cartography Laboratory
158 Derby Hall
Ohio State University
Columbus, Ohio
U.S.A. 43210

As part of this process of gathering comments from the profession, a special all day session is being planned for the 1984 ACSM meetings in Washington, D.C. where public hearings will be held.

At the Spring ACSM meetings the Committee will begin work on the third cycle, that of developing a proposed interim standard. In the work of the third year the Committee will carefully considered the comments received from the cartographic community while weighing the merits of various alternatives as the Committee moves towards making a decision for the interim proposed standard.

SUMMARY AND CONCLUSIONS

The committee is now in the final stages of the second cycle of examining the alternatives for digital cartographic data standards, with only the public comment and discussion remaining. The Committee will then move to the third cycle, that of developing an interim proposed standard.

REFERENCES

Canadian Council on Surveying and Mapping, 1982, National Standards for the Exchange of Digital Topographic Data, three volumes, Ottawa: Topographic Survey Division and Mapping Branch, Mines and Resources Canada.

General Accounting Office, 1982, Duplicative Federal Computer-Mapping Programs: A Growing Problem, Comptroller General of the United States Report #GAO/RCED-83-19, Gaithersberg, MD: General Accounting Office.

Moellering, Harold, 1980, "Strategies of Real-time Cartography" Cartographic Journal, 17(1), pp. 12-15.

__________, 1982, "The Challenge of Establishing a Set of National Digital Cartographic Data Standards for the United States," Proc. American Congress on Surveying and Mapping 42nd Annual Spring Meeting, Denver, pp. 201-212.

Nyerges, Timothy, 1980, Modeling the Structure of Cartographic Information for Query Processing, unpublished Ph.D. dissertation, Department of Geography, Ohio State University.

Office of Management and Budget, 1983, "Coordination of Federal Digital Cartographic Data Programs" Memorandum of April 4.

Standards Association of Australia, 1981, Interchange of Feature Coded Digital Mapping Data, Australian Standard 2482-1981, Sydney; Standards House.

DIGITAL CARTOGRAPHIC DATA STANDARDS:
ALTERNATIVES IN TERMS AND DEFINITIONS
A Progress Report

Work Group IV, NCDCDS

Part 1: Proceedural Alternatives
Prepared by Dean T. Edson

Part 2: The Definition of Fundamental Cartographic Objects
Prepared by Harold Moellering

WORK GROUP IV MEMBERSHIP

Mr. Dean T. Edson, Co-chairman
E-Quad Associates, Placerville, CA

Prof. Harold Moellering, Co-chairman
Ohio State University, Columbus OH

Mr. Erich Frey, Vice Chairman
National Ocean Survey, Rockville, MD

Mr. Frank Beck
U.S.Geological Survey, Reston, VA

Prof. Mark Monmonier
Syracuse University, Syracuse NY

Prof. Hugh Calkins
SUNY, Buffalo, NY

PART 1

PROCEDURAL ALTERNATIVES

Introduction

The primary problem facing the Terms Work Group was, is, and will remain the detection and resolution of confused communication due to terminology problems. It is axiomatic that the basis of communication is (1) the idea and (2) a common understanding of the language used to convey the idea.

Over the past two decades, the profession has become familiar with a new "word-of-mouth" or system specific vocabulary, and as we babble on and on, has sometimes contributed more to confusion than understanding. The real culprit is the accelerated growth of the technology on multiple fronts. Here, changes in syntax and idiom have followed this growth, and even when we start from an established, well-understood cartographic or mathematical root, we end up not "really" understanding one another.

There is a requirement, therefore, as we proceed with the task of defining standards, that we concurrently re-examine the entire vocabulary, compare existing terms with usage in the existing literature and consider each term for suitability in the conveyance of understanding.

This report on alternatives is presented in two distinct segments. The first segment covers the procedural alternative of the terms and definitions task with an eye on public involvement. The second segment focuses on the major terms problem related to the definitions of fundamental cartographic object. In this section we have attempted to present the concensus of term usage based on the profession, as a whole, through a thorough search of recent literature, discussions within the entire DCDS Committee, and a profound awareness of the volatility of the alternatives presented.

In view of the underlying critical nature of the Terms and Definition task, our work will continue to proceed in concert with the other Work Groups. The tune, we hope, will catch on.

Collecting and Validating Terms

The overall approach to be used in collecting appropriate terms and definitions consists of two channels from the other Work Groups to the Terms and Definitions Group.

The first channel is the members of WG 4 assigned to monitor each of the other three work groups. The monitoring process consists of personally attending work group sessions and reviewing appropriate Continuum statements, draft reports from the work groups, and recent technical literature referenced in the Committee's bibliography.

Work Group 4 member monitor assignments are as follows:

Work Group	W.G.4 Monitor
1. Data Organization	Prof. Hugh Calkins
2. Data Quality	Mr. Frank Beck
3. Data Features	Prof. Mark Monmonier

These monitor assignments have been operational since the Committee sessions in Salt Lake City.

The second channel consists of terms and definitions suggested by individual Work Group members through their Work Group chairman.

The terms collection method is not rigid. It provides a reasonable way of insuring that questionable terminology be addressed and when appropriately defined, placed in a Public Comment Glossary.

Most of the suggested terms and definitions have been collected through the Define Continuum of Cartnet. However, hand written notes and word of mouth also have been helpful in channelling significant problems and useful comment to the Terms and Definitions Group.

Definitions Procedures Using Cartnet

The basic requirement in collecting terms and proposing definitions will be to present them for public comment in a format that will encourage individual comment.

As a means of managing this task, three new continuums established as a part of Cartnet are:

Define Continuum, used to receive terms or definitions proposed by any member of the DCDS Committee. This continuum is available with both read and write access and provides a general terms collections point with no format rules but is reserved for terms and definitions commentary.

Terms In-Process Continuum, used to display terms selected for definition along with all proposed or alternate definitions. This continuum was established to operate as a read-only segment. Any comment at this stage would be entered in the Define Continuum. Write access to this continuum is reserved for use by members of the Terms Work Group.

Terms Defined, used to contain Terms and definitions which are ready for viewing and comment outside the DCDS Committee. Text from this continuum will be formated for hard- copy generation for both public comment (draft format) and Proposed Standard Terms and Definitions (published format). There again, the Define continuum is established for read only except for members of the Terms Work Group.

Figure 1 graphically relates these continuums and the flow of term-related information from the Work Group/monitor to a proposed Standard Term Glossary. As noted in this figure, we are currently using an off-line word processor to create a temporary working link between the Define Continuum and a Public- comment type document. We plan to implement a processing link from the Define continuum through the word processor located at the office of the Co-chairman (Dean Edson) and back to the Terms in-process continuum. This return data link will be accomplished by an appropriate modem and voice grade phone line. The off-line word processing capability uses the C/PM based WordStar software.

As public comment finds its way back to the Terms Work Group, we will, as a committee, distill, organize, and circulate these comments back through the system starting with the Work Group IV monitor to each of the Work Groups I, II and III.

Format for Public Comment Glossary

As a convenient means of presenting the preliminary results of DCDS Committee's effort to provide a standard definition for a large number of key terms, we will produce a public comment document based on the format illustrated in figure 2. This format is intended to provide public comment pages on which other DCDS participants can register their opinions and insights.

Using this approach, each public comment page will be provided with blank comment space and easy removal. The back of each page will be designed as a return envelope with folding instructions and a return address.

Glossary Distribution Alternatives

At the panel session held by DCDS Committee as part of the Fall ACSM-ASP meeting in Salt Lake city, comments made during the question/answer period suggested Terms and Definitions be made available to the public in machine readable form, or by direct access from a computer based text system. We plan to explore various methods of distributing both the Public Comment and Standard versions of the Glossary. In its present form, the Public Comment version is maintained on 5¼" floppy diskettes, which are easily duplicated and mailed at a minimum cost.

Status of Terms and Definitions

As part of this progress report, we are including the contents of the Term In-process file. Here, several terms and definitions are presented in our proposed Public comment format. These pages are intended to display the approach taken to obtain comment. They will be duplicated and greatly expanded for our Public Comment Sessions to be held during the Annual ACSM-ASP meeting in Washington, D.C. this coming spring.

TERMS AND DEFINITIONS PROCEDURE

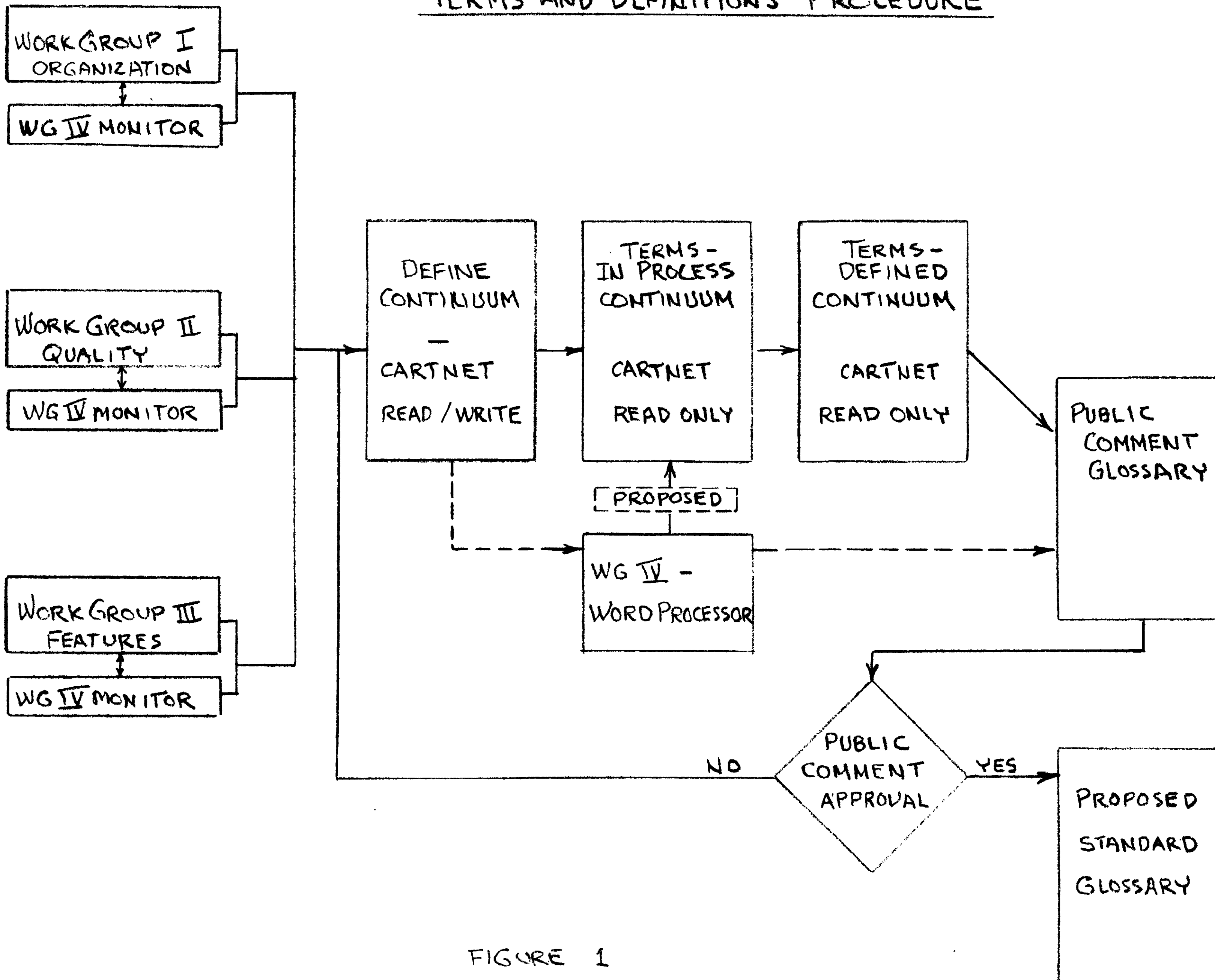


FIGURE 1

EDSON 11-83

PROPOSED STANDARD TERMS AND DEFINITIONS

TERM	DEFINITIONS
This column will contain the complete term with an explanation of acronyms or other unusual derivations.	This column will contain the committee approved definitons and will use an underscore to identify a term used in the definition which is self defined in this glossary. REFERENCE: This space will also contain any meaningful comment which might clarify source, usage, aliases or any unusual characteristics.

Space for Comment

This is a proposed sample format for the preliminary publication of terms where public comment is invited. Final format could be loose leaf, updatable paging. All records to be maintained using machine readable media. (Initial files will be compiled using 5¼ floppy diskettes formated using C/PM WordStar).

Figure 2.

PART 2

THE DEFINITION OF FUNDAMENTAL CARTOGRAPHIC OBJECTS

Introduction

When one represents a spatial entity from the real world as an object in a data structure, a wide variety of terms, many of which conflict in one way or another, have been used to name those objects. Table 1 illustrates the current situation of conflicting terms for fundamental cartographic objects. (This table is based on the one originally presented by Anderson and Calkins, (1982).)

	GIRAS (USGS)	DLG (USGS)	CGIS (Canada)	POLYVRT (Harvard)	Cook (1978) (C.S.I.R.O.)
0-D	node	node	point	node/point	junction
1-D	arc	line	face/edge	chain	line
2-D	polygon	area	polygon	polygon	region

Table 1. Diverse Names for Similar Cartographic Objects

Anderson and Calkins note seven different terms that have been used to describe "a line": arc, line, chain, segment, edge, face, link. Such terms have been used to name objects which are essentially the same, but have used different names, and also to name different objects with the same name. (See, for example, a discussion from the Harvard proceedings, Anon. (1978).) Upon careful reflection on this problem, it has become clear that one of the primary sources of such different usage of these terms is the diversity of backgrounds of the individuals in the field of cartography itself. For example, many traditional cartographers are primarily interested in map production and other surface structure representations of such cartographic information. This generally led individuals to lean very heavily on geometry and associated coordinates, and hence most of the time these individuals use the terminology from geometry. A second set of individuals in the field has been more interested in analytical approaches to cartographic information in terms of data organization requirements and other deep structure approaches which depend on mathematical approaches such as topology and graph theory. This approach brings with it a somewhat different point of view and, consequently, a different set of terminology has been used. Although verbal clashes occurred at meetings a few years ago, the sentiment which has developed in recent years has recognized the need for both approaches. It is safe to say that most digital data work in cartography currently uses concepts from geometry, topology and graph theory.

The purpose of this paper is to harmonize the terms used for cartographic objects into a compatible whole and at the same time to recognize the diverse needs of cartographers in the area of geometry, topology and graph theory. While most work today utilizes all three areas, one must also provide terms for objects which can be used solely for geometric applications in cartography, or can be used solely for topological and graph theoretic applications in the field. In this discussion, a review of primitive cartographic objects will be undertaken as well as a number of simple and very commonly used objects which can be built up from the primitive objects. More complex and compound cartographic objects can be built up from the simple and primitive objects (See for example, Youngmann (1978) or Nyerges (1980)), but will not be directly discussed here. However, consideration will be given to be sure that such compound and complex cartographic objects can be successfully constructed from the simpler objects defined.

For convenience, the discussion in the next section will begin with the 0-dimensional cartographic objects and will work through the 1- and 2-dimensional objects. The goal is to produce a compatible set of names and definitions for well understood cartographic objects. Three-dimensional objects are much less well understood, and since this work is only attempting to sort out names and definitions of well understood objects, 3-dimensional objects will not be discussed here.

As an alternative these objects could be discussed under the rubric of n-cells (0-cell, 1-cell, 2-cell, etc.) from topology as presented by White (1979). Here, as will be seen, some objects reflect only geometry and no topology, so it is not clear whether the n-cell terminology applies in all cases. Therefore, the n-cell terminology will not be used in the balance of this discussion.

0-Dimensional Cartographic Objects

Punctiform cartographic objects are all primitive objects that cannot be subdivided. However, one must be cognizant of both geometric and topological applications in cartography.

<u>Alternative 0-1</u>

point - A 0-dimensional object that specifies geometric location. A set of coordinates specifies the location.

node - A 0-dimensional object that acts as a topological junction. No coordinates are associated with it.

nodal point - A 0-dimensional object that is a topological junction and specifies geometric location. A set of coordinates specifies the location.

<u>Alternative 0-2</u>

point – same as above

node - same definition as nodal point above.

Discussion

Since cartographic work can occur in three modes, geometry only, topology only, and both combined, one possibility is that there could be three classes of 0-dimensional objects. This alternative explicitly defines objects for all three kinds of work, although the term "nodal point" had to be coined to resolve the question.

The second possibility is to have just two classes: geometry only and topology with coordinates. The problem with this scheme is that cartographers working only with topology, (see White and Griffen, 1979) do not have a clearly defined object term that can be used for their work without ambiguity. However, it is probably true that for a fairly large percentage of cartographic applications the problem is not that severe.

1-Dimensional Cartographic Objects

Linear objects are bounded by and defined by 0-dimensional objects. The generic term for a 1-dimensional object is that of a line. The question now is just how that linear object is defined. It should be recognized at the outset that continuous lines utilize discrete elements when processed by digital systems.

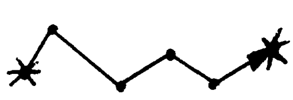
Alternative 1-1

line segment - A 1-dimensional object that is a straight line between two points. Used for geometric drawing.

link - A 1-dimensional object that is a direct connection between two nodes. Used for topological analysis. Alias: edge.

arc - A directed link between two nodes.

string - A series of line segments strung together.



chain - A directed series of line segments strung together with nodal points at each end of the string. (See Alternative 0-1)

Note: Nodal Points may also be used as the bounding points of arcs and links.

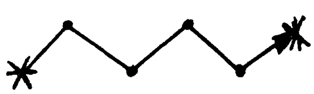
Alternative 1-2

 line segment - same as above

 link - same as above

 string - same as above

 directed link - A link between two nodes with direction specified.

 chain - same as above. Alias term: arc.

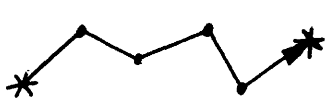
Alternative 1-3

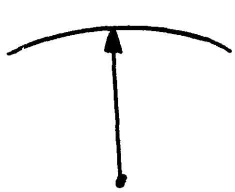
 line segment - same as above

 link - same as above

 string - same as above

 directed link - same as alternative 1-2

 chain - same as above, with no alias for arc

 arc - a locus of points with a constant radius

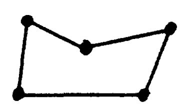
Discussion

The definitions for the terms, segment, link, string, and chain are clear cut with no adjustment required. Line segments are used for geometric applications while link are used for topological applications. A chain can be used for both. The primary problem is associated with the term, arc. The definition which takes historic precedence comes from Euclidean geometry as a locus of points which have a constant radius and has been used to define an object in the CAD/CAM standards. The common usage in digital cartography is with a topological object which has nothing to do with curvature. However, the term is well enough entrenched in cartography as a topological object that an effort has been made to utilize the term in a way which is less ambiguous. Alternative 1-1 uses the term arc as a directed topological link, while Alternative 1-2 uses the term arc as an alias for the term chain, a term which does not carry with it the complicated historical baggage of the term arc, while Alternative 1-3 uses it in a way that is consistent with the CAD/CAM usage. This question needs full discussion in the profession.

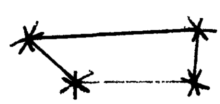
2-Dimensional Cartographic Objects

Areal objects can be defined in two fundamental ways, one by building up a simple object from 0- and 1-dimensional objects, and the second by recognizing a separate primitive called a pixel.

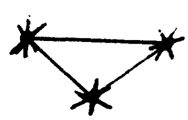
polygon - A 2-dimensional planar object that can be formed in four ways:



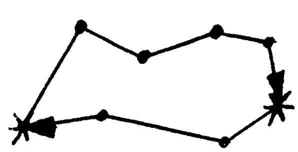
1) area bounded by a sequence of points and line segments with closure



2) area bounded by a sequence of nodes and links with closure

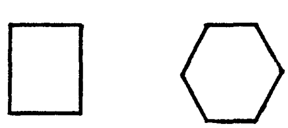


3) area bounded by a sequence of nodal points and links with closure



4) area bounded by a chain (s) which have closure

pixel - A picture element of an area on the ground in a nondivisible measurement. An array of pixels will form a regular tesselation of a plane. Common shapes are quadralaterals and hexagons, although other shapes are possible



Discussion

The specification of 2-dimensional cartographic objects implies that they are either polygons or pixels. Polygons have been defined in several ways with different kinds of objects as shown in the definitions. One question that arises is whether a polygon can really be defined with nodes that contain no coordinates. The answer to that question is not clear at this point. A second question that arises has to do with holes in 2-dimensional objects. It has been argued that a hole, if it exists, is an integral part of an areal object, but that the object should not have artificial cuts in it (White, 1979). White referred to the work of Corbett (1979) which relies on homology theory to solve the problem. The implications for numerical cartography are not all that clear. It seems that holes in cartographic objects constitute a gap in our knowledge. Both of these questions need further discussion in the profession.

Summary of the Work on Cartographic Objects

It is clearly recognized that the current confusion in terminology for cartographic objects has its origins in the diversity of the field itself. This discussion is an attempt to devise a compatible set of terms for cartographic objects which are internally harmonious and externally mesh with the terms used in other fields. As is described by the alternatives presented, this goal can be met in several ways. It is now up to the profession to provide comments so that an efficient choice can be made.

LITERATURE

Alexandroff, P., 1965, Elementary Concepts of Topology, translated by A. Obolensky, New York: Federick Ungar Publishing Co.

Anderson, K., and H. Calkins, 1982, "Cartographic Data Standards: Implications for GISs", paper presented at 42nd annual meeting of the American Congress on Surveying and Mapping, Denver, CO.

Anon., 1978, "Proposed Terminology for Symposium Discussion", in G. Dutton, ed., Summary of Proceedings, First International Study Symposium on Topological Data Structures for Geographic Information Systems, Vol. 1, Cambridge, MA: Laboratory for Computer Graphics, Harvard University, pp. 128-131.

Biggs, R., 1978, "Cartographic Data Structures for Thematic Applications of Automated Cartography", in G. Dutton, ed., Data Structures: Hierarchial and Overlaid, First International Advanced Study Symposium on Topological Data Structures for Geographic Information Systems, Vol. 4, Cambridge, MA: Laboratory for Computer Graphics, Harvard University.

Billingsley, F., P. Anuta, J. Carr, C. McGillem, D. Smith, and T. Strand, 1983, "Data Processing and Reprocessing", Chap 17 in R. Colwell, D. Simonett, and F. Ulaby, eds., Manual of Remote Sensing 2nd ed., Vol. I, Falls Church: ASP, pp. 719-724.

Brassel, K., 1978, "A Topological Data Structure for Multi-Element Map Processing", in G. Dutton, ed., Data Structures: Hierarchial and Overlaid, First International Advanced Study Symposium on Topological Data Structures for Geographic Information Systems, Vol. 4, Cambridge, MA: Laboratory for Computer Graphics, Harvard University.

Chinn, W., and N. Steenrod, 1966, First Concepts of Topology, New York: Random House.

Claire, R., and S. Guptill, 1982, "Spatial Operators for Selected Data Structures", International Symposium on Computer Assisted Cartography (Auto-Carto V), Falls Church: ASP & ACSM, pp. 189-200.

Cook, B., 1978, "The Structural and Algorithmic Basis of a Geographic Data Base", in G. Dutton, ed., Data Structures: Hierarchial and Overlaid, First International Advanced Study Symposium on Topological Data Structures for Geographic Information Systems, Vol. 4, Cambridge, MA: Laboratory for Computer Graphics, Harvard University.

Corbett, J., 1979, Topological Principles in Cartography, U.S. Department of Commerce, Bureau of the Census, Technical Paper No. 48, Washington: Government Printing Office.

Coxeter, H., 1961, Introduction to Geometry, New York: John Wiley & Sons.

Dangermond, J., 1982, "A Classification of Software Components Commonly used in Geographic Information Systems", Proc. US/Australia Workshop on the Design and Implementation of Computer Based GISs, Honolulu: IGU, Commission on Geographical Data Sensing and Processing, pp. 70-91.

Dueker, K., 1972, "A Framework for Encoding Spatial Data", Geographical Analysis, 4(1), pp. 98-105.

Dueker, K., and J. Noynaert, 1978, "Interactive Digitizing, Editing, and Mapping: Data Structure Considerations", in G. Dutton, ed., Spatial Semantics: Understanding and Interacting with Map Data, First International Advanced Study Symposium on Topological Data Structures for Geographic Information Systems, Vol. 7, Cambridge, MA: Laboratory for Computer Graphics, Harvard University.

Edwards, R., R. Durfee, and P. Coleman, 1978, "Definition of a Hierarchial Ploygonal Data Structure and the Associated Conversion of a Geographic Base File from Boundary Segment Format", in G. Dutton, ed., Data Structures: Hierarchial and Overlaid, First International Advanced Study Symposium on Topological Data Sturctures for Geographic Information Systems, Vol. 4, Cambridge, MA: Laboratory for Computer Graphics, Harvard University.

Fegeas, R., 1978, "The Graphic Input Procedure - An Operational Line Segment/Polygon Graphic to Digital Conversion", in G. Dutton, ed., Spatial Semantics: Understanding and Interacting with Map Data, First International Advanced Study Symposium on Topological Data Structures for Geographic Information Systems, Vol. 7, Cambridge, MA: Laboratory for Computer Graphics, Harvard University.

__________, 1982, "Modularization of Digital Cartographic Data Capture", International Symposium on Computer-Assisted Cartography (Auto-Carto V), Falls Church: ASP & ACSM, pp. 297-306.

Freeman, H., 1961, "On the Encoding of Arbitrary Geometric Configurations" IRE Transactions on Electronic Computers, Vol. EC-10, pp. 260-268.

__________, 1978, "Some New Map Data Encoding Schemes", in J. Moneta, ed., Information Technology, Amsterdam: North Holland, pp. 689-696.

[illegible] F., 1969, Graph Theory, Reading, MA: Addison-Wesley.

Henderson, T., 1983, "Efficient 3-D Object Representations for Industrial Vision Systems", IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-5(6), pp. 609-618.

Huang, X., and S. Hsu, 1983, "A Chained Data Structure for Efficient Extraction and Display of Information from Polygon Maps", Proc. ASP Spring Meetings, Washington, Falls Church: ASP, pp. 171-176.

Morehouse, S., 1979, "An Application of ODYSSEY to DIME File Mapping", Mapping Software and Cartographic Data Bases, Harvard Library of Computer Graphics, 1979 Mapping Collection, Vol. 2, pp. 157-162.

Nyerges, T., 1980, Modeling the Structure of Cartographic Information for Query Processing, unpublished Ph. D. dissertation, Dept. of Geography, Ohio State University.

Peuquet, D., 1982, "Vector/Raster Options for Digital Cartographic Data", Proc. US/Australia Workshop on the Design and Implementation of Computer Based GISs, Honolulu: IGU, Commission on Geographical Data Sensing and Processing, pp. 29-35.

__________, 1983, "A Hybrid Structure for the Storage and Manipulation of Very Large Spatial Data Sets", Computer Vision, Graphics, and Image Processing, 24(1983), pp. 14-27.

Pfaltz, J., 1976, "Surface Networks", Geographical Analysis, 8(1), pp. 77-93.

Requicha, A., and H. Voelcker, 1982, "Solid Modeling: A Historical Summary and Contemporary Assessment", IEEE Computer Graphics and Applications, 2(2), pp. 9-24.

Taketa, R., 1979, Structure and Meaning in Map Generalization, unpublished Ph. D. dissertation, Dept. of Geography, University of Washington.

U.S. Department of Energy, 1978, "Appendix A - Structures for Representing Geographical Features", in Geographic Exchange Standard and Primer, prepared by Battelle, Pacific Northwest Laboratory, Interlaboratory Working Group for Data Exchange, PNL-2748, UC-32, Springfield, VA: National Technical Information Service.

U.S. Department of the Interior, Geological Survey, 1983a, Digital Line Graphs from 1:2,000,000-Scale Maps, by M. Domaratz, C. Hallam, W. Schmidt, and H. Calkins, in R. McEwen, R. Witmer, and B. Ramey, eds., USGS Digital Cartographic Data Standards, Circular 895-D, Reston, VA: USGS.

__________, 1983b, Land Use and Land Cover Digital Data, by R. Fegeas, R. Claire, S. Guptill, K. Anderson, and C. Hallam, in R. McEwen, R. Witmer, and B. Ramey, eds., USGS Digital Cartographic Data Standards, Circular 895-E, Reston, VA: USGS.

White, M., 1979, "A Survey of the Mathematics of Maps", International Symposium on Computer Assisted Cartography (Auto-Carto IV), Falls Church: ASP & ACSM, pp. 82-96.

__________, 1983, "Tribulations of Automated Cartography and How Mathematics Helps", Proc. 6th International Symposium on Automated Cartography (Auto-Carto Six), Vol. 1, Ottawa: Steering Committee for 6th International Symposium on Automated Cartography, pp. 408-418.

White, M., and P. Griffin, 1979, "Mapping Without Coordinates", Proc. Urban and Regional Information Systems Association, pp. 431-442.

Youngmann, C., 1978, "A Linguistic Approach to Map Description", in G. Dutton, ed., Spatial Semantics: Understanding and Interacting with Map Data, First International Advanced Study Symposium on Topological Data Structures for Geographic Information Systems, Vol. 7, Cambridge, MA: Laboratory for Computer Graphics, Harvard Univ.

PUBLIC COMMENT GLOSSARY

---Preliminary Draft---

TERMS AND DEFINITIONS FOR PUBLIC COMMENT

TERM	DEFINITIONS
Accuracy	The overall error in a measurement. Accuracy includes both precision and bias, but it also includes other concerns. There cannot be a single figure to qauntify accuracy, if only because it should include the higher order moments of the probability of distribution. See also logical consistency. REFERENCE: Based on Eisenhart,1962
Attribute	Defined characteristics of a feature.
Attribute value	Defined value of an attribute.
Bias (or systematic error)	Bias is defined in terms of the difference between the mean of a measurement and the "true value." In practice the true value may only be approximated through measurements of higher accuracy or recourse to arbitrary standards.
Deep Structure	The underlying relationships among the marks or symbols on a map of Type 1 or Real Map, as well as the underlying relationships of digital data elements stored in a virtual map of type 3.

Digital Cartographic Data Model	A digital model for representation of cartographic features or cartographic processes. REFERENCE: Glossary of Terms in Computer Assisted Cartography, ICA 1980
Digital Map Registration	The spacial relationship of groups, classes or layers of digital data which can be referenced in either a relative or absolute control framework.
Entity	A geographic feature as it exists in the real world.
Feature	A defined real-world entity of interest that is not further subdivided. Identified in the real world as an entity or in the digital domain as an object.
Feature class	Defined group of related entities.
Interchange Format	A well defined data transfer philosophy which conveys such salient requirements as: Logical organization, deep structure completeness, common logical treatment of all data types, applicability to general purpose data bases and coding standards.

Object	A digital cartographic representation of a feature (entity.)
Precision	The closeness of measurements of the same phenomenon repeated under essentially the same conditions and using the same techniques. The precision of a measurement can be well described by the standard deviation. REFERENCE: Paraphrased from Eisenhart 1962.
Point	A zero-dimensional object which specifies geometric location by coordinate location.
Quality	Fitness for use.
Real Map	Any cartographic product which has a directly viewable cartographic image and has a permanent tangible reality (hard copy). There is no differentiation as to whether the real map was produced by mechanical, electronic or manual means. REFERENCE: Moellering, Harold; Real Maps, Virtual Maps, and Interactive Cartography
Repeatability	The positional accuracy of these devices includes (at least) a precision and a bias. If the bias is strictly zero, then the precision (the confidence in returning to the same spot) is the "accuracy." Often applied to describe performance of devices such as plotters and digitizers.

Resolution	Resolution is the smallest unit which can be detected. Resolution provides a limit to precision and accuracy.
Surface Structure	The symbols as viewed or touched (as in tactile) on a map (see Real Maps and Virtual Maps, Type 1.)
Virtual Map (type 1.)	A directly viewable cartographic image but only with transient reality as in a CRT map image. This is what Riffe called a temporary map. REFERENCE: Moellering, Harold; Real Maps, Virtual Maps, and Interactive Cartography.
Virtual Map (type 2.)	A product with permanent tangible reality, but cannot be directly viewed as a cartographic image. These are all hard copy media, but in all cases these products must be further processed to be made viewable. REFERENCE: Moellering, Harold; Real Maps, Virtual Maps, and Interactive Cartography
Virtual Map (type 3.)	This type map has neither of the characteristics of the earlier classes, but can be converted into a real map as readily as the other two classes of virtual maps. REFERENCE: Moellering, Harold; Real Maps, Virtual Maps, and Interactive Cartography

DIGITAL CARTOGRAPHIC DATA STANDARDS

ALTERNATIVES FOR CARTOGRAPHIC FEATURES

A PROGRESS REPORT - WORKING GROUP III

Prepared by

Warren Schmidt

MEMBERSHIP

Warren Schmidt, Chairman, Rand McNally & Company
Prof. Robert Rugg, Vice Chairman, Virginia Commonwealth University
Dr. Joel Morrison, U.S. Geological Survey
Maj. Robert Jacober, U.S. Air Force
Richard Hogan, National Ocean Survey
Dr. Beth Driver, Technology Service Corp.
Frederick Tamm-Daniels, Tennessee Valley Authority
Mary Clawson, IIT Research Institute

INTRODUCTION

As part of the National Committee for Digital Cartographic Data Standards, Working Group III - Features is committed to study cartographic feature classification systems and to specify a model that is consistent and comprehensive. The initial charge to the Working Group was to examine existing standards, specifications, and agreements and to address issues such as scale independence, organization, definitions and basic data. This has been accomplished and was previously reported (Morrison 1982, Schmidt 1983). Subsequent efforts concentrated on the alternatives and it is the purpose of this paper to document the recent deliberations of the Working Group. At the 1984 ACSM Annual Meeting, these findings will be publicly discussed.

FINDINGS

The Working Group examined many classification schemes and initially proposed the following five-part scheme:

Feature Class
Feature
Attribute Class
Attribute
Attribute Value

These classes were tested and it was found that, other than location, no common attributes applicable to all features could be identified. Therefore, it was decided to eliminate "Attribute Class". However, broad descriptive categories such as measure, serviceability, structure, and composition would be beneficial to producers of cartographic data when considering the kinds of attributes to include. The notion of "feature classes" such as culture, transportation, hydrography, etc. was also regarded as useful but would both introduce redundancy and vary among different users. Such a "feature class" need not be embedded in the definition of cartographic features, but the various classes into which a given feature might fall could be maintained as an aid for the selection and plate separation of relevant features. Based on the schema test using islands/shorelines and ports/jetties feature lists, the Working Group simplified its original rubric to include only three categories: feature, attribute, and attribute value. These were defined as follows:

Feature - a defined entity of interest that is not further subdivided

Attribute - a defined characteristic of a feature

Attribute Value - a defined value of an attribute

The use of the word "defined" in these definitions is in response to the widespread absence of definitions in digital cartography.

The definition of feature terms is a very major undertaking. The features, attributes, and attribute values must be collected, classified, defined, and approved. The Terms and Definitions effort of the National Committee (Working Group IV) will not deal with the feature sets. Even if the initial input is limited to the basic features contained on topographic maps and hydrographic charts, it is beyond the scope of a volunteer group. However, to be able to estimate the requirements of preparing feature standards, some prototype work must be done. Accordingly, a graduate student at Ohio State University was funded by the Standards Committee to devote 15 hours weekly for two months to these definitions. In addition, eight months funding has been requested for 20 manhours weekly beginning in April. The students will collect the definitions for features, attributes, and attribute values identified by the Working Group. These definitions and proposed final ones will be entered into the

Cartnet computer communication system. A sample of the initial results appears in the Appendix. The eventual operation of the Features Definition process is discussed in Alternatives.

Forty-two published articles on the topic of features were reviewed. The significant references are contained in the Bibliography at the end of this paper.

ALTERNATIVES

The second phase of the Working Group's efforts has been concerned with identifying and discussing the alternatives for cartographic features. Thirteen issues have been examined and alternatives developed. Note that nine of these issues contain the Working Group's recommendations. These choices reflect the cartographer's emphasis on solutions rather than any construed bias.

1. Scale Independence versus Scale Specific

Should feature lists be limited to certain scales or contain features from all scales? The Canadian Council on Surveying and Mapping (1982) felt that scale independence was both feasible and desirable. The Working Group concurred, finding that feature classification is an attempt to describe the real world, a place where features are independent of graphic scale and cartographic representation. The alternative, a scale-specific list, may be attractive to those producers of standardized map products but the open-ended, universal approach will serve all users.

2. Data Organization - Hierarchal or Relational?

The original issue concerning data organization questioned whether feature organization should be hierarchal or relational and just how should the data be stored for efficient retrieval. Following lengthy discussion, it was decided that because data is not retrieved in the manner which it is stored, data organization in the data processing sense was not relevant to features. What is needed for features is a logical coding scheme not necessarily tied to any existing formal hierarchal or relational model. The design should be open-ended, flexible, and provide for the entry of features and associated attributes. The system, once created, will then be accommodated by the organization of the cartographic data base.

3. Basic Feature Set - Selected or Universal

Originally one issue dealt with basic cartographic data, but the Working Group substituted "features" because the meaning of the word "data" was too broad. Is there a basic set of cartographic features--a foundation for our future classification scheme? Implied is a uniform and universal series, a perfect data set that obviously doesn't exist. However, most maps were originally derived from topographic maps or hydrographic charts. Why not start with those two map types as sources for our basic cartographic features? The features shown on those series are well documented, have stood the test of time, and apply to most scales. This "basic" information, however, would only be a start and additional features would be entered from sources at different scales and showing themes such as soils, climate, and population. In the case of thematic maps, the individual features should be classified according the theme discipline if standards exist and are contemporary.

4. Feature - Attribute Relationships

Should features be separated from attributes? In its review of the issue of feature definitions, the Working Group came to the conclusion that each feature group should be a single class and explicitly defined. If more than one definition for a feature exists, such as that for "shoreline" in the Canadian Standards, the difference should be captured in the attributes. Thematic features, e.g., aeronautical, geologic, land use, should be classified according to standards for the theme displayed. This will make for compatibility and encourage adoption of the eventual system. Attributes should be multiple and appropriately describe feature characteristics.

5. Feature - Identifier Relationships

The Working Group determined that the name of a feature is an attribute. For example, "New York" is an attribute of a feature "city." When pushed to a logical ending, this issue caused the group to conclude that relationships between features could, in some instances, also be considered as attributes. The preference of most members was to keep the notion of relations between features separate from feature and attribute definitions as much as possible. The relational aspect is further discussed in Alternative 10.

6. "Island" vs "Shoreline"

Given that an island can be defined either as a feature in its own right, or indirectly, as the area within the feature "shoreline," the group discussed this problem as one aspect of feature definition. The conclusion was that both the island and the shoreline are features. Other linear phenomena, like boundary lines, were also considered and it was decided that features need not be tangible solid objects to be defined as features.

7. "Standard Product" versus "Shopping List"

Two alternative notions of the purpose of arriving at standard feature definitions emerged. The "standard product" idea involves the definition of a minimum set of features that must be included in all "standard" cartographic products. The "shopping list" approach involves an open-ended list of features that may or may not be included in a given product. If included, the list would adhere to standard definitions. The Working Group rejected the idea that its purpose should be to develop a "standard product" for use of standards. Rather the goal was thought to be a potentially universal list of features that would be defined in the same way by various producers. Although the list of features would be open-ended and potentially all-inclusive, a beginning could be made with the features routinely included in USGS and NOS products.

8. Minumum Attributes

The Working Group discussed the possible identification of a minimum set of attributes required for a given feature. For example, should the width and surface material of a road always be specified in any data set to be exchanged under the National Standard? This goes beyond the "truth-in-labeling" approach previously enunciated, but does present an alternative to be considered.

9. Completeness

The issue of "completeness" of a feature set, raised by the Chairman of Working Group II, was considered. Although this appears to be an issue of data set quality, the group recognized that collection criteria might be included as a part of feature definition. This would provide the basis for testing "completeness" as a measure of data set quality.

10. Collocation of Features

A river which also serves as a national, state, and county boundary is an example of a collocated or redundant feature. This can be handled either by flagging in the feature file or provision for in the data structure. The former method is complicated but may be useful in certain applications. The Working Group, however, felt that data structure was the preferable approach.

11. "Pure Attributes"

Another conceptual problem discussed was "pure attributes," such as bare earth, forest cover, or gravity. Such attributes are unrelated to any particular feature. They can be viewed as features for which location is an attribute; or alternatively, locations can be viewed as features for which gravity, etc. are attributes. There was no resolution of this issue.

12. Interface - Uni- or Bi-Directional

Will the feature lists be in exchange format only or transferable in both directions. Because of the proposed detail and universality of the standard feature and attribute lists, the thinking of the Working Group is that the correspondence with other coding schemes would be in one direction from non-standard to standard but not the reverse. However, very large cadastral files might prove the exception.

13. Maintenance

The creation and monitoring of the feature definitions mechanism is not within the realm of the Working Group's charter. What organization should maintain the features list once it is adopted as a standard? Should it be the Bureau of Standards, the U.S. Geological Survey, the Defense Mapping Agency, the American Congress on Surveying and Mapping, or some other body? We believe this list should be formally maintained and that a Federal agency or board be given the responsibility. The U.S. Board on Geographic Names is a good model and, possibly, even might be a vehicle for this work.

FUTURE WORK

With two exceptions, Working Group III personnel live in or near the Washington, D.C. Metropolitan Area. Given this proximity and an unusually high degree of dedication, the members have been able to meet bi-monthly and press on with the work at hand. As a consequence the Features Working Group is approaching the end of its original goals, somewhat ahead of the other Working Groups. The work remaining will be to examine coding, evaluate the trial definitions process, and to make the final recommendations.

SUMMARY

A volunteer Working Group III, representative of government, industry, and academe was organized to deal with the issues of cartographic features and to develop alternatives. In the second phase, it has concerned itself with developing alternatives, defining and creating a prototype, and searching the literature. Tasks remaining are the evaluation of the prototype, a look at coding, and final recommendations. Public evaluation of the Working Group's efforts is needed at this point. Such comments are welcome and will be seriously considered.

APPENDIX - First page of feature definitions file

Feature Terms and Definitions

This segment contains terms for features along with the definitions for those terms. A numbered list is included which contains the sources for the definitions.

Sources

1. Houghton Mifflin, 1982, The American Heritage Dictionary, 2nd College Edition, Boston: Houghton Mifflin.

2. Canada, Energy, Mines, and Resources, Topographic Survey Division, Surveys and Mapping Branch, 1982, Canadian Council on Surveying and Mapping, National Standards for the Exchange of Digital Topographic Data, 1 - Standards for the Classification of Topographic Features, Appendix C, Dictionary of Topographic Terms, Ottawa: Energy, Mines and Resources Canada, Earth Sciences, Surveys and Mapping Branch.

3. USGS, date?, Geographic Names Information System, Appendix B-Feature Class Definitions.

Definitions

airfield
   2. Landing facility for aircraft, usually without a passenger terminal. The services offered for aircraft supply and maintenance are substantially less than those of an airport.

airport
   2. Landing facility for aircraft, usually with more than one runway and with facilities for handling passengers and air freight and for servicing aircraft.
   3. Manmade facility maintained for the use of aircraft (airfield, airstrip, landing field, landing strip).

airstrip
   2. Landing facility for aircraft consisting of a single runway which is usually of gravel construction. Airstrips rarely have a boundary fence or a delineated legal limit.

alley
   2. A narrow lane between buildings, esp. through the middle of a city block, giving access to the rear of buildings.

anchorage
   2. An area in which vessels, seaplanes, etc. may anchor. An anchorage is usually a sheltered position in which the depth, and nature of the bottom is suitable for ships or planes to anchor.

antenna
   2. A communications structure.

aqueduct (see also canal)
   1. A pipe or channel designed to transport water from a remote source, usually by gravity.
   1. A bridgelike structure supporting a conduit or canal passing over a river or low ground.
   1. A fluid channel or passage.
   2. A conduit for carrying a large quantity of flowing water.

APPENDIX - First page of attribute definitions file

Attribute Terms & Definitions
(plus some Attribute Values)

This segment contains terms for attributes (corresponding feature name is in parentheses) along with the definitions for those terms. A numbered list is included which contains the sources for the definitions. Supplied values for attributes are also included.

Sources

1. Houghton Mifflin, 1982, The American Heritage Dictionary, 2nd College Edition, Boston: Houghton Mifflin.

Definitions

access/egress
   1. The right to enter or make use of.
      The right of going out.
   (road) values: limited, unlimited, dead end

class
   1. A division by quality, rank, or grade.
   (railroad) values: 1st, 2nd
   (road)

composition
   1. A putting together of parts or elements to form a whole.
   1. The manner in which such parts are combined or related; constitution; make-up.
   (flat) values: mud, sand, gravel, unknown
   (lake) values: fresh, salt
   (reef) values: coral, rock
   (road) values: concrete, macadam, gravel, dirt, ice

cover
   1. To serve as a covering for; occupy the surface of.
   1. To extend over.
   (building) values: covered
   (pier) values: covered
   (land) values: barren, cleared, crop, desert, farm, flooded, grass-land, meadow, pasture, rangeland, snowfield

direction
   1. The distance-independent relationship between two points that specifies the angular position of either with respect to the other; the relation-shipb by which the alignment or orientation of any position with respect to any other position is established.
   1. A position to which motion or another position is referred.
   1. A line leading to a place or point.
   1. The line or course along which a person or thisg moves.
   1. The statement, in degrees, of the angle measured between due north and a given line or courseo on a compass.
   (road)

elevation
   1. The height to which something is elevated above a point of reference, such as the ground.
   1. Altitude.
   (rock)

function

BIBLIOGRAPHY

Adame, T., and D. Rhind, 1981, "The Projected Characteristics of a National Topographic Digital Databank", The Geographical Journal, 147(1), pp. 38-53.

Canada, Energy, Mines and Resources, 1982, Standards for the Classification of Topographic Features, Canadian Council on Surveying and Mapping, National Standards for the Exchange of Digital Topographic Data, Technical Committee No. I, Draft, Ottawa: Topographical Survey Division, Surveys and Mapping Branch.

Canada, Statistics, 1981, Standard Geographical Classification, Vol. 1, Ottawa: Statistics Canada, Standards Division.

Fegeas, R., 1982, "Modularization of Digital Cartographic Data Capture", International Symposium on Computer-Assisted Cartography (Auto-Carto V), Falls Church: ASP & ACSM, pp. 297-306.

Hagan, P., 1981, "The Evaluation of a Network Data Structure for Cartographic Features", Proc. ACSM Fall Meetings, San Francisco & Honolulu, Falls Church: ACSM, pp. 16-27.

Jacober, R., 1979a, Standardization of Map Symbology for Large Scale Maps, Master's Thesis, Dept. of Geodetic Science, Ohio State Univ.

Kleckner, R., 1982, "Classification Systems for Natural Resource Management", Proc. Pecora VII Symposium, Falls Church: ASP, pp. 65-70.

Morrison, J., 1982, "Report on the Activities of Working Group III, Cartographic Features", in H. Moellering, ed., Selected Papers from the Session on Digital Cartography - Planning and Standards at the Fifth International Symposium on Computer-Assisted Cartography, Crystal City, Virginia, August 23, 1982, Issues in Digital Cartographic Data Standards, Report No. 2, Columbus, Ohio: National Committee for Digital Cartographic Data Standards, pp. 13-14.

Payne, R., 1982, "Geographic Names Information System: An Automated Procedure of Data Verification", International Symposium on Computer-Assisted Cartography (Auto-Carto V), Falls Church: ASP $ ACSM, pp. 575-580.

Reneman, F., 1975, "A System for Digital Mapping", Commission III, Automation in Cartography, Enschede, The Netherlands: International Cartographic Association, pp. 254-269.

Schlichtmann, H., 1979, "Codes in Map Communication", The Canadian Cartographer, 16(1), pp. 81-97.

Schmidt, W., 1983, "Report of Working Group III: Cartographic Features", in H. Moellering, ed., Digital Cartographic Data Standards: Defining the Issues, Completion of the First Year of Work, Issues in Digital Cartographic Data Standards, Report No. 3, Columbus, Ohio: National Committee for Digital Cartographic Data Standards, pp. 33-36.

____________, 1983, "The Study of Digital Cartographic Data Features In Preparation For a National Standard", Sixth International Symposium on Automated Cartography (Auto-Carto Six), abstract p. 427, Ottawa. Also to be published in The Canadian Surveyor.

Standards Association of Australia, 1981, Interchange of Feature Coded Digital Mapping Data, Australian Standard 2482-1981, North Sydney, N.S.W.: Standards Association of Australia.

Tom, H., 1981, "Geographic Codes: Federal Information Processing Standards", Proc. ACSM Fall Meetings, San Francisco & Honolulu, Falls Church: ACSM, pp. 10-15.

U.S. Department of Commerce, National Bureau of Standards, 1970, States and Outlying Areas of the United States, Federal General Data Standard Representations and Codes, Federal Information Processing Standards Publication 5-1, Springfield, VA: National Technical Information Service.

____________, 1974, Standard Metropolitan Statistical Areas, Federal General Data Standard Representations and Codes, Federal Information Processing Standards Publication 8-4, Washington: Government Printing Office.

____________, 1979, Counties and County Equivalents of the States of the United States and the District of Columbia, Federal General Data Standard Representations and Codes, Federal Information Processing Standards Publication 6-3, Springfield, VA: National Technical Information Service.

U.S. Department of The Interior, U.S. Geological Survey, 1983, USGS Digital Cartographic Data Standards, USGS Circulars 895-A: Overview and USGS Activities; 895-B*: Digital Elevation Models; 895-C*: Digital Line Graphs from 1:24,000-Scale Maps; 895-D: Digital Line Graphs from 1:2,000,000-Scale Maps; 895-E: Land Use and Land Cover Digital Data; 895-F: Geographic Names Information System; 895-G*: Digital Line Graph Attribute Coding Standards, Reston, VA. (*draft, unpublished at submission time)

Werner, P., and Bruck, H., 1975, "National Geocoding Systems: Design of A Converter", Urban Systems Laboratory, MIT, Cambridge, Mass.

Yoeli, P., 1978, "The Drawing of Topographical Maps with Basic Cartographic Subroutines", in J. Moneta, ed., Information Technology, Amsterdam: North Holland, pp. 697-703.

ALTERNATIVES FOR SPECIFYING QUALITY STANDARDS FOR DIGITAL CARTOGRAPHIC DATA

Progress Report for Cycle 2
Working Group II: NCDCDS

Prepared by N. Chrisman

Working Group II on Data Set Quality is composed of:

Nicholas Chrisman (chair)	University of Wisconsin
Charles Poeppelmeier (vice-chair)	Defense Mapping Agency
Frederick Broome	Bureau of the Census
Wallace Crisco	Bureau of Land Management
John Davis	Kansas Geological Survey
George Johnson	National Ocean Service
Dean Merchant	Ohio State University
George Rosenfield	U.S. Geological Survey
John Stout	Petroleum Information Inc.

BACKGROUND AND PURPOSE

Working Group II on Data Set Quality has the mission to develop standards for describing and testing the quality components of digital cartographic data standards. The working group takes a broad view of quality, encompassing completeness, logical consistency, and lineage along with accuracies of position and attributes. The goal is a comprehensive scheme that can serve all forms of data, so the standards will not prescribe specific numerical thresholds for any particular product. (For background on the overall mission of the NCDCDS, see Moellering, 1982; 1983; Chrisman and Moellering, 1983.) This document reports on the completion of Phase 2 of our deliberations: defining the alternatives of the problem. Phase 2 was based on the previous phase: defining the issues, reported earlier (Chrisman, 1983). The next phase involves drafting an interim standard, so comments on this report are requested from all interested parties. A public discussion will occur at the ACSM Annual Meeting in March 1984.

Definitions of quality

> "Quality has been variously defined as 'fitness for use', 'meeting an expectation', 'degree of excellence', and 'conformance to a standard', along with other phrases. These all have merit, depending on one's point of view." (Hayes and Romig, 1977,p. 9).

The definition of the quality component of a digital cartographic data standard provides the first "alternative" to examine. The first definition in the quotation above, fitness for use, has received the most attention from the working group. We find "quality" to be a wide-ranging concern which can cover any issue affecting the use of cartographic data. A digital data base transmits data from a producer to a consumer. Other working groups are concerned with defining mechanisms to aid interchange, and the quality component fits into this general purpose.

Fitness for use is one useful definition for quality, but it is not comprehensive. Quality information forms the basis of quality control and thus should serve the producer. Information compiled to assess fitness for use also answers a producer's concerns. Furthermore, the demands of maintaining a data base over the long-term probably require more detailed quality information than a user would need. In addition to the producer and consumer, an intermediary organization often acts as "supplier" (or wholesaler). The role of a supplier who is not a producer depends heavily on quality information to understand the product.

In order to create viable, effective quality standards, a combination of user's and producer's perspectives must be maintained. The self-interest of a producer will lead to attention to quality issues, because data exchange standards will not coerce any internal changes. Considering this analysis of producers, the working group will concentrate on communicating "fitness for use" to a consumer.

In the previous report (Chrisman, 1983), we moved quickly from "fitness for use" to our general attitude towards a standard. The potential uses of digital cartographic data are so diverse that a fixed set of numerical thresholds could not adjust to the potential uses. In more circumscribed application areas (for example, a multipurpose cadastre or a forest inventory), a set of thresholds might be fruitful. Because these standards must serve the whole profession, we forsee a "truth in labelling" standard instead. The idea is to communicate actual numerical properties of the data in a way that potential users can make their own informed decisions on fitness.

The truth in labelling concept may seem less rigorous in that it blesses the status quo. Any imprecise, inaccurate data base could meet the standard in the formal sense by proclaiming those imprecisions and inaccuracies. While this is possible, we suspect that the publication of quality information will promote greater attention to quality. A truth in labelling standard will push producers whose product is not fully inspected towards greater quality control and higher standards, lest their product be judged inferior.

Relationship to quality control

A well-developed literature exists in the area of quality control, mostly oriented towards industrial and engineering applications. The working group has found a useful basis in works such as Eisenhart (1963). The quality control literature emphasizes the identity of the user's and the conscientious producer's attitude towards quality. The practice of industrial quality control has been codified into a set of standards for using control chart methods (ANSI Z1.1 (and .2) - 1958 R1975). From published reports, few cartographic producers seem to use this type of management tool to monitor and ensure quality.

The working group could recommend the adoption of this formal method of quality control in a producer's operations. This kind of intrusion could be counterproductive, although we encourage consideration of the control chart technique. The role of this committee is to promote standards for the interchange of products, not to reform the practice of cartography. Standards suggested by this working group could be used to develop the measures of accuracy needed to carry out a control chart approach.

Organization of this report

In the previous report, we defined a set of "issues" that contribute to data quality. The most important components (lineage, spatial accuracy, attribute accuracy, logical consistency and completeness) form separate sections below. In each section a set of alternatives are discussed which could contribute to a standard. Before entering these discussions there are a number of topics which are more general and apply to the whole process. These topics are grouped under headings of levels of testing, generic types of data and specificity.

LEVELS OF TESTING

Achievement of quality results requires some form of quality assurance. In some cases, assurance can be provided by standard specifications and procedures, but this working group is convinced that a set of standardized tests would provide much greater confidence in the quality estimates. To reiterate some recent titles: a specification is not a standard, particularly in a fast changing field such as digital cartography. Technology is changing too rapidly to permit rigid specifications of procedures.

At our earliest meeting (see Report 2, Merchant, 1982b), the working group recognized the need for a quality standard which separated different categories of testing. The categories we defined fall roughly along a continuum of rigor. The lowest level of "testing" originally included the lineage report (see below), which provides a narrative of the origins of the data. In recent deliberations lineage has taken an important, but essentially distinct role. Each of the four categories of testing (deduction, internal evidence, comparison to source, and independent source of higher accuracy) should be considered as alternatives. In addition, some of these categories raise more detailed alternatives.

Deduction

In the draft standards for Canadian topographic data (Canadian Council on Surveying and Mapping, 1982), quality estimates are not presumed to come directly from sources of higher accuracy. Each instrument or procedure is expected to have a separate calibration study. Then a quality statement for a particular product is made by propagating the separate error effects. This approach has also been applied to the assessment of the error inherent in distributed digital data bases (Chrisman, 1982a). These approaches use deductive logic to generate an estimate of quality from calibration studies of higher rigor. To a large extent, an approach of standard specifications and procedures is a form of deductive logic, but it does not even provide a numeric estimate.

Extrapolation from calibration tests (or deductions of possible error magnitudes) to their combination in a particular application can be realistic or misleading. There is no particular way to tell. However, some form of deductive logic is required for any test procedure which is not exhaustive. It should be possible to conduct detailed tests for limited areas (or for selected sheets) and use deductive logic to extrapolate to the rest of the data. Well designed tests are needed to obtain the best results. Considering the potential cost of exhaustive testing, the deductive approach may have its place.

Internal evidence

A test based on internal evidence requires some form of redundancy in the data. Sophisticated parity checking schemes for data transmission can ensure that errors can be detected and corrected from internal evidence, but this applies simply to strings of bits. For an example closer to mapping, some standards for surveying specify closure tolerances. Constructing a survey to provide closure requires extra effort, but it provides a basis for adjustments and estimates of positional accuracy. Similarly, a topological data structure includes independent encoding of relationships. This redundancy permits automatic checking of consistency and completeness.

The specific form of internal evidence will vary from situation to situation. Simple data structures have no redundant structure to test. For some aspects of the data, such as positional accuracy, internal evidence is a weak test compared to the next levels of rigor. This has been emphasized in recent papers on surveying standards (eg. Crossfield, 1980). However, a weaker test should not be completely disregarded. If it is the only test available it provides some indication of the fitness for use. It also indicates that testing was not carried out at the highest levels of rigor. For other aspects, particularly issues of data structure and logical consistency, there is no more rigorous test.

Comparison to source

While the preceding tests use internal evidence, the next more rigorous tests compare the product to the source material. These tests examine the fidelity of the processes performed, but cannot discover faults in the source information. A chain of comparisons back to the raw observations could eventually describe the overall precision, but deduction would be required to combine the error estimates. Tests by comparison to source material, such as "check plots" for digitizing, are commonly performed, but the results are rarely formalized and reported to a user. It will be a challenge to develop these standards.

An independent source of higher accuracy

To obtain a comprehensive test of overall quality, it is necessary to use an independent source of higher accuracy. In some situations, more accurate (and expensive) methods can be applied. Under ideal circumstances, the more accurate method measures identical entities and can be applied in a well-designed statistical sample. The measures are then used as the "true value", because their accuracy is so much higher than the values to be tested.

Some of the assumptions can be tested themselves, but some rest on faith. A more expensive, higher technology approach is not necessarily more reliable. One way to establish higher accuracy is through tests of internal evidence. (A reader may notice a circularity here.) As an example, the Federal Geodetic Control Standards rely upon standard procedures and relative errors of closure (a form of internal evidence), because there cannot be a source more accurate.

In actual applications there should be some solution to these questions. The strict assumptions may have to be relaxed somewhat. A standard based on independent sources of higher accuracy often establishes the ultimate standard by essentially arbitrary definition. International standards such as the meter are now more scientific than the length of a metal bar in Sevres, but they are no less arbitrary. Measurement standards for positions abound, but other aspects of cartographic data do not have a recognized standard of higher accuracy. Although a standard may be lacking, any information that has an objective basis is best tested by an independent derivation. This working group cannot promulgate all the possible standards for these sources, but it can lay a framework.

GENERIC TYPES OF DATA

In discussions involving all the working groups a list of generic types of data emerged. The list was mostly designed to capture the overall categories for defining logical consistency (see below). However, this list has applications elsewhere. Decisions about data structures reflect fundamental choices on ontology - what features exist and their relationships. While the list developed handles data structure issues, some refinement may be needed to handle all quality questions.

List of generic data types

- Point
- Grid / Raster
- Elevation (with subtypes: matrix, TIN)
  - Vector data types
- Cartographic Spaghetti (line data stored as "strings")
- Polygon (self-closing loops)
- Topological (with subtypes coverage, network)

SPECIFICITY

A final concern (voiced in the last report) was the specificity required for a quality report. Quality information could be taken to extreme limits where it could equal or exceed the rest of the data base. This issue leads to a set of alternatives.

Quality reports and tests could be performed on a very broad scale by occassional examination of isolated parts of a whole map series. The tests done on this basis are deduced to apply to the untested data. This approach to quality control may best describe the current approach of major federal mapping establishments. Does it provide enough information to address the user's questions on fitness for use applying to a specific area?

As an intermediary, some tests could be applied to each product (say map sheet). What would be "enough" to provide the right assurance about the quality information? In positional accuracy standards, the ASP spatial accuracy standards suggest twenty sample points, properly distributed, as a minimum (Merchant, 1982a). Statistical arguments could be made for thirty, forty, et cetera. What would be done if only a dozen high accuracy test values can be obtained? Some tests of classification accuracy have the same dilemma. It may be feasible, even necessary, to perform some less expensive tests exhaustively, particularly internal evidence tests for logical consistency. Some exhaustive tests need not be overly expensive to perform or to report, but there are extreme examples. The profession would be unlikely to accept a standard requiring individual accuracy ellipses stored with each coordinate.

The alternatives for specificity require complex tradeoffs.

LINEAGE

The basis of any quality report is a narrative of the lineage of the data. Lineage includes the original source material and all the processes and transformations leading to the final product. This information is required for a user to evaluate fitness, and it is required by a producer to maintain and update the data. With the development of digital cartography, the lineage information becomes increasingly important, but there are virtually no reports of software to maintain this information.

Lineage in the traditional process

The needs of digital cartography can be clarified by examination of lineage information currently maintained by producers for the traditional process. Currently the state of lineage information is highly variable. In a few limited cases, such as the National Ocean Service (NOS) shoreline maps, detailed lineage information is generated and published. For example, the NOS shoreline map TP-00466 (covering a five minute quadrangle at 1:10000) is covered by a Descriptive Report of 26 pages with attachments. The descriptive report includes a number of forms standardized by the producer to report the various phases of producing the map: an overall data record, a summary, a field report, a photogrammetric plot report with references to control, a compilation report, a field edit report, a review report. The first six pages of this report are reproduced as Appendix 1, below. These pages provide an alternative level of detail which lays out the basic raw materials used. The rest of the NOS report includes additional detail, down to the errors of closure for each control point used, for example. This further level of detail offers a very clear description of fitness for use. The working group commends this particular producer for their attention to lineage information.

This example provides some alternatives, but this level of effort may not be justified for all products. The actual context of a full lineage report would be quite product- and technology-specific, and hence beyond the mandate of this working group. The NOS information is probably as comprehensive as any records currently maintained by a producer. Other examples would include Census political boundaries and local property maps when they are based and linked to official plats, registered surveys and title information. Such records are not simple to maintain. Detailed lineage seems to accompany maps where each detail may be scrutinized by someone like a property owner who is directly affected. These map users are not interested in average errors.

In the more usual case, lineage consists of a few pithy statements printed on the map border. In many cases of maps produced by local governments, the lineage information is perpetuated only in the minds of a few public officials. A simpler alternative is to automate the notes found on the borders of topographic maps. A few examples are included in Table 1.

Table 1: Example of annotations on a USGS 1:24000 map

Belleville, Wis. N4245 - W8390/7.5
1962
PHOTOREVISED 1971
PHOTOINSPECTED 1981

THIS MAP COMPLIES WITH NATIONAL MAP ACCURACY STANDARDS
Mapped, edited, and published by the Geological Survey
in cooperation with State of Wisconsin agencies
Control by USGS and USC&GS
Topography by photogrammetric methods from aerial
photographs taken 1961. Field checked 1962
Polyconic projection. 1927 North American datum
10,000-foot grid based on Wisconsin coordinate system, south zone
1000-meter Universal Transverse Mercator grid ticks, zone 16 shown in blue
Fine red dashed lines indicate selected fence and field lines where
generally visible on aerial photographs. This information is unchecked
Revisions shown in purple compiled from aerial photographs
taken 1971. This information is not field checked
Map photoinspected 1981
No major culture or drainage changes observed

To place on the predicted North American Datum 1983
move the projection lines 4 meters north and
10 meters east as shown by dashed corner ticks

(On a nearby sheet:)
Red tint indicates area in which only landmark buildings are shown
..
Purple tint indicates extension of urban areas

An additional example, found in Appendix 2, is the record of cartographic work maintained by NOS for a nautical chart. This example involves much more detail than the topographic example, partially because updates are performed in a more piecemeal manner for this type of map. Both sets of notes are designed for a traditional sheet-based production system. Future data bases may be maintained on a sheetless basis over larger areas. Even where old sheets are perpetuated, updates will not be as tortuous as they used to be. Frequent update will lead to increased size of lineage records. Notice also that the marginal notes include cautions which deal with issues of attribute accuracy and completeness of certain features. To assist digital use, these records should be made more uniform and predictably available.

In some existing map series, a "reliability diagram" (not a map) is included on the map margin. This diagram shows lineage information in its spatial distribution, but it cannot be registered as a map overlay to use the information more directly. A modern revamping of the reliability diagram is a challenging alternative.

To complement the marginal notes, a producing agency often produces a guide to the compilation of the map series. These documents are largely used internally and receive little attention beyond the producing agency. However, these manuals are the real key to the map series, setting forth rules such as minimum sizes, interpretation keys, and other standards. There is a need to transfer this information to the users so that they can understand the utility and limitations of the series. Recently the specifications for compilation of the USGS Land Use/ Land Cover series moved from an Open File Report of little circulation (Loelkes, 1977) to be incorporated with the other discussion of the program in Circular 895-E (Fegeas and others, 1983).

Lineage in a digital age

Some of the new municipal digital data bases include a recognition of the need for quality records. For example, Hanigan (1983) describes a reliability code to describe the quality of each original map sheet. In this data base one byte is allocated for this quality information per map sheet. One byte will not strain the digital system, but it will soon be entirely inadequate. As a tax parcel data base, there will be continual activity modifying parcels as roads get built and as parcels are subdivided. The single byte will not be able to contain a reasonable record of this fragmented lineage.

A major challenge will come from the advantages of an automated system. Digital data bases are very easily modified, particularly with current graphic work stations. Yet our professional attitudes are built around the very different situations of traditional map production where updates and modifications had to be stockpiled for a long time before a revised edition could be justified. Recently computer scientists have paid attention to the problem of "versioning" or how to record the modifications and different states of a data base. Still, a full lineage record can not be efficiently generated from a digital data base's operation, but this development should be encouraged. Until that time, there should still be some type of audit trail to trace the history of a data base. The auditor may often be the producer's own staff performing quality control, but the audit trail provides valuable information to a user on currency and other aspects.

In an ultimate extension of the argument, quality records including lineage and other accuracy estimates have the potential of becoming as detailed as the smallest units of the data base. Certainly map sheet generalizations will not be sufficient for all users. A revival and extension of the reliability diagram is an alternative to complete disaggregation. A lineage overlay, registered as an overlay in a standard multi-layer data base, could describe zones sharing common sources, dates and standards. These attributes could be stored as other thematic data are. Map producing software could be used to examine quality or analytical software could include quality concerns in other operations. As updates and modifications occur, the reliability overlay must be modified.

POSITIONAL ACCURACY

This section of the report concerns the alternatives available for measurements of location (position). All geographic information includes some form of positional reference. This section will be mostly concerned with the planimetric (2-dimensional) situation, although it will also consider topographic accuracies as well. Since many standards exist for these measurements, the first portion of this section reviews some of these standards as alternatives in their own right. Returning to the levels of testing, a further section reviews the needs for each level of rigor and the relationships to existing standards. A final section deals with other aspects of positional accuracy mostly related to transformed coordinates.

Existing standards

Standards for positional accuracy have been developed by a number of groups. This working group is resolved not to add another layer of potential confusion. In so far as possible, we intend to use compatible or identical definitions and standards. The trouble is that existing standards have incompatible overlaps and they leave some cartographic problems uncovered. There is a need to interpret these standards and make them more compatible.

FGCC. The Federal Geodetic Control Committee (1974) has established a set of standards for geodetic control. This standard is a conventional one that creates classes of control based on following the specified field survey procedures and on thresholds of a test result. As mentioned earlier, the FGCC test uses closure of traverse, a measure of internal evidence. The users of geodetic control are relatively homogeneous, and some diversity is handled with the notion of separate classes of precision. This standard is well-accepted in the United States and is consistent with similar standards elsewhere.

This working group has two recommendations relative to the FGCC standard:

> Quality of control for all digital cartographic data bases should be expressed in terms of the FGCC classification. The class of each control point used to construct a product should be noted in the lineage report. (In addition, the class should be an attribute for the feature "control point", "benchmark", "triangulation corner", etc.) If the control comes from the National Geodetic Network, the identity of the points in that data base is needed. When a separate control survey has been used, that survey should be reported in the form used for the national network, even if the results fail to achieve the FGCC thresholds.
>
> Changing technology for geodetic surveying may alter many constraints on control. The standard may be redefined with more emphasis on absolute standards and less on relative errors. These changes should be the responsibility of the FGCC and standards for digital cartographic data should adjust to the current

geodetic system.

Along with the standards for geodetic control, standard reference ellipsoids define horizontal datum(s) and standard geoids define vertical datum(s). This working group, with its charge to promote interchange, strongly supports the use of these standards. This statement would seem obvious if a horizontal datum were enshrined forever. The current standard is dated 1929 and is in the process of readjustment. New technology, as well as revised procedures for reduction of existing observations, will make substantial improvements in the geodetic network. The new system will be dated 1983, although it will not be available for some time.

The conversion of the horizontal datum will alter the definition of state plane coordinates and other projections used in digital cartography. Some argue that the new datum creates unnecessary trouble. Relabeling all coordinate-dependent features would be costly with a manual cartographic product. Digital technology should reduce the problem. Our working group supports the use of the new datum and the continual conversion of digital data bases to reflect improved control. Software to support these needs must be developed.

Implied standards. Acceptance of the National Geodetic Network as the basis for digital cartography implies a few more basic standards. Ultimately the geodetic standard implies the use of latitude and longitude with the quasi-arbitrary origin in the Bight of Benin (Greenwich and Equator).

> As a standard, all coordinates used for the transfer of digital cartographic data should bear a known (and expressed) mathematical relationship to latitude and longitude.

This basic standard permits such common projections such as UTM and the state plane systems, along with many others. Some future committee may want to establish prefered projections and units of distance, but we do not find this worthwhile at this juncture.

NMAS. The National Map Accuracy Standards (NMAS) (Bureau of the Budget, 1947) are accepted in principle by the major federal map producers. These standards set thresholds of acceptable error in horizontal and vertical measures linked to graphic scale. There has been continuing debate over the formulation that 90% of tested points should lie within the tolerance. On first view, the tolerance seems straight-forward, but this specification does not directly convert into a more accepted statistical treatment of errors. NMAS also fails to specify standards for testing.

In addition, the NMAS is strongly tied to the graphic scale. The fixed tolerances translate into varying ground distances, which may be realistic for traditional map products. However, a digital data base has certain scale-free characteristics, because its coordinates are usually in absolute form. The resolution of digital systems usually exceeds accuracy by at least an order of magnitude, whereas the graphic resolution of a traditional product may only be a factor of two different from the NMAS threshold. One alternative is to rely on NMAS for all products derived from a given product, which creates additional needs for lineage information to transmit scale dependencies. In new production, based on direct digitizing from stereomodels, a sheetless data base is a current reality.

> The NMAS is obsolete in a number of respects. It is not mathematically and statistically sound. It is scale dependent in a less scale dependent digital age. It fosters the use of a simple threshold and a very simple report (THIS MAP COMPLIES...) where we believe more information should be generated and presented to the user.

Koppe's formula. In a number of European countries, the accuracy of topographic maps is described using a formula derived by Koppe at the beginning of this century. This formula was mathematically relatively advanced for its time, as it splits errors into two parameters controlling planimetric and slope-related errors. Koppe's formula can be applied as a structure for a deductive statement, in which case it is not very different from NMAS. Koppe's parameters can also be derived from fitting the test results derived from an independent source of higher accuracy. The working group wants to encourage testing, but it is not clear that Koppe's formula provides the best means to report test results. The parameters relate to a producer's problems, but do not directly fit a user's needs.

Recent standards efforts. Alternative standards for map accuracy have been developed by the ASCE and the ASP (Merchant, 1982a). This working group does not plan to promulgate another variation on this theme. The recent work stresses reporting error in ground units for users understanding and the statistical treatment of test results. The test procedures and their numerical treatment are more germane to digital cartographic data standards than a particular classification scheme. For example, the suggested ASP standard tests coordinate errors in two or three dimensions using statistical tests of bias and precision.

Levels of testing for positional accuracy

Deduction. The use of deduction for positional accuracy statements is not mentioned in the standards above, but it is developed in the Canadian draft standards (Canadian Council of Surveying and Mapping, 1982). If all of the procedures used in production have been carefully calibrated, it would be possible to apply error propagation approaches to develop an estimate of overall error. In practice, this method has many assumptions which must be met.

> First, there is a need for rigorous calibration tests for all equipment. For example, in spite of the large number of digitizers currently in use and the competition of scanning technologies, there is scant literature on the errors expectable from digitizing hardware and its normal use.
>
> Second, the nature of error propagation between procedures is not well known. The Canadian approach assumes uncorrelated errors, but statistical correlations between some tasks (drafting and line following, for example) are quite plausible.

To implement a deductive approach combining calibration-derived errors, it would be necessary to perform a test of the whole system to discover possible correlated errors or biases.

Another type of deductive logic is the extrapolation from one intensive test to cases sharing similar conditions. The process of testing cannot apply exhaustively to each individual measurement, of course. Alternatives to derive workable sampling strategies will be discussed below, but they all leave room for uncertainty.

Internal evidence. Some forms of testing can rely on either repeated measurements or other forms of redundancy, such as closure of traverses. Standards such as the FGCC provide an adequate basis for these tests.

Comparison to source. In many cartographic procedures, information is graphically transcribed or traced with no explicit standards governing the fidelity to an original. Digitizing contracts require a means to specify the closeness of digitizing results to original linework. Current contracts call for errors less than one line width, or some such, but there are no accepted procedures for testing.

Graphic inspection of two products is very difficult work. It is probably impossible to detect all errors until they exceed about two linewidths at the minimum. In principle, a comparison to source can be conducted exhaustively, but a visual scan for gaps will not produce a numerical estimate of the fidelity of the product. To produce a numerical result, a more accurate measurement is needed, entailing a test more like the next level of rigor. Further research is needed if these tests form a major part of the quality control of our digital data bases.

Independent sources of higher accuracy. The working group supports the use of judiciously sampled tests using an independent source of higher accuracy. It is important to stress each part of this term. The testing source must be independent of the source to be tested. Independence can be more or less complete, using different operators, different control points, or best, different technologies that have different error properties. The testing source must also achieve a higher accuracy. In some cases a few test points of higher accuracy can be obtained with little trouble, only higher cost. One current example is described by Petersohn and Vonderohe (1982); ground survey was used to verify digitized property maps. In other cases it may be harder to find a practical independent source of higher accuracy. A standard for positional quality should provide a guideline for selecting the independent source. The ASP draft standards consider higher accuracy to imply one third or less error (Merchant, 1982a). The standard also calls for at least twenty points for the test and certain rules of spatial distribution. These rules need to be considered as a basis for our standard.

One of the main stumbling blocks of existing accuracy tests is that they apply best to "well-defined points". A well-defined point is defined by NMAS in terms of sharp identity on the graphic product and on the ground (or source material). In a digital age, the need for definition on the graphic product can be relaxed. Many locations are measured in a digital data base which do not fit the definition of well-defined point. Some maps, such as wetlands and soils, contain very few points which can be tested. A standard should prescribe techniques for less defined points.

From one respect, evaluating positional error for well-defined points reduces the chance for confusing attribute error with a positional one. However, all positional error may not mimic what can be measured at well-defined points. Some evidence suggests that cartographic detail is less accurately captured than the few isolated well-defined points (Thompson, 1981). Furthermore, some of the well-defined points sampled may be the features used to control the compilation process (in photogrammetry or ground survey). If so, these points received special treatment and their errors would not be indicative of the rest of the map. The distribution rules of the ASP draft standards offer one solution to this problem by insisting that all check points be separated from each other and from map control by a specified distance. A detailed lineage report is required to preserve information on control points so that tests of positional accuracy can be properly designed.

Slivers. In a multi-layered data base there will inevitably be features which are intended to be the same, but which are not identically encoded. The graphic result is a set of "slivers" - narrow polygons formed by lines which may be intended to be the same. Slivers are noted in many applications and can occupy a large portion of storage in some data bases. At least partially, the existence of a sliver problem can be blamed on a lack of realistic quality standards. Sometimes, in order to preserve graphically pleasing smooth curves, inordinate positional accuracy is imputed to a data source. In other cases, slivers are caused by errors in interpretation or change of features over time.

To the extent that slivers represent independent versions of identical features, they provide a source for testing positional accuracy. If one source is clearly superior (from a more detailed scale, more accurate technology, etc.), it could provide a test in the most rigorous category. If the sources are more similar, it may be more realistic to design a test to discover precision without imputing higher accuracy to one or the other. Since slivers are considered to be a widespread problem, the raw material for these tests is plentifully available.

Coordinate transformations

Quality in positional information does not stop with accuracy in source material. The accuracy must be maintained through later stages of processing. Storage in all digital forms necessarily restricts resolution, sometimes rather dramatically. Positional error from roundoff of coordinates is unavoidable, and sometimes it can be large enough to cause trouble. For example, some metropolitan DIME files were rounded off at some stage of the process, creating wavering effects in straight streets. A more difficult problem can be created through loss of precision in transformations. If a coordinate requiring the full resolution of the storage scheme (say 6 digits) is multiplied by a factor with the purpose of rotation or some other operation, it is quite easy to lose the precision of the value, even if the operation is mathematically exact. Even worse results can occur in calculating intersections or centroids which require multiplying two coordinates together. The problem is often caused by using absolute coordinates for all operations. Programming tricks, such as use of local offsets, can diminish the problem. Information on the nature of the programming is hard to come by, particularly for proprietary software only distributed in executable form. Perhaps a vendor of software should certify that care has been taken to preserve precision of coordinates through all potential transformations. Some systems adopt double precision storage to solve this problem, but this is not always necessary.

In the course of the history of a digital data base, coordinates are transformed from one system to another. The methods used to make the changes should be documented. Careful use of lineage could reduce propagation of errors. The locations of control points for rubber sheeting should be flagged, and the nature of the algorithms should be documented. The parameters for projection equations should also be reported in the lineage report.

The discussion of lineage mentioned the "reliability diagram" as an important alternative. Many of the components of this diagram would pertain to positional accuracy. In a multi-layer data base there might be a layer which described the nature of the control for the source material and describe the relationships of dependent objects. This later step is necessary to develop data bases which can adjust themselves to upgraded positional accuracy as it is discovered. An example would be the dependence of parcel lines in the Public Land Survey System to the coordinates for the adjacent section corners.

Raster registration and rectification

Most of the preceding discussion concerns the generic types of data that use coordinate measurements directly to represent position. When using a grid data structure, position is implicit in the addressing scheme. Some terrain data (DEM) and most imagery data bases are stored in this form. Quality in position is just as crucial to these sources, although the details and the techniques will vary. The process of registration is similar to the cartographic process used to compile maps from photographic sources, except that the process is more directly digital. A registered image merely tells the correspondence between pixels and ground positions. In terms of positional accuracy and spatial properties, a registered image may contain distortions. Any information produced from such a distorted base should contain a caution, just as it would if unrectified photographs were used.

Image rectification or geometric correction transforms the raster data to become a proper map data base. In such a data base each pixel has planimetric properties, and spatial measures such as area can be reported without a caution. Geometric correction is a complex task to perform for raster data, because of the grainyness of pixels. (With coordinate data, a distortion can be done numerically without altering the data structure.) This working group takes notice of the current interest in the topic of rectification in the remote sensing community. Most of our cartographic data standards (lineage, attribute accuracy, etc.) could apply to raster sources without much difficulty, but a geometric standard should develop from the specialized interest group.

ATTRIBUTE ACCURACY

In the scheme used by Working Group III on Features, a feature can have a number of attributes, one of which is location (position). For the purposes of this working group, the positional "attribute" must be considered separately, because the quality concerns and tests are different. This section considers all the other attributes of a feature, which must logically include the identity of the feature itself. Thus attribute in this section includes "feature code", "geocode", etc.

The most basic distinction to make about attributes is normally referred to as "levels of measurement". Attributes may be measured on nominal, ordinal, interval or ratio scales. Each scale determines the appropriate mathematical operations which can be applied. For example, it is only possible to determine equivalence of two nominal measurements, whereas a ratio measure may be added and divided. Almost all useful cartographic data bases require ratio measures for position, so the issue does not apply in the previous section. The limitations of each level of measurement must be recognized in a technique to determine quality. In statistical methodology, it is common to group nominal and ordinal scales under the general title of discrete or categorical measures, and interval and ratio scales as continuous. For the purposes of this discussion, these two broad groups will be recognized, although the more detailed distinctions would be important in a specific application.

Continuous measures

Continuous scales of measurement permit the mathematical treatment often associated with quantitative analysis. Statistics for bias and precision, and related descriptions of probability density, provide an adequate description of overall accuracy. Accuracy assessment for quantities of precipitation, for example, can be performed using the same numerical methods used for positional accuracy. Certain spatial effects can make the problem more complex than an exercise in elementary statistics.

The most developed tests for continuous attributes have been designed for topography. To some extent, elevation is a measure of position on the earth's surface, but on the planimetric map, elevation is simply an attribute. Continuous attributes are not easy to portray if they vary continuously over space. Traditionally, the continuous surface of elevation is sliced into discrete contour lines for display. However, in evaluating accuracy, even of a contour line, elevation is treated as a continuous quantity.

The statistics to report error in a continuous quantity, such as elevation are not controversial, but the method of testing is not so obvious. Ley (1981) reports on five different methods to assess the accuracy of a digital elevation model. There are other alternatives in the literature. The problem relates to the issue of not-so-well-defined points raised in regards to positional accuracy. On a continuous surface, few locations will be completely unambiguous on two sources from different technologies. Ley shows the dangers of forming an accuracy assessment of topography from only the well-defined photopoints, because these are the same points used to construct the maps (so the sources are not fully independent). A statement of fitness for use should apply to the whole surface, not just to privileged locations which are easy to check. Most alternatives involve some form of interpolation.

In general terms, the methods used to test topography can be applied to any other surface, except where the method relies on specific technologies such as stereomodels which are not possible for other surfaces such as subsurface geology or rainfall. Tests based on independent sources of higher accuracy are possible and suggested. Deductive estimates should be traceable to a well-designed calibration test. Tests based on internal evidence are possible for some instruments, such as the Gestalt Photomapper, which can be programmed to overlap sampling points for adjacent patches. Tests based on comparison to source are also employed when elevation data is derived from existing topographic maps. The different nature of these different levels of rigor should be emphasized.

<u>Categorical attributes</u>

Categorical attributes are very common in digital cartography, particularly nominal attributes such as feature codes, geocodes, land use types, parcel identifiers, etc. A full examination of quality must consider how reliable this information is. Standard methods of accuracy assessment applied to positional error (such as Root Mean Square Error) cannot be applied to categorical information because of the differences in mathematical operations permitted. A nominal code can only be right or wrong, there are no gradations of closeness. Instead of borrowing the estimation statistics used to deal with continuous measurements such as coordinates, categorical measurements must be treated with statistics developed for other applications. The largest literature in the fields related to cartography is in the analysis of thematic classification accuracy in remote sensing.

In remote sensing applications it is common to test the classification results against a source of higher accuracy (often called ground truth). Usually the percentage correct is reported. This figure offers one alternative which is easily understood. Recent work has shown, however, that the percent correct is an imperfect description of classification accuracy. It fails to account for differences in the number of classes and the distribution of area amongst them. The working group supports the search for more sensitive statistical measures so long as they communicate efficiently. One such possible measure is kappa, a measure of agreement which deflates the percentage correct to account for potential "random" errors (Congalton and Mead, 1981; Chrisman, 1982b).

Another alternative is to simply report the whole misclassification matrix. This matrix shows the ground truth data arrayed in columns and the tested classification arrayed in rows. The percentage correct merely reports the diagonal of this matrix. The rest of the cells show exactly which sorts of errors seem to be most prevalent. The information would be very useful in assessing fitness for use, because certain errors might be more worrisome than others. Also, any more complex method must be based on this information.

Tolerable errors rates for categorical data are probably lower than they are for continuous data, in some respects. If feature codes are entered incorrectly, the wrong features will be extracted to make selected products. This will cause significant trouble to the user. Some errors can be checked through the checking of data structures discussed in the next section on logical consistency, but others must be tediously examined.

Sampling strategies can be devised, using the classical alternatives discussed in Berry and Baker (1968; Rosenfield and others, 1982). These include random samples which ignore spatial properties, systematic samples which can run into wavelength problems in some landscapes, and stratified systematic unaligned which combine a number of factors to avoid these problems. This standard cannot dictate the proper sample to use for all applications, but the choice should be documented. In addition, the locations of the sampling points should be reported as a component to the reliability overlay.

Purity

A major problem with cartographic attributes is the interpretation of scale. As it was traditionally handled, with graphic products, scale restricted operations performed with a map. In digital form it is easy to push information beyond its intended scale of use. When polygon maps are made they are considered homogeneous, but they often are the product of conscious generalization. If the polygon is examined too closely, the homogeneity of the polygon will not hold up; the polygon will be found to be impure. The soil map is a frequent example of maps which have purity problems, but even political maps can exclude small outliers as a function of scale. When the whole polygon is examined, the choice of attribute is defensible, but details inside may not be completely pure. It seems unreasonable to demand complete purity of all polygons, but operations must be designed to deal with this form of error. Particular attention must be paid to polygon overlay logic, but that is not a topic of this report. Similarly, an attribute test could make improper assumptions about purity.

Many approaches for attribute accuracy use point samples to test the accuracy of a classification. It would be easy to find a point in water in a delta area which has been classified as land. Does this invalidate the whole delta? A proper test of attribute accuracy should use an independent source of information, but it must use the same system of classification. In addition it should use the same scale in its interpretation.

A recent paper by Cook (1983) presents a useful way to present problems with the purity of polygons. He proposes a graph of the relationship between probability of correct interpretation and the size of the polygon. This type of information in a quality report would make it harder to ignore the purity issue. A procedure for determining the size-probability graph should be established.

Slivers

Just as slivers can be used to test positional accuracy, they also demonstrate attribute accuracy. The complexities of discriminating some map categories leads to uncertainties over the locations of borders. "Fuzziness" of boundaries also relates to the choice of scale. Thus errors in boundaries cannot be easily ascribed to either the positional or attribute components. Difficulties of analysis do not make slivers any less real. A quality report should describe whatever errors are known, even if they are hard to tie back to an error source.

LOGICAL CONSISTENCY

The previous two sections have dealt mostly with the content of a digital cartographic data base at the scale of an individual feature. These individual features fit into a context, a spatial information system. A quality standard that considered only positional and attribute accuracy would ignore some of the fundamental user needs. Logical consistency is a general term for fidelity in representing features in a data structure. This section reviews consistency testing for a broad set of generic types.

As mentioned above, the properties of logical consistency cannot be subjected to tests beyond the level of internal evidence. There is the possibility of deduction based on a sample approach, but normally properties of logical consistency are so crucial that they must be tested exhaustively. A report on logical consistency should answer these questions:

What can be checked in the data structure?
What was checked, and with what results?

Point data

A large body of digital data is generated for the simplest geometric object: points. There is not much data structure to verify for these files.

One class of tests for legal values pertains to virtually any type of data. It is about all that can be done, other than tedious manual verification. A test for legal values assures that attributes are in the set of permissible ones or that positions are in the study area. This is a weak test, but it does trap gross errors, such as US cities located on the Greenwich meridian due to a blank field. A test for legal values is so easy to implement that it should not be left out. The master file of legal values must be very carefully scrutinized.

Grid / Raster data

The process of generating gridded data is different from other approaches because the arbitrary spatial object takes precedence over the attributes. To understand the logic of a particular set of gridded data it is important to know how it was generated and processed. For example, grid cell coding strategies include center point, presence/absence, predominant type, priority by attribute, aggregations or apportioned attributes, averages interpolations and more. Particular sensors have special properties that affect the results. Some of this information belongs in a full lineage report, but it has a link back to the issue of consistency. The fidelity of encoding can be checked against other information, if it is available. For example, the areas of features, such as cities or counties should approximate published figures. However, internal evidence provides little means to verify that cells are properly encoded. A sample test from an independent source might be more powerful.

Elevation data

Topographic data can be encoded in a number of ways, but there is a common thread. A large amount of the national digital data base consists of gridded terrain (DTM or DEM depending on one's predilections). Gridded elevation data is not easy to check for fidelity and consistency. It is possible to look for "block faults" on sheet boundaries, or adjacent cells with 1000 meter differences. These checks are rather weak. More sophisticated analysis of smoothness could work, but it is landscape dependent. The tests described under the section on positional accuracy are more useful to the terrain problem.

An alternative to the gridded approach is the Triangular Irregular Network. The TIN is essentially a topological structure built on a set of scattered point locations. As a topological structure it will be considered below. However, the two dimensional topology of the triangles bear little relationship to the terrain surface. The few constraints on the topology and geometry of triangles which can be verified rather easily.

A brand-new terrain encoding scheme devised by Dutton (1983) will verify that the elevations in a region are consistent, as a matter of its internal structure. This elegant new idea demonstrates that we are not stuck with the same old alternatives. A workable standard must be able to accomodate data structures which have not been designed yet.

Vector data

The main problems of digital cartography concentrate on files of lines or areas. Geographic information systems, usually the most sophisticated consumers of digital cartographic data, require reliable data structures. The next sections cover the major alternative data structures currently in use.

Topological data structures

The topological approach to digital cartography was developed largely to address the problems of logical consistency (see Corbett, 1979 for a treatment of the theory). White (1978) reduces questions of consistency into two major issues.

First, the objects incident at each node or polygon must form a closed ring.

Second, no lines can intersect except at their ends (at nodes).

These properties are important, but they are not the only ones which can be detected.

The first property is checked by examining all the chains incident at a node or around a polygon. Either approach will detect the same errors, but the node approach will localize errors more precisely. The polygon approach is needed to verify proper embedding of holes in the polygon, although a geometric approach to polygons may be more direct. These cycling checks ensure that the network is topologically planar.

The other property mentioned by White involves geometric planarity. No chain should intersect itself or any other chain. If chains are restricted to have no detail between nodes (to be straight line segments as in DIME files), then a node cycling check should detect any intersections except in degenerate cases. With more complex chains, a separate intersection procedure is needed.

The checks proposed by White verify consistency of topological structure for cases, like DIME, where the topology was entered manually. These checks also produce some useful side effects. Chains entered twice will be detected. Unnecessary nodes of degree two can be demoted to points by joining the two chains. Chains with the same feature on each side can be detected. This operation provides a simple method to aggregate zones from a detailed data base. In addition, small polygons can be detected and merged into surrounding zones if they fall under some minimum mapping criterion.

The topological structure carries many advantages, but the structure must be reliable, so that the user can take advantage of it without worrying about errors. Software systems can be rather fragile with respect to errors in data structure. As an example, with their antiquated CUE system, the Census Bureau had a difficult time removing all the topological errors in their DIME files. But their research has lead to a new generation of software which should give DIME a better name. At least one vendor is selling DIME-like data with an emphasis on removing all consistency errors.

A major alternative for a quality standard is whether to allow a proportion of encoding errors or whether to insist on 100% clean files.

Topological encoding has become increasingly common in geographic information systems. Even some producers, such as Ordnance Survey, have gone through the massive restructuring of old spaghetti-like files. In most systems, topologically coded data is not entered manually as the DIME files were. The GIRAS system (Fegeas and others, 1983) was an early example of building a topological structure from unstructured lines and centroids - spaghetti and meatballs, as it is sometimes called. Turning spaghetti into chains involves a process of intersection of the lines, but this check was required anyway. The intersection software can form proper nodes automatically. Polygon information cames from an independent source, a visual centroid with a tag. This separate source preserves the dual approach of DIME, but with some modifications. Consistency checks detect somewhat different properties.

The intersection process not only detects intersections as potential coding mistakes, it actively inserts a new node into the data base. To a DIME purist this puts geometry at a higher level than topology. Despite the theoretical impurity, this procedure seems to be practical for manual and automated digitizing. The polygon tags can create two kinds of errors: an area with no point, and an area with two (or more) conflicting points. The remedy for the first is usually to add a point, unless the polygon was unintentional due to geometric errors. The second case often involves a missing line.

Common errors with lines include duplication, overshoots and undershoots. In order to address these problems, recent software systems (particularly ODYSSEY, Morehouse and Broekhuysen, 1982; and ARC/INFO, Dangermond, 1983) have introduced a tolerance to detect near misses as intersections. This process will provide an automated correction to all three line problems, as long as the errors are within the tolerance. This tolerance can also implement or verify minimum width rules. While the tolerance offers reduced editing time, it has an impact on positional accuracy which should be noted. The tolerance selected must be acceptable within the overall error budget.

Topological data structures are often used for polygon coverages which are exhaustive. This application makes certain occurrences into errors: Nodes of degree one are not needed, and lines with the same zone on each side are not needed. However, the topological structure is also applicable to network data, such as streams or roads, where these events are not errors. The different situations should be part of a report on logical consistency.

Polygon files

A simpler approach to polygon coverages is to store each polygon as a separate loop. With this data structure, a user might be interested in the same properties discussed under topological structures, but they will be harder to check. A polygon loop can be checked for closure easily, but it takes a lot of processing to ensure that the boundaries shared with neighbors are identical. This processing essentially constructs the topological structure, so it would be best to retain it rather than returning to the polygon form. As another example, it is easy to find small polygons (below some minimum map unit tolerance), but it is hard to merge them into some neighbor.

Other peculiarities of polygon encoding are the treatment of inliers and outliers. Most systems vary; the approach taken should be documented. It is important to note that the topological structure does not need to use retracements, the common solution with the polygon approach.

Cartographic spaghetti

A large amount of digital cartographic data has been generated to serve the purposes of display - reproducing traditional products. The term "cartographic spaghetti" refers to lines which are "no more structured than spaghetti on a plate". This somewhat pejorative term was coined in part to point up a problem in quality assessment. A spaghetti file is judged only by the products it produces; there is no logic internally that can verify that the encoding is complete and consistent.

If a file represents a set of polygons, but only with a set of feature codes, the user will ask the same questions expected from a topological system:

- Do any lines intersect?
- Are any lines entered twice?
- Are all the polygons encoded?
- Are there undershoots or overshoots?
- Are any polygons too small, or any lines too close?

Reporting all these characteristics may require conversion to a topological structure. The standard does not intend to legislate specific data structures, but some are more rigorous in the quality delivered to a user.

Alternatives for logical consistency

The issue of logical consistency is a rematch of some of the data structure debates of the last fifteen years. Perhaps this is the time to move from theoretical advantages and concentrate on the user's needs - fitness for use. The working group recognizes the advantages of the topological approach, but it does not wish to reject all other forms of data storage. The set of properties which can be verified in a topological system do provide a possible standard for full disclosure. Less complex data structures will produce a less complete statement about logical consistency.

COMPLETENESS

The issue of logical consistency begins to move from properties of individual objects to more global concerns. However, most topological checks are rather local. Other issues relate to the entirety of a cartographic data base. Some of these are discussed here under the title of completeness. This issue was not a separate heading in the previous report, but the discussion is sufficiently different from logical consistency to merit an independent section.

In a polygon coverage application, the polygons should be exhaustive of the whole area. This property can be checked by the topological procedures mentioned above, so the issue of completeness seems simple in this case. However, the user may ask more probing questions which relate to attribute accuracy as much as logical consistency. Any polygon coverage may have the purity problem discussed above. In some cases, such as administrative areas or property parcels, the purity may be quite high, but in some cases it will vary with scale.

While purity deals with errors of commision (eg. a sub-area of type B wrongly generalized into A), completeness covers errors of ommission. Completeness raises questions about how exhaustively a data base captures each type of feature. For example, a user interested in census tract thematic mapping would want a data base which defines all census tracts, not just a selection. Another user might want to use the hydrography layer from a topographic sheet, but only if it includes all navigable waters. Completeness is a common concern in analytical applications, far removed from traditional constraints such as symbolism and scale.

The call for completeness information is not difficult for most coverage data bases, particularly those with topological structures. Completeness is a particular problem for base mapping. For instance, many topographic maps include buildings as features in rural areas, then switch to a red tint for built-up areas. The feature "building" is not consistently recognized. The scale dependent, non-local decision which substitutes the built-up tint must be explained to the potential user so that the digital data base is properly used. While the urban tint is an extreme example, completeness presents other more subtle potential problems. Sometimes features are selected not for reasons of their attributes, but because nothing else appears in that portion of the map. This sort of rule is the modern equivalent of the heraldic doodle. For use in geographic information systems, the features should be included for some consistent, defensible reasons.

Geocoding

Working Group III on features is studying codes to promote the exchange of cartographic features, at least those common to a class of base maps. A wide variety of other standards for features exist. Standard codes for states, counties, census tracts, hydrological units and more have been developed. These coding schemes do not raise a quality question by themselves. However, a user may be interested in knowing the relationship between the geocodes and the objects in the data base. For example, does this file encode all census tracts within a city once and only once? In a landuse map, what universe of codes was used? Is there some hierarchical relationship between the geocodes and some other universe (implicit or explicit)? These concerns fall between attribute accuracy and completeness.

Another quality issue is the meaning of these codes over time. If a census tract is split, two new codes are made. For historical purposes these two can be reaggregated, but there is no confusion because the old code is retired. In other cases, such as the creation of a Virginia independent city, the original county does not receive a new code although its definition is altered. These rules should be made explicit. The procedures for creating new codes should be specified so that parallel creations of conflicting codes are avoided. The toughest problems are historical consistency and maintenance into the future.

REFERENCES CITED

American National Standard Institute 1974, Control Chart Method of Controlling Quality During Production, ANSI Z1.1-158 (R1974): New York, ANSI

Berry, B.J.L., and Baker, A.M. 1968, Geographic Sampling: In Berry and Marble, eds., Spatial Analysis: Englewood Cliffs NJ, Prentice Hall, p. 91-100

Bureau of the Budget, 1947, National Map Accuracy Standards: Washington DC, GPO

Canadian Council on Surveying and Mapping 1982, Standards for Quality Evaluation of Digital Topographic Data: Ottawa, Energy, Mines and Resources

Chrisman, N.R. 1982a, A Theory of Cartographic Error and its Measurement in Digital Data Bases: Proc. AUTO-CARTO 5, p. 159-168

Chrisman, N.R. 1982b, Beyond Accuracy Assessment: Correction of Misclassification: Proc. ISPRS Commission IV, p. 123-132

Chrisman, N.R. 1983, Issues in Digital Cartographic Data Standards A Progress Report: in Report 3, NCDCDS, Columbus OH

Chrisman, N.R. and Moellering, H. 1983, Issues Addressed by the National Committee for Digital Cartographic Data Standards: paper presented at Harvard Computer Graphics Week

Congalton, R.G. and Mead, R.A., 1981, A Quantitative Method to Test for Similarity between Photointerpreters: Proc. 47th Annual Meeting, ASP, p. 263-266

Cook, B.G. 1983, Geographic Overlay and Data Reliability: in Peuquet D. and O'Callaghan, J. eds, Proc. U.S./Australia Workshop on Design and Implementation of Computer Based Geographic Infomation Systems, Amherst NY: IGU, p. 64-69

Corbett, J.P. 1979, Topological Principles in Cartography: US Bureau of the Census, Washington DC

Crossfield, J.K. 1980, Proposed Minimum Accuracy Standards for Land Surveying in Arkansas: Proc. ACSM Fall Tech. Meeting, p. 459-466

Dangermond, J. 1983, ARC/INFO, A Modern GIS System for Large Spatial Data Bases: Proc. ACSM Fall Tech. Meeting, p. 81-89

Dutton, G.H. 1983, Geodesic Modelling of Planetary Relief: Proc. AUTO-CARTO 6, vol. 2, p. 186-201

Eisenhart, C. 1963, Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems: Journal of Research, National Bureau of Standards, vol. 67C, p. 161-187

Federal Geodetic Control Committee, 1974, Classification, Standards of Accuracy, and General Specifications of Geodetic Control Surveys: Washington DC, GPO

Fegeas, R.G., Claire, R.W., Guptill, S.C., Anderson, K.E., and Hallam, C.A. 1983, Land Use and Land Cover Digital Data: Section E of McEwen, R.B., Witmer, R.E, and Ramey, B.S., (eds.) USGS Digital Cartographic Data Standards, Geological Survey Circular 895-E, Reston VA: USGS

Hanigan, F.L. 1983, Houston's Metropolitan Common Data Base: Four Years Later: Surveying and Mapping, vol. 43, p. 141-151

Hayes, G.E. and Romig, H.G. 1977, Modern Quality Control: Encino CA, Bruce

Ley, R.G. 1981, An Examination of Various Cartometric Vertical Accuracy Tests: Cartographic Journal, vol. 18, p. 25-31

Loelkes, G.L. 1977, Specifications for Land Use and Land Cover and Associated Maps: Open File 77-555, USGS, Reston VA

Merchant, D. 1982a, Spatial Accuracy Standards for Large-Scale Line Maps: paper presented at ACSM-ASP Annual Meeting, Denver CO, reprinted in Report 1, Columbus OH, NCDCDS

Merchant, D. 1982b, Report on the Activities of Working Group II, Data Set Quality: in Report 2, Columbus OH, NCDCDS, p. 11-12

Moellering, H. 1982, The Goals of the National Committee for Digital Cartographic Data Standards: Proc. AUTO-CARTO 5, p. 547-554

Moellering, H. 1983, ed. Digital Cartographic Data Standards: Defining the Issues: Report 3, Columbus, OH, NCDCDS

Morehouse, S. and Broekhuysen, M. 1981, ODYSSEY User's Manual: Cambridge MA, Harvard Laboratory for Computer Graphics

Petersohn, C. and Vonderohe, A.P. 1982, Site-Specific Accuracy of Digitized Property Maps: Proc. AUTO-CARTO 5, p. 607-619

Rosenfield, G.H., Fitzpatrick-Lins, K. and Ling, H.S. 1982, Sampling for Thematic Map Accuracy Testing: Photogrammetric Engineering and Remote Sensing, vol. 48, p. 131-137

Thompson, L.S. 1981, Digitizing and Automated Output Errors, Photogrammetic Engineering and Remote Sensing, vol. 47, p. 1455-1457

White, M.S. 1978, A Geometric Model for Error Detection and Correction: Proc. AUTO-CARTO III, p. 439-456

APPENDIX 1

NOAA FORM 76-35

U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL OCEAN SURVEY

DESCRIPTIVE REPORT

Type of Survey ..Coastal Zone Map..................

Job No. ..PH-7120............ *Map No.* TP-00466....

Classification No. Final Field Edited Map *Edition No.* ..1........

LOCALITY

State ....Florida....................................

General Locality ..Monroe County...................

Locality ..Big Pine Key.............................

....

19 74 TO 19 77

REGISTRY IN ARCHIVES

DATE ....

☆ U.S. GOVERNMENT PRINTING OFFICE: 1974-762-901

NOAA FORM 76-36A (3-72)

U. S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMIN.

DESCRIPTIVE REPORT - DATA RECORD

TYPE OF SURVEY	[X] ORIGINAL / [ ] RESURVEY / [ ] REVISED
SURVEY	TP-00466
MAP EDITION NO.	( )
MAP CLASS	Final
JOB	PH-7120

PHOTOGRAMMETRIC OFFICE

Rockville, Md.

OFFICER-IN-CHARGE

Cdr. J. Collins

LAST PRECEEDING MAP EDITION

TYPE OF SURVEY	[X] ORIGINAL / [ ] RESURVEY / [ ] REVISED
JOB	PH-____
MAP CLASS	
SURVEY DATES:	19__ TO 19__

I. INSTRUCTIONS DATED

1. OFFICE	2. FIELD
General Instructions-OFFICE-NOS Cooperative Coastal Boundary Mapping, Job PH-7000 December 9, 1975 Supplement I, November 4, 1974 Supplement III, October 24, 1974 NOTE: Office and field edit instructions (1975) incorporate applicable prior operational instructions	Instructions_FIELD-July 6, 1972 Field Edit (PH-7000 General Instructic for Florida Coastal Zone Mapping) 1973

II. DATUMS

1. HORIZONTAL:	[X] 1927 NORTH AMERICAN	OTHER (Specify)
2. VERTICAL:	[X] MEAN HIGH-WATER [X] MEAN LOW-WATER [ ] MEAN LOWER LOW-WATER [ ] MEAN SEA LEVEL	OTHER (Specify)

3. MAP PROJECTION

Transverse Mercator

4. GRID(S)

STATE	ZONE
Florida	East

5. SCALE

1:10,000

III. HISTORY OF OFFICE OPERATIONS

OPERATIONS		NAME	DATE
1. AEROTRIANGULATION	BY	R. Kelly	Oct. 76
METHOD: Analytic	LANDMARKS AND AIDS BY	Inapplicable	
2. CONTROL AND BRIDGE POINTS	PLOTTED BY	J. Taylor	Apr 76
METHOD: Coradi	CHECKED BY	Inapplicable	
3. STEREOSCOPIC INSTRUMENT COMPILATION	PLANIMETRY BY	Inapplicable	
	CHECKED BY		
INSTRUMENT:	CONTOURS BY	Inapplicable	
SCALE:	CHECKED BY		
4. MANUSCRIPT DELINEATION	PLANIMETRY BY	J. McClure	Jul 76
	CHECKED BY	C. Lewis	Jul 76
METHOD: Graphic-rectified photos	CONTOURS BY	Inapplicable	
	CHECKED BY		
SCALE:	HYDRO SUPPORT DATA BY	Inapplicable	
	CHECKED BY		
5. OFFICE INSPECTION PRIOR TO FIELD EDIT	BY	C. Lewis	Jul 76
6. APPLICATION OF FIELD EDIT DATA	BY	P. Dempsey	Mar 77
	CHECKED BY	C. Lewis	May 77
7. COMPILATION SECTION REVIEW	BY	J. Battley	June 77
8. FINAL REVIEW	BY	D. Brant	Jul 77
9. DATA FORWARDED TO PHOTOGRAMMETRIC BRANCH	BY		
10. DATA EXAMINED IN PHOTOGRAMMETRIC BRANCH	BY	D. Brant	Oct. 77
11. MAP REGISTERED - COASTAL SURVEY SECTION	BY		

NOAA FORM 76-36A SUPERSEDES FORM C&GS 181 SERIES

U.S. G.P.O. 1972-769383/582 REG.

NOAA FORM 76-36B
(3-72)

U. S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL OCEAN SURVEY

COMPILATION SOURCES

TP-00466

1. COMPILATION PHOTOGRAPHY

CAMERA(S): Wild RC-10 C 3.5" focal length

TYPES OF PHOTOGRAPHY LEGEND: (C) COLOR; (P) PANCHROMATIC; (I) INFRARED B&W

TIME REFERENCE — ZONE: Eastern; MERIDIAN: 75th; [X] STANDARD; [ ] DAYLIGHT

TIDE STAGE REFERENCE
- [ ] PREDICTED TIDES
- [ ] REFERENCE STATION RECORDS
- [X] TIDE CONTROLLED PHOTOGRAPHY

NUMBER AND TYPE	DATE	TIME	SCALE	STAGE OF TIDE
74C(C) 8251	14 Mar 74	1630	1:30,000	The stage of tide is inapplicable for the color photography.
74C(C) 8288, 90, 92	14 Mar 74	1650	1:30,000	
74C(C) 8457	16 Mar 74	1520	1:30,000	Refer to 76-36B(1) for tide information.
74C(C) 8526, 28, 30	16 Mar 74	1540	1:30,000	
74CR 2197, 98	8 Nov 74	1348	1:30,000	
74CR 2302, 03,04, 05	11 Nov 74	1339	1:30,000	
74CR 2450, 51	12 Nov 74	1447	1:30,000	
74CR 2475, 76	12 Nov 74	1522	1:30,000	
74CR 2674	22 Nov 74	0957	1:30,000	
74CR 2755, 56, 57	22 Nov 74	1111	1:30,000	

REMARKS

74CR 2785, 86, 87, 88	22 Nov 74	1150	1:30,000
74CR 2861	22 Nov. 74	1402	1:30,000

2. SOURCE OF MEAN HIGH-WATER LINE:

The source of the MHW line is the tide-coordinated black and white infrared photography listed under item 1. The rectified color photography was used as an aid for interpreting cultural features and compiling the limits of vegetation and shoal and shallow areas.

Where the shoreline is obscured by vegetation, such as mangrove, the apparent shoreline symbol was used.

3. SOURCE OF MEAN LOW-WATER OR MEAN LOWER LOW-WATER LINE:

The source of the MLW line is the tide-coordinated black and white infrared photography listed under item 1.

4. CONTEMPORARY HYDROGRAPHIC SURVEYS *(List only those surveys that are sources for photogrammetric survey information.)*

SURVEY NUMBER	DATE(S)	SURVEY COPY USED	SURVEY NUMBER	DATE(S)	SURVEY COPY USED

5. FINAL JUNCTIONS

NORTH	EAST	SOUTH	WEST
TP-00460	TP-00467	TP-00472	TP-00479

REMARKS

Final junctions will be made in the Coastal Mapping Section.

NOAA FORM 76-36B
(3-72)

NOAA FORM 76-36B(1)
(7-75)

U. S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL OCEAN SURVEY

TIDE – COORDINATED PHOTOGRAPHY

TP – 00466

LOCATION AND PHOTOGRAPHY	TIDE STATIONS (In operation at time of photography)	STAGE OF TIDE	MEAN RANGE
74-CR 2197	Big Pine Key W. Side	-0.08' MHW	0.71'
74-CR 2198	Summerland Key E. Side	-0.09' MHW	1.18'
74-CR 2302, 03	Big Pine Key W. Side	+0.18' MHW	0.71'
74-CR 2304, 05	Big Pine Key N. End	+0.16' MHW	1.27'
74-CR 2450, 51	Big Pine Key W. Side	+0.32' MHW	0.71'
74-CR 2475	Big Pine Key W. Side	+0.29' MHW	0.71'
74-CR 2476	No Name Key	+0.31' MHW	0.71'
74-CR 2674	Big Pine Key W. Side	+0.06' MLW	0.71'
74-CR 2755	No Name Key	+0.05' MLW	0.71'
74-CR 2756	Big Pine Key W. Side	-0.11' MLW	0.71'
74-CR 2757	Big Torch Key	-0.16' MLW	0.76'
74-CR 2785, 86	Big Pine Key W. Side	-0.10' MLW	0.71'
74-CR 2787, 88	No Name Key	+0.03' MLW	0.71'
74-CR 2861	Big Spanish Key	+0.23' MLW	2.67'

REMARKS:

U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL OCEAN SURVEY

HISTORY OF FIELD OPERATIONS

I. [X] FIELD INSPECTION OPERATION * Aug. 1972 [X] FIELD EDIT OPERATION Sept. 23, 1977

OPERATION		NAME	DATE
1. CHIEF OF FIELD PARTY		R.R. Wagner	
2. HORIZONTAL CONTROL	RECOVERED BY	J.D. DiMare	2/77
	ESTABLISHED BY		
	PRE-MARKED OR IDENTIFIED BY		
3. VERTICAL CONTROL	RECOVERED BY	J.D. DiMare	2/77
	ESTABLISHED BY		
	~~PRE-MARKED OR~~ IDENTIFIED BY	J.D. Di Mare	2/77
4. LANDMARKS AND AIDS TO NAVIGATION	RECOVERED (*Triangulation Stations*) BY	J.D. Di Mare	2/77
	LOCATED (*Field Methods*) BY	R.R. Wagner	2/77
	IDENTIFIED BY	J.D. Di Mare	2/77
5. GEOGRAPHIC NAMES INVESTIGATION	TYPE OF INVESTIGATION: [ ] COMPLETE [ ] SPECIFIC NAMES ONLY [X] NO INVESTIGATION — BY		
6. PHOTO INSPECTION	CLARIFICATION OF DETAILS BY	R.R. Wagner	2/77
7. BOUNDARIES AND LIMITS	SURVEYED OR IDENTIFIED BY	N/A	

II. SOURCE DATA

1. HORIZONTAL CONTROL IDENTIFIED

PHOTO NUMBER	STATION NAME
	Refer to the Field Report

2. VERTICAL CONTROL IDENTIFIED

PHOTO NUMBER	STATION DESIGNATION
74C8457	Z 272, A 273, B 273, E 273, K 328
74C8526	C 273
74C8292	L 273, M 273, P 273
	H 70 Reset

3. PHOTO NUMBERS (*Clarification of details*)

74C8288, 8290, 8292, 8451, 8457, 8526, 8528, 8530; 74CR2756, 2786, 2787

4. LANDMARKS AND AIDS TO NAVIGATION IDENTIFIED

Landmarks and non-floating aids were either located or verified by field methods

PHOTO NUMBER	OBJECT NAME	PHOTO NUMBER	OBJECT NAME
74CR2756	Micro Tower		

5. GEOGRAPHIC NAMES: [ ] REPORT [X] NONE

6. BOUNDARY AND LIMITS: [ ] REPORT [X] NONE

7. SUPPLEMENTAL MAPS AND PLANS

8. OTHER FIELD RECORDS (*Sketch books, etc.* **DO NOT** *list data submitted to the Geodesy Division*)

Refer to Field Report bound in this Descriptive Report

U.S. Government Printing Office 1976 - 665-661/1110, Region No. 6

NOAA FORM 76-36D
(3-72)

U. S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

RECORD OF SURVEY USE

TP-00466

I. MANUSCRIPT COPIES

COMPILATION STAGES			DATE MANUSCRIPT FORWARDED	
DATA COMPILED	DATE	REMARKS	MARINE CHARTS	HYDRO SUPPORT
Office - Class III	7/7/76	Special request from Requirements Branch	11/8/76	

II. LANDMARKS AND AIDS TO NAVIGATION

1. REPORTS TO MARINE CHART DIVISION, NAUTICAL DATA BRANCH

NUMBER	CHART LETTER NUMBER ASSIGNED	DATE FORWARDED	REMARKS

2. ☐ REPORT TO MARINE CHART DIVISION, COAST PILOT BRANCH. DATE FORWARDED: ____
3. ☐ REPORT TO AERONAUTICAL CHART DIVISION, AERONAUTICAL DATA SECTION. DATE FORWARDED: ____

III. FEDERAL RECORDS CENTER DATA

1. ☐ BRIDGING PHOTOGRAPHS; ☐ DUPLICATE BRIDGING REPORT; ☐ COMPUTER READOUTS.
2. ☐ CONTROL STATION IDENTIFICATION CARDS; ☐ FORM NOS 567 SUBMITTED BY FIELD PARTIES.
3. ☐ SOURCE DATA *(except for Geographic Names Report)* AS LISTED IN SECTION II, NOAA FORM 76-36C. ACCOUNT FOR EXCEPTIONS:
4. ☐ DATA TO FEDERAL RECORDS CENTER. DATE FORWARDED: ____

IV. SURVEY EDITIONS *(This section shall be completed each time a new map edition is registered)*

EDITION	SURVEY NUMBER / DATE OF PHOTOGRAPHY	JOB NUMBER / DATE OF FIELD EDIT	TYPE OF SURVEY / MAP CLASS
SECOND EDITION	SURVEY NUMBER: TP - ____ (2) DATE OF PHOTOGRAPHY	JOB NUMBER: PH - ____ DATE OF FIELD EDIT	TYPE OF SURVEY: ☐ REVISED ☐ RESURVEY MAP CLASS: ☐ II. ☐ III. ☐ IV. ☐ V. ☐ FINAL
THIRD EDITION	SURVEY NUMBER: TP - ____ (3) DATE OF PHOTOGRAPHY	JOB NUMBER: PH- ____ DATE OF FIELD EDIT	TYPE OF SURVEY: ☐ REVISED ☐ RESURVEY MAP CLASS: ☐ II. ☐ III. ☐ IV. ☐ V. ☐ FINAL
FOURTH EDITION	SURVEY NUMBER: TP - ____ (4) DATE OF PHOTOGRAPHY	JOB NUMBER: PH - ____ DATE OF FIELD EDIT	TYPE OF SURVEY: ☐ REVISED ☐ RESURVEY MAP CLASS: ☐ II. ☐ III. ☐ IV. ☐ V. ☐ FINAL

NOAA FORM 76-36D

☆U.S. Government Printing Office: 1975 - 665-661/1100, Region No

CHART 11536
DRAWING NO 20 & 20 [illegible]
SHEET 1 OF
PROJECT 8 3 2 3 0 1

NOAA FORM 76-71 (8-73)
USCOMM-DC 53687-P72

HISTORY OF CARTOGRAPHIC WORK

U.S. DEPARTMENT OF COMMERCE
NOAA
NATIONAL OCEAN SURVEY

REMARKS

NC REC CORR	TO REPRODUCTION	PROOF REC'D	PRINT DATE
FUTURE	NC NE NP RP TO	☐ NO PROOF ORDERED	
DATE	DATE	PROOF VER	

ITEMS	COMPILED BY (Signature)	STARTED DATE	COMPLETED DATE	TIME (Hours)
1,18-33	JAY SITERMAN	10/22/81	3/3/82	8.0
2	[illegible] Kohl	10-29-81	10-29-81	.5
3-17	Richard Thomas	2/12/82	2/20/82	32.0
[illegible]	[illegible]	[illegible]	[illegible]	8.[illegible]
38	B. Fernandez	9-25-82	9-25-82	0 5
[illegible]0	B. Fernandez	4-7-83	4-8-83	7.0
41	[illegible] Kohl	4/20/83		1.0
42	B. Richter	6/16/83		40
46-49	[illegible] Kausch	8 31 83		40
[illegible] 53	[illegible] Kausch	11-28-83		4.0

CHECK LIST	COMPILER (Initial)	REVIEWER (Initial)
Data applied thru all larger scales where applicable		
All overlaps at same scale in full agreement		
General comparison with next larger scales		
Photo corrections examined		
All colors checked for interference		
All color changes clarified for engraver		

ITEMS	REVIEWED BY (Signature)	STARTED DATE	COMPLETED DATE	TIME (Hours)
1	Fannie B Powers	10/20/81	10/20/81	—
12-17, 2, 3-11	JAY SITERMAN	2/22/82	2/24/82	6 5
18-33	Fannie B Powers	3/8/82	3/8/82	4 0
34-37	Allen [illegible]	8-18-82	8-18-82	1.0
38, 42-45	Fannie B Powers	1/20/82	1/23/82	[illegible]
40	[illegible] Kohl	4/20/83		1[illegible]

APPLIED PART	APPLIED FULL	ITEM NO.	SOURCE OF INFORMATION FILE NO	DATE	AUTHORITY	TYPE OF INFORMATION	LOCALITY (Use Lat./Long.)	INFORMATION APPLIED
	✓	1	CL-1242	1981	USPS	Wk	33°54'57" 78°01'08"	Added ED to visible Wk
	✓	2	CL-1152	1981	USPS	Misc	—	No Correction
	✓	3	Bp 113723	Dec 80	CE	Hydro	33°48' 78°03'	Applied thru Cht 11537 Revised sdgs
	✓	4	Bp 115079	Oct 80	C of E	Hydro	33°50.5' 78°03.0'	No corr outside channel } thru 11537
	✓	5	Bp 115080	Oct 80	C of E	Hydro	33°51.0' 78°02.0'	No corr outside channel }
	✓	6	Bp 113858	6/81	C of E	Hydro	Smith Island Chan	Revised 30 ft curve thru cht 11537
	✓	7	L-1170	1981	USPS	Misc	Cape Fear River	Revised sdg through 11537
	✓	8	L 103	1980	C of E	Bridge	33°54'30", 78°06'45"	Applied
	✓	9	L-1743	1980	USPS	pier	33°52'37" 78°00'07"	Examined no corr
	✓	10	Bp 112590	1980	NOS	Limits	Entire chart	Revised Territorial Sea limits and
								Contiguous Zone
	✓	11	L 1574	1980	USPS	TANK	33°54'46.9", 78°00'04.9"	Applied thru 11534A (Added TANK ⊙)
	✓	12	Bp 115914	1/6/82	C of E	Hydro	33°50'00", 78°03'30"	No corr to Hydro thru 11537
	✓	13	Bp 115915	1/6/82	C of E	Hydro	33°52'00" 78°00'40"	No Corr
	✓	14	L-966	1981	Fed Reg	Colregs	—	Changed Par 82 to 80
Cont'd	✓	16	CL-1152	1981	USPS	Tank	33°54'38", 78°06'51" ✓	Added ● } Tanks through 11537

Superseded by L-100 83

CHART

APPLIED PART	APPLIED FULL	ITEM NO.	FILE NO	DATE	AUTHORITY	TYPE OF INFORMATION	LOCALITY	INFORMATION APPLIED
			—			Tank	33°54'59", 78°09'40" ✓	Added ○ } Tanks through 11537
Cont'd	✓	17	L-1570	1980	USPS	Wk	33°54' 77°59'16"	Revised visible wk to Sub[illegible]
						Wk	33°56'51", 77°55'36"	Revised visible wk to "
	✓	18	L-1731	1980	USPS	TANK	33°54'52.5", 78°16'15.8"	Added TANK ⊙ through 11534A
	✓	19	Carto Order	3/80	NOS	Auth Note	—	Revised note
	✓	20	Carto Order	4/80	NOS	Comments	Request	Added note
	✓	21	New Print	11/79	NOS	Misc	Entire Chart	Applied
	✓	22	Bp 113012	3/81	CE	Sdgs	33°48'58" 78°04'19"	Applied thru 11537 - revised one sdg
	✓	23	Bp 113661	5/81	CE	Cond. Survey	Baldhead - Caswell Ch.	" " " " " "
	✓	24	L-730	1980	CE	Groins	33°54'36", 78°07'00"	Added groins thru 11534A
	✓	25	11537	11/81	NOS	HWL-SDG	Baldhead Island	Revised HWL, LWL to bring into
								agreement
							Dump Site	Revised sdgs in Dump Site
								vicinity to bring into agreement
								Revised date
	✓	26	Mag Var.	1982	NOS	Compass Roses	—	Revised roses to reflect 1982
								data
	✓	27	Bp 113011	3/81	C of E	After Dredge	Baldhead Shoal Channel	Exam/N.C.
	✓	28	Bp 112644	2/81	C of E	During Dredge	Baldhead Shoal Channel	Exam/NC
	✓	29	Bp 112645	2/81	C of E	"	" " "	Exam/NC
	✓	30	Bp 112646	1/81	C of E	After Dredge	" " "	Exam/N.C.
	✓	31	Bp 112647	1/81	C of E	Condition Survey	" " "	N/C
	✓	32	Bp 112648	12/80	C of E	"	Southport & Battery I	NC
	✓	33	Bp 113662	5/81	"	Condition Survey	Southport & Battery I.	Exam/N.C through 11537
	✓	34	BP 117172	82	C of E	Hydro	Lower Swash Channel	Apply thru Chart 11537
	✓	35	BP 117173	82	C of E	Hydro	BALD HEAD - CASWELL Ch	Apply thru Chart 11537
	✓	36	BP 117174	82	C of E	Hydro	BALDHEAD Shoal Channel	Apply thru Chart 11537
	✓	37	BP-117175	82	C of E	Hydro	BALDHEAD Shoal Channel	Apply thru Chart 11537
	✓	38	L-968	1982	USPS	COND PWR CAB	33°54'51" - 78°00'00"	Applied --- [illegible]

(p. 73 follows)

DIGITAL CARTOGRAPHIC DATA STANDARDS: ALTERNATIVES IN DATA ORGANIZATION

A Progress Report
Working Group on Data Organization

Edited by Tim Nyerges

ABSTRACT

In cycle 1, the National Committee for Digital Cartographic Data Standards (NCDCDS) Data Organization Working Group identified three major issues: terminology, modeling and data interchange. This paper presents work from cycle 2; current efforts directed at examining alternatives for the major issues. The problem of transfering data between systems is discussed initially in regards to application topics and functional uses of data. Alternatives for data transfer are presented as interchange scenarios, existing data interchage formats, and alternatives for generalized data interchange. Although primary concern is data interchage per se, the process of data transfer can be understood better when discussed in regard to interchange modeling. General models for the data interchange process, i.e. interchange models, are presented to summarize the relationship between data structures (data models) and intermediate formats used for transfering data. A general milieu for the development of a data interchange format is presented as a set of guidelines. These guidelines will focus our efforts when we propose an interim data interchange standard.

Working Group I is composed of the following members:

Dr. Timothy Nyerges (Chair)	GeoSystems Software Inc.
Dr. Donna Peuquet (Vice-Chair)	University of California, Santa Barbara
Mr. Fred Billingsley	Jet Propulsion Laboratory
Dr. A. Raymond Boyle	University of Saskatchewan
Dr. Hugh Calkins	S. U. N. Y. Buffalo (Terms Representative)
Mr. Robert Edwards	Oakridge National Laboratory
Mr. Robin Fegeas	U. S. Geological Survey
Mr. William Liles	Xerox Corporation
Mr. David Pendleton	National Ocean Service
Mr. Dan Rusco	Defense Mapping Agency

INTRODUCTION

Purpose

The Data Organization Working Group received its charter from the National Committtee for Digital Cartographic Data Standards (NCDCDS) Steering Committee in June, 1982. Its purpose is to identify issues, discuss alternatives and propose standards for digital cartographic data organization.

History of Meetings and Communication

The first meeting of the full membership of the working group was in August, 1982 and the second in the Spring of 1983. Several members wrote position papers to identify issues. Some papers were discussed at the initial meeting, and later papers were circulated among the members for discussion at the second meeting. Those discussions formed the basis of the previous report: "Issues in Digital Cartographic Data Standards". Subsequent meetings, discussions and literature searches form the basis of this report on the alternatives for data organization.

Major Issues

Three major issues have been identified: terminology, modeling and data interchange. The rapidly growing field of digital cartography has given rise to a large technical vocabulary with many words having similar meanings. This tends to confuse issues. Part of our effort is to identify these terms and to pass them on to Working Group IV on Terminology for clarification. This is a continuing process.

The modeling issue has been as much a problem of terminology as it has been of conceptual focus. Modeling may mean something different to mathematicians than it does to cartographers or computer scientists, especially in the context of data organization. This has been evident with the term "data model" which has received widespread use in the computer science literature. A data model is a framework for a data language (Kunii, 1983 p.1). A data model is the formal abstract definition of the data objects and the relationships together with the operations on the structures supported in the language. A data language consists of a data definition language (DDL) that defines the data objects and relationships, and a data manipulation language (DML) that implements operations on these objects. Although the concept of data model is an appropriate topic for data organization, our focus is somewhat different with respect to modeling. We use a new term called interchange model. An interchange model is a framework for the overall interchange process. Several interchange models will be presented later in the paper.

Data interchange is the most important of issues from a practical point of view. This importance is indicated by the numerous attempts by many organizations to define data interchange formats in addition to the documented waste of resources caused by duplication of data set building efforts (Government Accounting Office, 1982).

Alternatives

Cycle 2 of the working group efforts concerns identification and discussion of the alternatives for data organization. Several alternatives for each issue are being considered. Primary focus is on alternatives for data interchange. A major aspect of this effort is to discuss data interchange in terms of interchange modeling. As defined previously, an interchange model focuses on the overall process of interchanging data.

The outline of the paper is as follows. The first section provides background information; application topics and functional uses of data are discussed. Alternatives for data transfer are presented as interchange scenarios, existing data interchange formats, and alternatives for generalized data interchange. General models for the data interchange process, i.e. interchange models, are presented to summarize the relationship between data models and their respective formats. A general milieu is presented to provide guidelines for developing an interchange format. A summary and conclusion follow as the final sections.

THE SYSTEM PROBLEM

The system problem is discussed in terms of application topics and the range of functional use of the data.

Application Topics

Historically, map making and map analysis have been analog processes, operated by humans. Because of the ability of humans to grasp the interrelations among elements on maps, and the ability to readily interpret and interrelate maps of differing scales and content, there have been few attempts at coordinating the various cartographic archives. Now, however, at least four forces are causing a change to digital processes: 1) increasing amounts of cartographic data are being generated and must be stored, cataloged, and retrieved; 2) increasing amounts of cartographic data processing are being automated; 3) increasing amounts of related but non-cartographic data are being obtained in digital form; and 4) Increasing sophistication in digital map registration and analysis of multiple sets of data is resulting in a call for digital cartographic data. At the same time, increasing computer capabilties over the past decade have made the computer storage, cataloging, retrieval, and processing practical.

The incompatiblity of various archives, data base formats, and processing has been documented (Billingsley, F. C., 1980; Stockman, D. A. 1983; McEwen, R. B. et. al., 1983). Some characteristics of the present situation at this time are:

- Various data are produced in unique formats. In assembling data for internal use each data generator typically has formatted these with little consideration towards standardizing with anyone else.

- The data are stored in archives in incompatible formats and cataloging structure. There is little coordination between the various methods by which data are geographically referenced, or uniform ways to express what system is in use (see section on linneage in previous paper).

- Digital representation of cartographic data requires specific and unambiguous coding of the relationships among the various elements of the data. There are several distinct , viable methods and structures for digitally encoding these relationships. Two major structuring schemes are in use within the U.S. Geological Survey (McEwen et al. 1983).

- Other Federal agencies each have their own formats and structures.

- Many voluminous and popular digital data sources use their own structures for data storage and dissemination. Recent format definitions for imagery begin to consider the cartographic community, but the defintions are not complete.

- Data are geographically distributed. While this should not cause a problem, the absence of cross catalogs ensures that the data will be difficult to locate, and encourages further diversity in data structure.

Several critical factors emerge in the use of the various data, singly and in combination:

- Digital cartographic analysis requires a codification of spatial relationships and coding of the attributes of the cartographic features. This coding must be compatible with the needs of the computer assisted mappping systems.

- An important problem in the use of image data is the requirement to geometrically rectify the imagery, and to register the images to the ground. Most imagery in an archive is only approximately (if at all) georeferenced (geographic information is included with the data to allow subsequent registration), and very few data elements are placed within a recognized coordinate system.

- The use of combined data sets requires co-registration of one to another. For disparate types of data, this must be done via registration of each to a common reference system such as geographic coordinates. Co-registration is implicit. This is generally a user problem, as the required registration data is normally absent.

The concerns for common data formats and geocoding conventions cut across all disciplines from computer-assisted cartography to remote sensing. At this stage it is intractable, clumsy, and inefficient for diverse users to use a given data set. It is also difficult to assemble a data set from a variety of sources for general use. The time has come to consider coordinated data systems which retrieve data from the various archives and combine and deliver them to users in a common format. A common approach will eventually propagate into the data preparation and archiving stage.

Functional Uses of Data

Digital cartographic data may be used in at least two ways: 1) analysis in the digital domain; and 2) conversion to conventional map form for display. For the first, the topological information must be specifically included; for the second, the cartographic attributes such as line weight, color, symbology and labels must be included. Unless the use of the digital data is clearly known, both sets of information are required.

Analysis of the cartographic data usually requires geographic locations and topology . Some attribute data may be important, but data such as line weights or color is less important. Such analysis is the primary domain of geographic information systems. Sophistication of implementation varies widely with such systems. Multiple overlay problems are difficult, as are problems involving spatial relational operators. This is particularly true if the data are obtained from several sources.

Typically, conversion of cartographic data to map form will be done in specialized centers. Here the attribute data and geographic location of the cartographic data points are critical, but the topology is less important. Implementation of data reduction and map production capability will normally accompany the data generation capability, but the presence of various incompatible generating facilities prevents the ready interchange of this data.

Digital images and the availability of digital image based information systems have allowed the combined analysis of image and cartographic data (Nagy, G. and S. Wagle , 1979). In such systems the cartographic data can be converted to two dimensional grid-cell arrays. The cartographic data and the images are then co-registered in a common cell format (Billingsley and Bryant, 1975). Uses of this data vary considerably and are a function of the capabilites of the systems.

ALTERNATIVES FOR DATA TRANSFER

Data Interchange Scenarios

A scenario can be understood as a description of the major course of events pertaining to a process. Several alternative scenarios for data transfer exist. The optimal one is determined by the needs and resources of the organizations involved. On a general level, scenarios differ based on the answer to the question: "Where can two systems be interfaced?". Five possibilities are identified (Yan et. al., 1977).

1) Transfer in non-digital form:
Data may be transfered directly in map form.

2) User adopts suppliers format:
The data recipient builds a system around the supplier's existing file formats. This approach is usually not feasible unless a user is developing a new system which is a subset of the suppliers system or the file structures have been standardized previously.

3) Integrate user's data into supplier system:
Whenever more manpower or system sophistication is available from the data supplier, it is advisable to turn over the users data set to the supplier who will then process the data to the

user's specifications.

4) Transfer in a specialized format:
A specialized interchange format is specified and an interface program is written by the user or supplier, or both.

5) Transfer in a standardized format:
This usually involves an agreement on a stanadard format among several organizations who will be exchanging data. Each organization writes programs to convert between their internal format and the interchange format.

Of the five scenarios certainly number five would be the most resource efficient in the long term. Scenario five can be examined further in terms of the major characteristics which affect the final result of a data transfer. Some of these characteristics include:

1) Mandate - Without a clear mandate and resulting problem definition the task will proceed ill-defined.

2) Funds - One of the most important parameters of the interchange process is the level of funding needed to complete the task.

3) Time - A time frame need be specified to gauge the progress of the effort. The volume of data per time frame can be used to judge the progress.

4) Tools - The software and hardware environment can affect the productivity of data interchange. Better tools (more cpu cycles and structured software) can facilitate the data transferral process.

5) Documentation - Internal and external documentation must be available to communicate the process to everyone involved in the same manner. Incomplete documentation leads to misunderstanding and can lead to poor results.

6) Energy - The professional committement of the personnel is crucial to a well organized and timely process.

Lessons learned from a case study (Yan, et. al. 1977) between Statistics Canada and the Canada Department of Fisheries and the Environment help point out the importance of the above characteristics . These lessons are:

1) Estimates of time and commitment required for a succesful data transfer should be made very carefully. Three to four calender months may not be uncommon.

2) Prepare complete and precise specifications before the transfer process begins.

3) Close communication ties must be maintained throughout the entire exchange process.

4) The number of uncontrolled or unknown elements involved at any one time should be minimal.

5) Follow the simplest, most direct, transfer route with respect to systems and human resources. The liklihood of success seems to decrease geometrically as the number of intermediate systems and personnel involved increases.

6) Several alternate data transfer methodologies should be considered. The strategy requiring the least initial outlay of resources is not necessarily the best in the long run.

As discussed previously, the needs and applications of users in the cartographic community vary so widely that the number of alternatives for data formats is of considerable proportion. To define a single format for all applications is a monumental task, as can be indicated by the number of specialized formats in existence. Thus, the range of implementation for a standardized format can take on characteristics from the simple to the very sophisticated to meet the needs of users.

Extant Data Interchange Formats

Several data interchange formats which have some application to digital cartographic data have been identified:

Australian Feature Coded Digital Mapping Data Standard (Standards Association of Australia, 1981)

Canadian EDP Standards Applied to Digital Topographic Data (Canadian Council on Surveying and Mapping, 1982)

Computer Assisted Mapping and Records Activity System (American Public Works Association, 1979)

Digital Line Graphs (Domaratz, M. A. et. al., 1983)

Graphics Standards Planning Committee Metafiles (Graphics Standard Planning Committee, 1979)

Information Interchange Data Descriptive File (International Standards Organization, 1983)

Initial Graphics Exchange Specification (National Bureau of Standards, 1980)

Landsat CCT Family of Tape Formats (Lansat Technical Working Group, 1979)

Standard Format for Digital Linear Data (Defense Mapping Agency, 1983)

Standard Interchange Format (Intergraph, 1980)

Standard Interchange Structures (Consultative Committee for Space Data Systems, 1983)

Generalized Data Interchange for Digital Cartographic Data

The extant formats discussed in the previous section have been designed to fulfill certain needs, and generally do not satisfy all needs for many different organizations. Within the context of the level five data transfer scenario the proliferation of specialized data interchange formats is becoming a problem much like the proliferation of internal system formats, implying a lack of standards. With the recognition of this problem, there has come a move to ameliorate it with a concept called the "Family of Formats" first discussed and implemented in terms of Landsat imagery (Landsat Technical Working Group, 1979). Although the concept is not a complete solution for cartography (let alone imagery interchange), it is a beginning which has provided some direction for organizing unrelated interchange formats. Two major alternatives exist for a baseline interchange standard in the context of a family of formats. These are:

1) Define a suitable superstructure surrounding the data which can identify the data source, format, and other ancillary information as required for the translation. This superstructure would be an "envelope" around the data sets as currently being distributed. Recognizing that there are numerous data archives, data bases and data interchange formats in existence, this would not require them to translate the data into some

new format. This would not require new software at both the archive and user facilities if data interchange software already exists.

2) Define a new, expandable format family, to work within the superstructure. This would include standard ways of specifying the various data types. Such a solution would closely follow the more popular current formats. With this, define standard ways, such as ANSI X3.61-1978 for Geographic Point Locations, or ANSI X3.43-1977 for Local Time of Day, for coding the various elements. This structure would need to be defined and maintained and updated at the National level.

INTERCHANGE MODELING

General Model of the Interchange Process

The extent of the interchange process can be brought into focus through the development of a general model of the interchange process. Such a model draws from similar efforts in the computer science literature that concern generalized data translation techniques for data base interchange among diverse systems and data models (ANSI, 1975; ANSI, 1977; Fry, 1981).

The interchange of data from one computer system to another can be represented initially as a transformation function T operating upon a source set S to produce an object set of data O, as depicted below:

$$S \dashrightarrow T \dashrightarrow O \qquad [1]$$

This interchange process would be adequate in the case where S is a set of data files of known format in a known environment, i.e. a single vendor's hardware and systems software, T is a special-purpose reformatting program, and O is the set of data files in the desired format. Clearly, this process is too limited to be of any general use in interchanging structured data between diverse systems. Therefore, the model must be extended to one in which S is a set of formatted files possibly including a structured data base with a known schema. T is decomposed into T1 and T2 in order to take into account different source and object hardware systems, and O is a set of files in a different format possibly including a structured data base with the same schema:

$$S \dashrightarrow T1 \dashrightarrow I \dashrightarrow T2 \dashrightarrow O \qquad [2]$$

After these extensions are made to model [1], a new set I is required. This set represents an intermediate form of the data (Fry, 1981) created on one hardware/software configuration by T1, suitable for reading on a different hardware/software configuration using T2 to produce O. This model is used by many of the extant interchange formats discussed previously.

At this point it is important to understand the full implications of the use of the concept of schema. As used in the computer science literature, this term refers to the description of the organization of data relationships in an integrated data base under a particular data model as implemented in a particular data base management system (DBMS). In practice, it will be required that data be interchanged between systems using not only different hardware, but different schemas and different data base management systems. These DBMS's could be based upon different data model's,.e.g. S could be a network-based DBMS and O might be a DBMS based upon the relational model. Therefore, interchange model [2] must be revised once more to account for these requirements:

S --> T1 --> I1 --> T2 --> I2 --> T3 --> O [3]

As before, new constructs appear in the model to serve the additional requiremenets listed above. In this case, "I" has been decomposed into I1, which includes the schema structures as well as data in S, and I2, which contains the schema structures as well as the data values. The function T2 maps the S schema structures of S onto the schema structures of O. T3 creates the object data base O from the information contained in I2.

An examination of the data forms and transformations required by this interchange model reveals several major issues facing the development and standardization of data interchange on a wide scale:

- The data models and their implementations as data base management systems must evolve to a state suitable to allow the definition of a standard specification for each. Progress is being made in this area and draft standards for the relation and network models are under consideration for adoption by the American National Standards Institute (ANSI, 1983a, 1983b).

- Standard intermediate forms for data and structure interchange would be the next step after the data models have been standardized.

- The availability of standard software packages implementing the transformation functions and intermediate forms for standard models over a variety of vendor hardware and data base management systems would complete the move to generalized data interchange.

Model [3] above meets the interchange requirements of the typical user of available data base management systems as applied to business and much scientific data processing. Studies have been carried out to determine the relative completeness of this approach, including implications bearing upon the various data models and their manipulation languages (Taylor, 1982). Model [3] represents the state-of-the-art in interchange model development in the computer science field.

Model for Digital Cartographic Data Interchange

The distinguishing characteristic of spatial data systems is the frequent requirement to capture, store and operate upon topological information inherent in the data elements. The goal of exchanging this kind of data is related in a fundamental way to the exchange of this topological information, which in turn, is related in a fundamental way to the deeper representation of knowledge in a computer.

Any attempt to extend model [3] into this domain is risky, since this would imply that a data model and suitable intermediate forms to cover this additional and very complex requirement exists. However, one might envision what such a model might be like in general terms and from the representation deduce the kinds of developments that would have to occur in order to implement it.

Extrapolating from the previous models and proceeding in the same manner:

S --> T1 --> I1 --> T2 --> I2 --> T3 --> I3 --> T4 --> O [4]

In devising this model, two additional constructs are needed to account for the process of decomposing and recreating topology, as well as having the schema information mapped into the intermediate forms. The elements of the model are as follows:

S: The source set of data files and/or data base structured under a data model with the capability to represent the topological information inherent in cartographic data. This sopurce data base is to be transported from its resident hardware configuration to the object data base O under a different hardware configuration without loss of data content or distortion of cartographic meaning.

T1: A transformation function implemented as a standard software procedure on the source data base S and produces the first intermediate form I1.

I1: The first intermediate form driven by the S schema consisting of a data format containing data items, a format containing schema transfer rules for the S data model, and a format containing rules for the transfer of the topological information under the S data model. The I1 structures represent a one-to-one mapping with the S data base structures.

T2: A transformation function implemented as a standard software procedure for the data model used by S which reads the I1 form and produces I2, the standard interchange form.

I2: This is the standard interchange form which is independent of hardware and data model and consists of formatted data items, schema information, and topology information for the data set originally from data base S. Information in this form can be transferred to any system that has T3 and T4 software procedures for completing the transfer to the hardware configuration and data model being used for that system.

T3: A transformation function implemented as a standard software procedure for the data model used by O which reads the I2 form and produces I3.

I3: The last intermediate form driven by the O schema consisting of a data format containing data items, a format containing schema transfer rules for the O data model, and a format containing rules for the transfer of the topological information under the O data model. The I3 data structures represent a one-to-one mapping with the O data base structures.

T4: A transformation function implemented as a standard software procedure on the object hardware which reads the I3 form and produces the object data base O.

O: The object data base and/or set of files consisting of the data and topological information from the source data set S. The data interchange process is complete after the object data base has been brought up under the object data model and can be operated upon by the spatial data operators of that model.

An examination of the data forms and transformations required for the Cartographic Data Interchange Model in [4] reveals the deep complexities unique to the exchange of spatial data. These are intimately related to the nature of digital cartographic data analysis and the requirement for specific encoding of spatial relationships in addition to the coding of feature attributes. The use of digital cartographic data for analysis by computer places the burden on the data model and its structures to represent the topological relationships that the human perceptual process automatically sorts out when viewing a cartographic display. In some sense, then, advances in digital cartographic data models will parallel the advances made in the computer science field for knowledge-base systems (Codd, 1979). Such systems are identified by features that enable the interpretation of meaning, the drawing of inferences, and the planning of high-level objectives with little human intervention. Cartography will push the state-of-the-art in the development of this kind of data base management system in much the same way as the large computational requirements of meteorology and similar fields pushed the developments leading to present day super-computers.

Milieu for Defining an Interchange Format

Within the interchange process we develop a general milieu as a guide for data interchange definition. An understanding of the general milieu for data interchange is as important as the data interchange format itself. The alternatives for the milieu provide the background and basic motivations underlying the definition of the format. Such a milieu might consist of:

1) The data interchange format must meet a defined philosophy. The goals of the format must be clearly stated and agreed upon by all concerned. Then the details can be worked out. Without this, arguments over the details will have no ground for resolution. A suitable philosophy for our purpose might be:

   - Provide an interchange format which will allow the local user to read the data set to determine the basic logical structure.

   - Provide a format which will allow the inclusion of all necessary data such as: feature information, data quality, spatial data type, locational definitions, spatial relationships, and ancillary data. - Provide a format that will be expandable to include any future types of information which will arise.

- Provide a data organization which will handle, raster, vector and text data in the same format.

- Provide a structure which is useable within a local system or general purpose data base as well as for data interchange, thus minimizing reformatting problems for the user.

- Provide coding standards to handle the various information and attributes required.

2) The data interchange must be easy to use.

3)The data interchange must be applicable to both archive requirements and data bases maintained in an incremental fashion.

4) The data interchange must be applicable to present distribution formats and present and future media. The format should be media independent to allow transmission by various means such as tape, digital video disks, and electronic transmission, and to allow storage on any of the available storage media. Any specific translations required would be the function of the media controller.

5) The data interchange must be compatible with standards in other disciplines. Mutual cooperation should be sought wherever this is possible.

6) The data interchange must meet computer imposed requirements. If the field of cartography is to exploit technological advances in automatic data processing, digital cartographic data must be structured, coded and formatted to be reasonably compatible with the data standards set forth by the ADP industry. Some of the more important computer imposed requirements are:

- The format must be independent of the target (user) computer to allow maximum transportability of the data.

- The transfer of logically different sets of data must be possible on one physical medium or in one equivalent electronic transmission. It must be possible to add these onto the medium incrementally.

- The structure must be able to handle multi-tape or multi-transmission volumes.

- Adopt ANSI or other tape labeling standards where applicable.

- The physical medium description should not be a part of the format.

- The logical structure should be self-describing such that the translation should be performed without recourse to written documentation.

- Variable record size should be accommodated, although fixed size may be recommended.

- No maximum record size should be imposed . However, due note should be taken of the maximum record sizes which can be handled by the mini- or micro-computers which may be in user facilities. Similarly, due note must be taken of the minimum record size of various computer systems.

- No records are required to be present except the first, a Volume Directory, plus one File Descriptor for each file included. All others may be omitted. Thus the structure is maximally compactable.

The above characteristics of the milieu is a working list, and should not be assumed to be the final milieu.

SUMMARY

Applications for digital cartographic data have been discussed in terms of problems with interchnaging data between systems. Two major uses of data were discussed: analysis and display. The alternatives for data transfer involve five major scenarios. Each scenario describes a method for transferring data. Many of the existing interchange formats appear to be closer to scenario four, specialized transfers. Specialized formats are fine for specific application needs. The most complete approach is scenario five, a multi-organization standardized data interchange. This approach was further discussed in terms of characteristics affecting the interchange process.

Alternatives for data interchange range from the simplest of extant formats, which will apply to only a limited application area, to the more sophisticated "family of formats" which will apply to many applications.

Several interchange models were presented to provide a focus for the overall process of data interchange. These models can be used to better understand the nature of the alternatives under consideration.

A milieu for the data interchange process was presented. The milieu consists of specific alternative guidelines which will be used to help orient the data interchange decision process.

CONCLUSIONS

The number of existing data interchange formats demonstrates a considerable interest in being able to share resources. The varied applications and orientation of these extant formats precludes any one format being chosen as the best one to work for all data.

The family of formats concept facilitates a common approach to a flexible data format definition. The concept provides the mechanism which can help guide organizations to a better definition of their interchange format, no matter which is chosen for the application specific data records.

The two major alternatives for specifying a format with the family of formats concept will allow organizations to continue to use their existing interchange formats. However, should a better format be devised, the concept will permit the use of such a structure.

The interchange model and the general interchange milieu will help us in chosing an approach for the proposed interim standard to be defined in the next cycle(cycle 3) of our efforts.

REFERENCES

American Public Works Association, Utility Location and Coordination Council, 1979. "File Format for Data Exchange Between Graphic Databases", Vol. 3 of Computer Assisted Mapping and Records Activity System, January, 1979.

American National Standards Institute, Inc. 1975. American National Standard for the Representation of Numeric Values in Character Strings for Information Interchange, ANSI X3.42-1975, August 4, 1975.

American National Standards Institute, Inc. 1977. American National Standard Code for Information Interchange, ANSI X3.4-1977, June 9, 1977.

Amercian National Standards Institute, 1983a. "Draft Proposed Relational Database Language", ANSI X3H2 Document X3H2-83-49, February, 1983.

American National Standards Institute, 1983b. "Draft Proposed Network Database Language with Proposed Introduction and Concepts Section", ANSI X3H2 Document X3H2-83-48, March, 1983.

Billinsley, F. C. 1980. "Data Base Systems for Remote Sensing", in Lecture Notes in Computer Science, Vol 81, Proceedings of a Meeting "Data Base Techniques for Pictorial Applications", A. Blaser, Ed., Springer Verlag, 1980, pp. 299-313.

Billingsley, F. C. and N. Bryant, 1975. "Design Criteria for a Multiple Input Land Use System", NASA Earth Resources Symposium, NASA TM X-58168, Vol I-B, pp. 1389-1396.

Canadian Council on Surveying and Mapping, 1982. Canadian EDP Standards Applied to Digital Topographic Data, Energy, Mines and Resources, Canada, April, 1982.

Codd, E. F. 1979. "Extending the Database Relational Model to Capture More Meaning'" ACM TODS 4:4(December, 1979), pp. 397-434.

Defense Mapping Agency Aerospace Cartography Department, 1983. "DMA Standard Format for Linear Digital Data", Defense Mapping Aerospace Center, November, 1980.

Domaratz, M. A. et. al. , 1983. USGS Digital Cartographic Data Standards. Digital Line Graphs from 1:2,000,000-Scale Maps", U. S. Geoglogical Survey Circular 895-D, U. S. Geological Survey.

Energy, Mines and Resources Canada, 1977. "Standard Format for the Transfer of Geo-coded Polygon Data", available from Canada Centre for Remote Sensing, 717 Belfast Road, Ottawa, Ontario, K1A OY7, Document code For-SDP-0001.

Fry, J. P. et al., 1981. Draft Specification for a Structured Data Interchange Form, rep. NBSIR 81-2315, National Bureau of Standards, July, 1981.

Government Accounting Office, 1982. "Duplicative Federal Computer-Mapping Programs: A Growing Problem", Report GAO/RCED-83-19, November 22, 1983.

Intergraph Corporation, 1980. "Graphics Standard Interchange Format", November, 1980.

International Standards Organization, 1983. "Information Processing - Specification for a Data Descriptive File for Information Interchange'" ISO DIS8211, 1983-05-24.

Kelley, J. C., Wolf, R., Kennicott, P. and R. Nagel, 1981. "A Technical Briefing on the Initial Graphics Exchange Specification (IGES)", rep. NBSIR 81-2297, Natioanl Bureau of Standards, July 1981.

Landsat Technical Working Group, 1979. "The Standard CCT Family of Tape Formats", Document CCB-CCT-0002, available through Ms Lynn Buhler, ORI, Inc. 1300 Spring Street, Silver Spring MD, 20910.

McEwen, R. B., Calkins, H. W. , Ramey, B. S., 1983. USGS Digital Cartographic Data Standards, Overview and USGS Activities, USGS Circular 895-A.

Nagy, G. and S. Wagle, 1979. "Geographic Data Processing", Computing Surveys, Vol 11:2 June, 1979, pp.139ff.

Standards Association of Australia, 1981. "Interchange of Feature Coded Digital Mapping Data", Australian Standard 2481-1981. Standards Association of Australia Standards House.

Stockman, D. 1983. "Memorandum M-83-12", on the Coordination of Federal Digital Cartographic Data Programs, OMB, April 4, 1983.

Taylor, R. W., 1982. "Experience Using Generalized Data Translation Techniques for Database Interchange", Computers and Standards 1 pp. 111-118.

Yan, J. , S. Witiuk, and T. Fischer, 1977. "Exchanging Spatial Systems Technology: Related Issues and a Case Study", <u>Applications of Geographic Information Processing</u> papers <u>presented at the Second Annual Symposium Workshop</u> of the National Capital Geographic Information Processing Group. University of Ottawa, Dept. of Geography and Regional Planning, April 1, 1977.

General Comments on the Work of the National Committee for Digital Cartographic Data Standards, Cycle 2

NAME: ____________________ WORK PHONE: ____________________

ADDRESS: ____________________ EVENING PHONE: ____________________

____

____

____

Your Comments Please:

SIGNED: ____________________

DATE: ____________________

Cut Here

------------------------------Fold------------------------------

Place
Postage
Here

Prof. Harold Moellering, Chairman
Nat. Com. Dig. Cart. Data Stds.
Numerical Cartography Laboratory
Ohio State University DB 158
Columbus, Ohio 43210

------------------------------Fold------------------------------