ISSUES IN DIGITAL CARTOGRAPHIC DATA STANDARDS

Report #6

Digital Cartographic Data Standards:
An Interim Proposed Standard

Harold Moellering, Editor

87-306

January, 1985

National Committee for Digital Cartographic Data Standards
Numerical Cartography Laboratory
The Ohio State University
158 Derby Hall
154 North Oval Mall
Columbus, Ohio
U.S.A. 43210-1318

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-951. Supplementary funding for Working Group supplied by Defense Mapping Agency. Production assistance has been provided by the Ohio State University Research Foundation (RF Project #716442).

PREFACE

This report is the sixth in the series which discusses the work of the National Committee for Digital Cartographic Data Standards. It contains an Interim Proposed Standard for cartography for the United States. The report is divided into two major sections, the first is the Interim Proposed Standard itself, and the second is the supporting documentation for the standard that is a discussion of the logic and rationale of the standard along with other supporting material.

The Committee has organized a special set of public hearings on this Interim Proposed Standard to be held at the AUTO-CARTO 7 meetings in conjunction with the Spring American Congress on Surveying and Mapping meetings in Washington, D.C. on Thursday March 14, 1985 beginning at 10 A.M. All interested parties are invited to participate in these hearings.

This report represents the work of the Committee for the third year of operation, that of developing the Interim Proposed Standard. 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

Page

PART I. THE STANDARD

Digital Cartographic Data Standards: Interim Proposed Standards: Introduction 1

An Interim Proposed Standard for Digital Cartographic Data Organization 5

An Interim Proposed Standard for Digital Cartographic Data Quality 17

An Interim Proposed Standard for Digital Cartographic Features 23

An Interim Proposed Standard for Terms and Definitions 37

PART II. SUPPORTING DOCUMENTATION

Background and Organization for the Development of an Interim Proposed Standard: Supporting Documentation, by Harold Moellering 45

An Interim Proposed Standard for Digital Cartographic Data Organization: Supporting Documentation, edited by Timothy Nyerges 67

An Interim Proposed Standard for Digital Cartographic Data Quality: Supporting Documentation, by Nicholas Chrisman 113

An Interim Proposed Standard for Digital Cartographic Features: Supporting Documentation, edited by Warren Schmidt 135

An Interim Proposed Standard for Terms and Definitions: Supporting Documentation, edited by Harold Moellering 147

DIGITAL CARTOGRAPHIC DATA STANDARDS: INTERIM PROPOSED STANDARDS

0.0 INTRODUCTION

The National Committee for Digital Cartographic Data Standards was founded in 1982 to develop standards that would be useful to the cartographic profession for the task of data base use and exchange. To date the Committee has completed the first two cycles of work, the first of defining the issues involved, and the second of examining the alternatives, and has almost completed the third of defining an Interim Proposed Standard. This report sets forth that Interim Proposed Standard for examination, discussion and comment by the profession. Following this series of events the Interim Proposed Standard will serve as the basis for empirical testing in cycle four.

0.1 ORGANIZATION AND STRUCTURE OF THIS REPORT

This report is divided into two major sections: the body of the report that contains the proposed specifications for the Digital Cartographic Data Standards, and a second part that contains supporting material and discussion of the standards proposed. Section 1. by the Working Group on Data Organization describes an approach to cartographic data base exchange. Most of the attention has been focused on developing an exchange superstructure with the widest applicability in cartography while some attention has been devoted to a couple of test exchange modules to verify the concept. Section 2. by the Working Group on Data Set Quality focuses on a "truth in labeling" approach to data quality. Fundamentally, the producer of a data base is expected to produce a statement that describes the quality levels of the data contained in the file. Section 3. by the Working Group on Cartographic Features sets forth the principals of an amalgamated set of feature definitions and presents about 20 example definitions and associated terms. Section 4. by the Working Group on Terms and Definitions proposes a standard set of primitive and simple cartographic objects which are contained in digital cartographic data bases. A second part contains a set of terms and definitions associated with the standards proposed by the other three Working Groups.

* Note: As the reader examines the standards proposed, the supporting material should also be perused so that the logic of the recommended standards can be clearly seen.

0.2 MEMBERS OF THE COMMITTEE

The Committee is made up of a Steering Committee , four Working Groups and an Executive Committee. The Steering Committee is the primary organizational structure for the effort and its members are the ones who created the working groups in 1982 and defined the scope of their activities. The Steering Committee is also the group that formally votes on the standards according to the American National Standards Institute rules being followed. The Executive Committee is composed of the Chairs and Vice Chairs of the Working Groups and the Committee itself. This group leads the work of the Committee on a day to day basis. The Working Groups focus on specific aspects of the standards problem and are composed of experts knowledgeable about those specific aspects of the problem.

The members of the Steering Committee are as follows:

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

The members of Working Group I, Data Organization, are as follows:

Tim Nyerges, Northwest Cartography, Chair
Donna Peuquet, Univ. of California, Vice Chair
A.R. Boyle, University of Saskatchewan
Robert Edwards, Oak Ridge National Laboratories
Fred Billingsley, Jet Propulsion Laboratories
Bill Liles, Xerox Special Information Services
Robin Fegeas, Geological Survey
David Pendleton, National Ocean Service
Dan Rusco, Defense Mapping Agency

The members of Working Group II, Data Set Quality, are as follows:

Nicholas Chrisman, Univ. of Wisconsin, Chair
Charles Poeppelmeier, Defense Mapping Agency, Vice Chair
Dean Merchant, Ohio State University
John Davis, Kansas Geological Survey
George Rosenfield, Geological Survey
George Johnson, National Ocean Service
Wallace Crisco, Bureau of Land Management
John Stout, Petroleum Information Inc.
Gunter Greulich, Survey Engineers of Boston
David Meixler, Bureau of the Census

The members of Working Group III, Cartographic Features, are as follows:

Warren Schmidt, Digital Mapping Unlimited, Chair
Robert Rugg, Virginia Commonwealth Univ., Vice Chair
Joel Morrison, Geological Survey
Richard Hogan, National Ocean Service
Beth Driver, Technology Service Corp.
Fred Tamm-Daniels, Tennessee Valley Authority
Mary Clawson, ITT Research Institute
Leslie Kemp, Defense Mapping Agency

The members of Working Group IV, Terms and Definitions, are as follows:

Dean Edson, E-Quad Associates, Co-Chair
Harold Moellering, Ohio State Univ., Co-Chair
Erich Frey, National Ocean Service, Vice Chair
Hugh Calkins, SUNY Buffalo
Frank Beck, Geological Survey
Mark Monmonier, Syracuse University

0.3 AMERICAN NATIONAL STANDARDS INSTITUTE (ANSI) PROCEDURES

The National Committee for Digital Cartographic Data Standards operates under the auspices of the American Congress on Surveying and Mapping and is not an ANSI committee. However the Committee operates under the ANSI rules in Appendix A "Model Procedure for an Accredited Standards Committee" as they apply to the work of the Committee because these rules are those most widely accepted.

0.4 MAINTAINING AGENCY

Although these standards statements are still in the proposed stage, it is clearly recognized by the Committee that the inherent nature of these standards require that a maintaining agency and procedures be identified in a future cycle of work.

(p.5 follows

AN INTERIM PROPOSED STANDARD
FOR DIGITAL CARTOGRAPHIC DATA ORGANIZATION

1.0 A GENERAL DIGITAL CARTOGRAPHIC DATA INTERCHANGE FORMAT

The data interchange format must meet a defined set of objectives. The objectives must be clearly stated then the details can be worked out. Without this, arguments over details will have no ground for resolution. A suitable set of objectives for a general digital cartographic data interchange format is:

1) Provide an interchange format which is documented within the transfer media which will allow the recipient to read the data set to determine the basic logical and physical organizaton with minimal specific external information.

2) Provide a format which will allow the inclusion of all necessary data such as: feature information, data quality, spatial and other data types, locational definitions, spatial and other relatonships, ancillary data, and map symbology.

3) Adopt industry accepted standards to handle the various informaton required: e.g., Federal Information Processing Standards, American National Standards Institute standards, International Standards Organization standards or other labeling, coding ad transmission standards where applicable.

4) Provide an interchange format 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, disks and electronic transmission. The format shall allow storage on any of the available storage media.

5) Provide an interchange format that is computer independent to allow maximum transportability of data and permits flexibility in tradeoffs among transportability, storage overhead and processing efficiency.

6) Provide an interchange format that will allow transfer of logically different data sets on one physical medium or in one equivalent electronic transmission. It must be possible to add incrementally onto the medium.

These objectives for data interchange are accomplished by using an approach to data definition as proposed in the International Standards Organization (ISO) Draft International Standard (DIS) 8211. This standard follows a general Tag-Length-Sequence convention for defining data structures and the data formats specifying the storage of data within the structures.

1.1 Data Definition Sections

A data interchange shall consist of one or more files. A file is a collection of related records. A record is a collection of related data fields. Each data field stores an item of data. Each of these is treated as a logical unit.

The first record of a file must contain at least a record CORE section, and optionally a CORE EXTENSION section and DATA section. Coding in the CORE and in the CORE EXTENSION shall use the American Standard Coded Information Interchange (ASCII) 7-bit character set (ANSI X.34,1977) or where applicable, special control codes as specified in this document. Coding in the DATA section may be ASCII or Binary, as defined by the context of the interchange. This shall be noted in the Data Descriptive information of the CORE EXTENSION section.

When data interchange takes place through electronic transmission in an open networked environment (ISO/TC97/SC16, 198Ø) between machines, confusion between binary codes and special control shall be eliminated by using only ASCII characters in the CORE and CORE EXTENSION sections.

1.2 Data Definition Fields

The CORE Fields and CORE EXTENSION Fields shall consist of a series of field sets, each set composed of a two-character (decimal integer) size-field, a data field. and a unit separator (US). The data field may be subdivided to include a Tag subfield, an ASCII group separator (GS), and other data. Absence of the GS shall indicate no tag. In the last field set, the US shall be replaced by a record separator (RS).

(Note: Non-printable control characters 1/13 (ASCII GS), 1/14 (ASCII RS) and 1/15 (ASCII US) are represented in this document by GS, ";" (or RS) and "&" (or US) respectively.)

This logical format for data definition is recognized as the format of the file by assuming that this format is being used and finding the USs and the RSs in their proper places. If not found, then the data is defined with another format, and external documentation must be consulted.

1.2.1 Core Fields

The CORE section shall consist of specific field sets. The meaning of a field set is determined from the contents of a field set except for the first two field sets: Maintenance Authority and Format ID. In all cases, the first two field sets shall be the Maintenance Authority and Format ID (defined below).

A CORE field set may be constructed with or without tags which provide for the reference to an interpretation of the fields. When a tag is used in the field set, the field set shall be subdivided with the tag appearing in the first subdivision. In this case a field set would be composed of the following:

Data Length, Tag, ASCII GS, Data Field, US.

The Data Length shall indicate the length of the Tag, GS, Data Field and US by using a two-character (decimal integer) size-field. A GS missing from a field set shall indicate that no Tag is being used. In the absence of a tag, the order of field sets in the CORE mst be determined by reference to external documentation.

Alternatively, one field set can be designated as a tag list, eliminating the need for tags in other field sets. This tag list shall appear early in the CORE, immediately following the Format ID. Its data field would be subdivided with the first section being a TAGLIST tag. The ASCII GS, and variable length sub-fields, separated with commas, containing the series of CORE Tags would then follow in their respective order. This will allow the local implementor to use any desired Tag structure, including none at all. The CORE TAGLIST field set list is then considered to be an ordered list of CORE field sets. The field sets may include the tags defined in the TAGLIST as an option.

The CORE fields shall be specified in the following order (if no tags are used):

MAINTENANCE AUTHORITY

This field defines the authority under which the format is defined and maintained.

FORMAT IDENTIFICATION (FORMAT ID)

This field defines the specific format identification, including revisions as defined by the maintenance authority.

DATA START POSITION

This field defines the byte postion of the data in the record, relative to the first byte of the record.

RECORD TYPE AND SUBTYPE

A four byte code using ASCII characters.

DATA DEFINITION INDICATOR (DDI)

This field defines the style and location of Data Definition (DD) information pertaining to the data fields in the CORE EXTENSION. The DDI may take one of three forms, depending on the data defintion style to be used:

	DDI	Entry Map
No Data Definition At All	Zero	None, or all zero
Inline DD, Long Form	Terminator Symbol	Non-zero Integer
Inline DD, Short Form	Terminator Symbol	None, or all zero
DDR/DR Groups	Grouping Code	Non-zero Integer

The DDI Terminator Symbol shall be an ASCII printable symbol, but not), (or *, which have specific reserved meanings. The DDI Grouping Code shall be an ASCII alphanumeric character, but not zero.

RECORD NUMBER

The first record in each file shall be number "1".

RECORD LENGTH

The total logical record length, in bytes.

EXTERNAL AUTHORITY (EXTAUTH)

Under some conditions, the structure and meaning of the data fields can be determined by reference to known external documents. This field defines the source of those documents. This field shall apply to the entire file when found in the Volume Descriptor, or only to those data records carrying the same DDI, when found in a DDR.

EXTERNAL FORMAT (EXTFMT)

This field defines the document identification version as defined by the EXTAUTH. This field shall apply to the entire file when found in the Volume Descriptor, or only to those Data records carrying the same DDI, when found in a DDR.

PACKING FACTOR

Defines the division of physical records into logical records.

RECORD REPETITION INDICATOR

Allows subsequent data records to omit the CORE. This indicator shall contain the number of subsequent data records without a CORE or a CORE EXTENSION. The Data fields of those records shall be the same structure as those of the current record.

1.2.2 Core Extension Fields

The CORE EXTENSION section in a record shall be defined according to the type of record, Data Definition Record (DDR) or Data Record (DR). The major components and functions of these records are as follows:

Record	Component	Function
DDR	Leader	Identifies the DDR Contains the entry map (sizes of the tag, length, and position fields of the corresponding directory entries in this record)
	Directory	Gives tag, length, and position (relative to the start of the Data Description component) of each Data Descriptive field in this record.
	Data Descriptions	Structure of each corresponding Data Field in the DR.
DR	Leader	Identifies the DR Contains tge entry map (sizes of the tag, length, and position fields of the corresponding directory entries in this record)
	Directory	Gives tag, length, and position relative to start of the Data component) of each data field in this record.
	Data Fields	These fields have the format as described in the corresponding DDR Data Descriptions component.

The Tags described in the DDR and DR above relate the DDR Directory entries to the corresponding DR Directory entries. This implicitly links the data fields to their respective data descriptive fields. No tags appear in the DDR Data Descriptions or in the DR Data Fields.

The 8211 data definition capability presented here has been extended to include an ability to indicate the grouping and uses of various types of records, define records of various formats within one file, enclose and identify existing data in formats created by different organizations, and provide for transmission of binary data.

1.3 Different Forms of Data Definition

1.3.1 DDR Components with DR Components

When the optional DDR Components are included in the record, the fields will appear in the following order:

DATA DEFINITION RECORD (DDR) LEADER

Identifies a DDR. Contains an entry map and certain format information. See the 8211 defining document.

DIRECTORY

Multiple entries, one for each identified data field in the Data Description section, as follows (styled after 8211):

Data Description Field Tag	ASCII Chars of t length
Data Description Field Length	ASCII Integer of m digits
Data Description Field Position	ASCII Integer of n digits

DATA DESCRIPTION section of the DDR

Defines the structure of the DR data fields, and has a structure for each entry (corresponding to Directory entries), in accordance with 8211:

Field Control	Defines the type of field: integer, complex, etc.
Separator	RS/US
Field Name	Optional, a user supplied name
Label	Optional labels for subfields
Format Control	Optional Fortran Style (See Note 1) Format Designations needed

Note 1: Because the Fortran structure definitions do not include binary fields, the following designation for binary fields is used:

nZx...x(L or R)y...y

where,

n is the number of repetitions (ommitted if =1)
Z indicates a binary field
x is the field length in bits (decimal)
L or R indicates left or right justification of the bits
y is the number of significant bits (decimal)

1.3.2 Inline Structure

When a shorter form than the entire DDR/DR is desired, the structure would use an Entry Map in the Long Form, but would not use and Entry Map in the Short Form. The Entry Map consists of fields describing the sizes of Tag, Length, Position in the Data Directory. These are:

	Field Size	
Size of Directory Length Field	I1	Integer (=m length)
Size of Directory Position Field	I1	Integer (=n length)
Reserved	I1	ASCII zero
Size of Directory Tag Field	I1	Integer (=t length)

The Tag, Length and Position are then:

Tag	Same tag as used in the Directory, followed by the US.
Length	Length of corresponding DR data field, followed by the US.
Position	Relative position of the corresponding DR data field in the data area (first position is zero), followed by the US.
Format Control	Fortran Style Format Designation (See Note 1), followed by the RS. A format label for a linear array is one in which a series of individual label subfields describes the components of a major data field. Fields for linear arrays shall use the chosen Terminator Symbol of the DDI to separate the sub-fields.

1.3.3 Short Form Inline Structure

"Format Control" is the only structuring used in the short form of the inline structure. For each data field, a Fortran Style Format Designation (See Note 1) shall be used, followed by a unit separator. The last Format Control Field shall be followed by a Record Separator, instead of the unit separator. Linear array format control fields shall use the Data Definition Indicator terminal symbol as internal separators.

1.3.4 No Data definition At All

This form of structural definition requires the CORE as introductory information, but does not require any other information for definition.

1.4 Different Forms of Data Records

The Data Component shall consist of data fields in the same manner as described in Section 1.3 for each of the different structural definitions corresponding to CORE and CORE EXTENSION fields. The CORE EXTENSION is interpreted in the Data Component to be a "Data Record CORE EXTENSION".

In Data Records using the Inline Structure, Short Form Inline Structure, and/or No Data Defintion At All, the data fields shall be contiguous without unit separator or record separator terminators. Extenal format control definitions shall be the only means of encoding and decoding the data field values in the transmission when these definitions are used.

REFERENCES

ANSI X3.4 (1977) "American National Standard Code for Information Interchange", American National Standards Institute.

ISO DIS 8211 (1983) "Specification for a Data Descriptive File for Information Interchange", Draft International Standard, International Standards Organization.

ISO/TC97/SC16 (1980) "Open Systems Interconnection, Basic Reference Model", International Standards Organization, Technical Committee 97, Subcommittee 16, December, 1980.

APPENDIX A. Examples of File Definition

This Appendix is not part of this standard and is used for explanation purposes only.

Files may be defined and transmitted in several versions, depending upon the robustness desired and the degree of field structure specification. These forms will use the components described in the standard in various combinations:

1. No Data Definition At All

This would be used when little flexibility in data interchange is needed by an organization. This would be coded as zero in the Data Definition Indicator. External documentation must be sought for all data record structures.

Major components are:	CORE DATA (Untagged fields)
Recognition Coding in DDI:	Zero
CORE Entry Map Length	= Ø (No Entry Map)

2. Data Definition (DD) included within the Data Record

The DD information would follow the core and precede the data fields. This structure might be used when each data record has a unique structure, and/or where the overhead of the full 8211 structure is undesireable. Thus, the shortened definition is appropriate. The core and a data definition structure are derived from the 8211 structure.

Major components are:	CORE, including Entry Map DIRECTORY INLINE STRUCTURE DATA (Untagged Fields)
Recognition coding in DDI:	Non-conflicting Terminator Symbol
Core Entry Map Length	> Ø

The short form of the inline data structure record is:

Major components are:	CORE, including Entry Map SHORT FORM STRUCTURE DATA (Untagged Fields)
Recognition coding in DDI:	Non-conflicting Terminator Symbol
Core Entry Map Length	= Ø (No Entry Map)

3. Sets of Data Definition Records (DDR) and Data Records (DR)

Data Definition Records (DDR)

The 8211 definition requires only one DDR per data record group, with all fields tagged in the Directory and Data Descriptions. In the structure defined in the standard, multiple groups are allowed with all of the records in a record group having the same format. Thus, one such DDR would precede a set of data records, each of which has the structure as defined in the preceding DDR. This requires recognition coding in the DR which indicated that there is an external DDR. This will be indicated by a Grouping Code in the DDI of each DDR and DR, using the same symbol throughout the DDR-DR group.

The 8211 structure will start immediately after the core. Although there is some overlap between the purposes and information in the core and the 8211 DDR leader, the 8211 leader will be included in its entirety.

When the 8211 DDR or DR core is used, the Core Entry Map is not needed; the 8211 Entry Map (part of the leader) shall convey the required information.

DDR major components are:	CORE 8211 DDR LEADER including ENTRY MAP DDR DIRECTORY DDR DATA DESCRIPTION
Recognition Coding in DDI:	Desired Record Grouping Code
Core Entry Map Length	= 0
8211 Entry Map	> 0

For 8211 details, see the 8211 document.

Data Records

The entire DR leader will be included (or not), as indicated in Par. 5.3.1.3 of the 8211 document. The DDR/DR repetitive characteristics would be coded in RP 6 of the Leader, using the coding of Par 5.3.1.3 of the 8211 document. However, note that the DR Base-Address-of-Data will be different in the DRs, depending on the repetitive (or not) inclusion of subsequent leader and directories. Therefore, the DR leader address-of-data field shall be defined to apply to the concurrent record, and the data start must be calculated by the software for subsequent records. Data locations as given in the Directory are relative to the start of the Data area (first position = zero).

DR major components are:	CORE 8211 DR LEADER including ENTRY MAP DR DIRECTORY DATA FIELDS, WITH TERMINATORS
Recognition coding in DDI:	Record Grouping Code to match DDR DDI
Core Entry Map Length	= 0
8211 Entry Map	> 0

For details see the 8211 document.

SUMMARY OF INTERPRETATION OF THE DATA DEFINITION INDICATOR CODING

	DDI	Entry Map
No Data Definition At All	Zero	All Ø
Inline DD, Long Form	Terminator Symbol	Non-zero integer
Inline DD, Short Form	Terminator Symbol	All Ø
DDR/DR Groups	Grouping Code	Non-zero integer

AN INTERIM PROPOSED STANDARD
FOR DIGITAL CARTOGRAPHIC DATA QUALITY

2.0 COMPONENTS OF A QUALITY REPORT

Digital cartographic data shall include a quality report. This standard describes the five sections required in the quality report: lineage, positional accuracy, attribute accuracy, logical consistency and completeness. Each section of the report will contain reference to temporal information and currency.

The purpose of the quality report is to provide detailed information for a user to evaluate the fitness for a particular use. This style of standard can be characterized as "truth in labelling", rather than fixing arbitrary numerical thresholds of quality. To implement the standard, a producer is urged to include the most rigorous and quantitative information available on the components of data quality described below.

The statement prescribed by the National Map Accuracy Standards (Bureau of the Budget, 1947) does not provide a complete quality report. It is recognized that the National Map Accuracy Standard statement may constitute the sole quality report available for certain existing products.

2.0.1 Form of a Quality Report

The quality report can be issued as a paper document or encoded on computer-compatible media in the form prescribed by Section 1.2.5.2 of this Standard. Since the quality report will function in the assessment of fitness for use, it shall be obtainable separately from the actual data. The digital data transmission may contain the quality report, in whole or in part, but, as a minimum, it must contain a reference to the quality report and how to obtain it.

2.0.2 Testing

In sections 2.3 and 2.4 of this Standard, there are options described for different categories of testing. Informed assessment of fitness for use is best served by the most rigorous types of tests. However, this standard leaves the level of testing optional.

2.1 LINEAGE

The lineage section of a quality report shall include a description of the source material from which the data were derived, and the methods of derivation, including all transformations involved in producing the final digital files. The description shall include the dates of the source material and the dates of ancillary information used for update. The date assigned to a source shall reflect the date that the information corresponds to the ground, however, if this date is not known, then a date of publication can be used, if declared as such.

Any data base created by merging information obtained from distinct sources must be described at sufficient detail to identify the actual source for each element in the file. In these cases, either a lineage code on each element or a reliability overlay will be required. A reliability overlay is a collection of points, lines and areas organized to represent quality information for another set of map information. If a reliability overlay is transmitted in digital form, it shall be encoded according to the standards of Section 1 of this Standard.

The lineage section shall also include reference to the specific control information used, whether benchmarks or triangulation stations. Control from the National Geodetic Reference Network shall be identified according to identifiers in that system, while other points used for control shall be described with sufficient detail to allow recovery.

The lineage section shall describe the mathematical transformations of coordinates used in each step from the source material to the final product. The locations of any control points for coordinate transformations shall be given. The methods used to make coordinate transformations must be documented. To fulfill this standard, it is acceptable to make reference to separate documentation for the coordinate transformation algorithm used, but the specific parameters applied must be described for the particular case. Documentation of a transformation algorithm must include the nature of computational steps taken to avoid loss of digits through roundoff and must include a set of sample computations including numerical values of coefficients to confirm equivalence of transformations. The documentation of a transformation algorithm must be available on request by a user obtaining digital data even if that user is not licensed to use the particular software.

2.2 POSITIONAL ACCURACY

All coordinates used for the transfer of digital cartographic data must have a known (and expressed) relationship to latitude and longitude. This standard is implemented by the use of currently recognized standard reference ellipsoids (for horizontal measurements) and standard geoids (for vertical measurements). These standards are set by the Federal Geodetic

Coordinating Committee (1974). The dates of the geodetic standards and of the datum used must be referenced.

Quality of control surveys must be reported using the procedures established in the geodetic standard. If a separate control survey has been used, it must be described in the standard form, even if results fall below the recognized classification thresholds.

Descriptions of positional accuracy must consider the quality of the final product after all transformations. The information on transformations forms a part of the lineage section of the quality report.

Measures of positional accuracy can be obtained by one of the following optional methods:

2.2.1 Deductive Estimate
(based on knowledge of errors in each production step)
Any deductive statement must include reference to complete calibration tests and must also describe assumptions concerning error propagation. Results from deductive estimates must be distinguished from results of other tests.

2.2.2 Internal Evidence
FGCC procedures will be used for tests based on repeated measurement and redundancy such as closure of traverse or residuals from an adjustment.

2.2.3 Comparison to Source
When using graphic inspection of results ("check plots") the geometric tolerances applied must be reported, and the method of registration must also be described. Use of check plots shall be included in the lineage section.

2.2.4 Independent Source of Higher Accuracy
The preferred test for positional accuracy is a comparison to an independent source of higher accuracy. The test must be conducted using the rules prescribed in the proposed Accuracy Specifications for Large-Scale Line Maps (American Society of Photogrammetry, 1985). The definitions of independence and higher accuracy in the ASP standard apply. When the dates of testing and source material differ, the report shall describe the procedures used to ensure that the results relate to positional error, not to temporal effects. The numerical results for precision and bias, as well as the number and location of the test points must be reported. A statement of compliance to a particular threshold is not adequate in itself.
This test may only be applicable to well-defined points.

The report of any test of positional accuracy shall include the date of the test.

2.3 ATTRIBUTE ACCURACY

Accuracy assessment for measures on a continuous scale shall be performed using procedures similar to those used for positional accuracy.

Accuracy tests for categorical attributes can be performed by one of the following methods. All methods shall make reference to map scale in interpreting classifications.

2.3.1 Deductive Estimate

Any estimate, even a guess based on experience, is permitted. The basis for the deduction must be explained. Statements such as "good" or "poor" should be explained in as quantitative a manner as possible.

2.3.2 Tests Based on Independent Point Samples

A misclassification matrix must be reported as counts of sample points crosstabulated by the categories of the sample and of the tested material. The sampling procedure and the location of sample points must be described.

2.3.3 Tests Based on Polygon Overlay

The misclassification matrix must be reported as areas. The relationship between the two maps must be explained; as far as possible, the two sources should be independent and one should have higher accuracy.

The report of a test of attribute accuracy shall include the date of the test and the dates of the materials used. In the case of different dates, the report shall describe the rates of change expected in the phenomena classified.

2.4 LOGICAL CONSISTENCY

A report on logical consistency shall describe the fidelity of relationships encoded in the data structure of the digital cartographic data. The report shall detail the tests performed and the results of the tests.

Tests for permissible values can be applied to any data structure. Such a test can detect gross blunders, but it does not ensure all aspects of logical consistency.

A data base containing cartographic lines can be subjected to the following general questions:

Do lines intersect only where intended?
Are any lines entered twice?
Are all areas completely described?
Are there any overshoots or undershoots?
Are any polygons too small, or any lines too close?

Different tests can be applied to address these questions, but the quality report shall contain a description of the tests applied or a reference to documentation of the software used.

The report shall state whether all inconsistencies were corrected or it shall detail the remaining errors by case.

The report must include the date on which the tests were applied. When corrections and modifications occur after the test for logical consistency, the quality report should indicate how the new information is checked for logical consistency.

2.5 COMPLETENESS

The quality report must include information about selection criteria, definitions used and other relevant mapping rules. For example, geometric thresholds such as minimum area or minimum width must be reported.

In encoding cartographic features, standard geocodes (such as the feature codes described in Section 3.2 or in the FIPS codes for states, counties, municipalities and places) shall be employed as far as possible. Deviations from standard definitions and interpretations must be described.

The report on completeness shall describe the relationship between the objects represented and the abstract universe of all such objects. In particular, the report shall describe the exhaustiveness of a set of features. Exhaustiveness concerns spatial and taxonomic (attribute) properties, both of which can be tested. A test for spatial completeness can be obtained from topological tests for logical consistency that respond to the questions in 2.4. Tests for taxonomic completeness operate by comparison of a master list of geocodes to the codes actually appearing in the file. The procedures used for testing, and the results, shall be described in the quality report.

REFERENCES

American Society of Photogrammetry, Committee for Specifications and Standards, 1985, Accuracy Specifications for Large-Scale Line Maps, to appear in _Photogrammetric Engineering and Remote Sensing_.

Bureau of the Budget, 1947, _National Map Accuracy Standards_: Washington DC, GPO reprinted in M.M. Thompson, 1979, _Maps for America_: U.S. Geological Survey, Reston VA, p. 104

Federal Geodetic Control Committee, 1974, _Classification, Standards of Accuracy, and General Specifications of Geodetic Control Surveys_: Washington DC, GPO 1980-0-333-276 (also NOAA--S/T 81-29)

Appendix 3A

SAMPLE FEATURE DEFINITIONS

The following list is an illustrative sample of approximately 300 of the 1,000 feature terms considered to date by the Working Group on Cartographic Features. These terms have been assembled into generic features that differ from each other only in their attributes: for example, a "stream" is a generic feature of which "river," "brook," "creek," "slough," "rivulet," etc. are examples of different sizes of streams that are given different names in various parts of the country. The definitions suggested for each generic feature are based on a search for the simplest, most general definition either found in an existing source, or compiled by the Working Group's staff from a combination of existing sources. Attributes have been derived from these definitions as have the lists of "included" terms. The principal sources used in this process include:

The American Heritage Dictionary of the English Language
Bruder (ed.), Nautical Chart Manual
Canadian Council on Surveying and Mapping, Draft Report, Volume I, Standards for the Classification of Cartographic Features
Defense Mapping Agency, Feature File (DMAFF) Data Collection Guide
Defense Mapping Agency, Product Specifications for DLMS Data Base
Geographic Names Information System Documentation, Appendix B, Feature Class Definitions
Monkhouse, A Dictionary of Geography
Moore, A Dictionary of Geography
Schmieder et al., A Dictionary of Basic Geography
Snead, Coastal Landforms and Surface Features
Stamp, Dictionary of Geography
U.S. Naval Oceanographic Office, Navigation Dictionary
U.S. Naval Oceanographic Office, Glossary of Oceanographic Terms

In the course of the feature definition work, a data base is being created that includes not only preferred definitions, but also a complete list of the definitions of each feature and attribute considered by the staff. This data base is intended for use by the Working Group and other review bodies in the process of confirming, or recommending changes in, the suggested selection of feature and attribute categories.

AN INTERIM PROPOSED STANDARD FOR

DIGITAL CARTOGRAPHIC FEATURES

3 CARTOGRAPHIC FEATURES

The purpose of feature classification is to describe entities as they occur in the world and not as they appear on a graphic representation. The lists of Features, Attributes and Attribute Values are not limited to any map series or scales.

3.1 Cartographic Feature Descriptive Model

Cartographic features shall be described by the following three categories: Feature, Attribute, and Attribute Value. These are defined as follows:

- o Feature - a defined entity of interest that is not further subdivided.
- o Attribute - a defined characteristic of a feature. The only mandatory attribute shall be location.
- o Attribute Value - a specific quality or quantity assigned to an attribute.

Two additional categories, Feature Class and Attribute Class are provided as user options. These are defined as follows:

- o Feature Class - a specified group of features (e.g., hydrographic, land use, transportation)
- o Attribute Class - a specified group of attributes (e.g., those describing measure, serviceability, composition, or structure)

3.2 Cartographic Feature Definitions

A comprehensive list of feature and attribute definitions is being prepared. Appendix 3A immediately following this page describes and lists a sample of the feature definitions. Appendix 3B in the Supporting Documentation contains a sample of the attributes. Maintenance of the standard list of features and attributes will be provided by a national body which will rule on all additions and changes to the standard.

3.3 Cartographic Feature Codes

The assignment of codes for the features and attributes will be made upon completion and review of the definitions. These codes shall not impose a structure upon the features, but are intended only for retrieval and maintenance. A preliminary version will be assigned for testing and the final form afterwards.

It should be emphasized that the following list is not complete with respect to features, nor to the attributes listed for each feature. It is for the purpose of illustration only, and has not been formally endorsed by the Working Group, the National Committee, nor the American Congress on Surveying and Mapping. Work on these and the remaining 700 feature terms is continuing, with scheduled completion of an interim standard set of generic feature terms by June, 1985. Also ongoing is the process of identifying a separate list of attributes and attribute defintions to be associated with each generic feature (see Supporting Documentation, Appendix 3B), and assigning of alphanumeric codes to each feature and attribute.

SAMPLE FEATURE DEFINITIONS

Acacia: see Scrub
Aerodrome Beacon: see Beacon
Aeronautical Beacon: see Beacon
Airport Beacon: see Beacon
Anabranch: see Stream
Anse: see Inlet
Aqueduct: see Artificial Watercourse
Arm: see Inlet
Arroyo: see Natural Watercourse
Artificial Harbor: see Harbor
Artificial Watercourse:
    Definition: A manmade or artificially improved waterway through which water may or does run
    Source: new term, no existing definition
    Attributes: location, name, width, depth, volume, length, relationship to ground level, composition, charted depth, covered, slope, shape, navigable, irrigation, drainage, water supply, commercial shipping, passenger transportation, water body connection, buoyed, lighted, waterage, recreation, flood control, hydroelectric power
    Includes: aqueduct, canal, channel, culvert, ditch, drain, draw, fairway, flume, lode, moat, overflow channel, seaway, ship canal, viaduct
Awawa: see Stream
Back Marsh: see Wetland
Backswamp: see Wetland
Bald: see Cleared Area
Bar Buoy: see Buoy
Barranca: see Natural Watercourse
Barrier:
    Definition: A fence or other obstacle to passage
    Source: Canadian Council on Surveying and Mapping
    Attributes: location, length, height, composition
    Includes: fence, gate, guard rail, hedge, hedgerow, wall
Barrier Flat: see Wetland
Bascale Bridge: see Bridge
Bay: see Inlet
Bayou: see Stream
Beacon:
    Definition: A fixed aid to navigation that emits a signal
    Source: Adapted from Navigation Dictionary
    Attributes: location, characteristic of signal
    Includes: aerodome beacon, airport beacon, aeronautical beacon, circular beacon, code beacon, continuous

radio beacon, day beacon, directional beacon, homing beacon, landmark beacon, lighted beacon, marine radiobeacon, obstruction beacon, omnidirectional beacon, pile beacon, post beacon, radar beacon, radar responder beacon, rotating beacon, rotating loop beacon

Beacon Buoy: see Buoy
Beck: see Stream
Bell Buoy: see Buoy
Bend: see Natural Watercourse
Blanket Bog: see Wetland
Bight: see Inlet
Boat Basin: see Harbor
Boat Harbor: see Harbor
Bog: see Wetland
Braided Stream: see Stream
Brake: see Woodland
Branch: see Stream
Bridge:

Definition: A structure erected over a depression or an obstacle such as a body of water, railroad, etc. to provide a roadway for vehicles or pedestrians

Source: Navigation Dictionary

Attributes: location, name, transportation mode, name of road or river crossed, composition

Includes: bascale bridge, causeway, culvert, draw bridge, foot bridge, lift bridge, navigating bridge, pontoon bridge, signal bridge, suspension bridge, swing bridge, trestle bridge, viaduct, weigh bridge

Brigalow: see Scrub
Brook: see Stream
Brush: see Woodland
Buoy:

Definition: A floating object, other than a light ship, moored or anchored to the bottom as an aid to navigation

Source: Navigation Dictionary

Attributes: location, shape, color, characteristic sound, characteristic light, characteristic radio signal, color pattern

Includes: bar buoy, beacon buoy, bell buoy, can buoy, cask buoy, checker buoy, combination buoy, conical buoy, danger buoy, dredging buoy, fairway buoy, fishnet buoy, horn buoy, ice buoy, junction buoy, keg buoy, lighted buoy, light sound buoy, mooring buoy, obstruction buoy, position buoy, quarantine buoy, radar buoy, radiobeacon buoy, sea buoy, sono buoy, sound buoy, spar buoy, special buoy, spoil ground buoy, swinging buoy, telegraph buoy, thermobuoy, topmark buoy, transo buoy, trumpet buoy, trunk buoy, warping buoy, whistle buoy, winter buoy, wreck buoy

Bush: see Scrub
Caatinga: see Woodland
Can Buoy: see Buoy
Canal: see Artificial Watercourse
Canal Port: see Port
Cask Buoy: see Buoy
Causeway: see Bridge
Channel: see Artificial Watercourse/Natural Watercourse
Chaparral: see Scrub
Checker Buoy: see Buoy
Cienega: see Wetland
Circular Beacon: see Beacon
City: see Municipality
Cleared Area:

Definition: An open area in a woodland
Source: Adapted from Canadian Council on Surveying and Mapping
Attributes: location, elevation, natural or manmade, area
Includes: bald, clearing, glade

Clearing: see Cleared Area
Code Beacon: see Beacon
Combination Buoy: see Buoy
Conical Buoy: see Buoy
Coniferous Forest: see Woodland
Coniferous Woodland: see Woodland
Continuous Radio Beacon: see Beacon
Copse: see Woodland
Cove: see Inlet
Creek: see Stream
Crop Land:

Definition: Land that has been plowed or otherwise cultivated for the planting of crops
Source: Canadian Council on Surveying and Mapping
Attributes: location, crop grown, growing patterns, area, growing season
Includes: cultivated area, cultivated field, farm, field, garden, market garden, orchard, paddy field, truck farm, truck garden, vineyard

Cultivated Area: see Crop Land
Cultivated Field: see Crop Land
Culvert: see Artificial Watercourse/Bridge
Cut Off: see Natural Watercourse
Danger Buoy: see Buoy
Day Beacon: see Beacon
Deciduous Forest: see Woodland
Deciduous Woodland: see Woodland

Dike: see Embankment
Directional Beacon: see Beacon
Dismal: see Wetland
Distributary: see Natural Watercourse
Dike: see Embankment
Directional Beacon: see Beacon
Dismal: see Wetland
Distributary: see Natural Watercourse
Dike: see Embankment
Directional Beacon: see Beacon
Dismal: see Wetland
Distributary: see Natural Watercourse
Dike: see Embankment
Directional Beacon: see Beacon
Dismal: see Wetland
Distributary: see Natural Watercourse
Ditch: see Artificial Watercourse
Down: see Grassland
Downland: see Grassland
Downs: see Grassland
Drain: see Artificial Watercourse
Draw: see Artificial Watercourse
Drawbridge: see Bridge
Dredging Buoy: see Buoy
Dyke: see Embankment
Embankment:

Definition: A raised structure of earth, ground, etc. used to hold back water

Source: Canadian Council on Surveying and Mapping

Attributes: location, length, height, composition, width

Includes: dike, dyke, levee, seawall

Entreport: see Port
Equatorial Forest: see Woodland
Equatorial Woodland: see Woodland
Everglade: see Wetland
Exposed Wreck: see Wreck
Fairway: see Artificial Watercourse/Natural Watercourse
Fairway Buoy: see Buoy
Farm: see Crop Land
Fen: see Wetland
Fence: see Barrier
Fenland: see Wetland
Field: see Crop Land
Fishing Harbor: see Harbor
Fishing Port: see Port
Fishnet Buoy: see Buoy
Floating Marsh: see Wetland
Flume: see Artificial Watercourse
Foot Bridge: see Bridge

Forest: see Woodland
Fork: see Stream
Garden: see Crop Land
Garigue: see Scrub
Gate: see Barrier
Glacial Stream: see Stream
Glade: see Cleared Area
Grass Field: see Grassland
Grassland:

Definition: An area covered mainly with grass

Source: Adapted from Canadian Council on Surveying and Mapping

Attributes: location, annual precipitation, area, predominant species, name

Includes: down, downland, downs, grass field, hay meadow, intermediate grassland, long grass prairie, meadow, midlatitude grassland, pampas, pasture, plain, prairie, puszta, range, savanna, short grass prairie, steppe, temperate grassland, tropical grassland, veld, veldt

Grove: see Woodland
Guard Rail: see Barrier
Guide Rail: see Barrier
Gulch: see Natural Watercourse
Gulf: see Inlet
Gully: see Natural Watercourse
Hamlet: see Municipality
Harbor:

Definition: Sheltered area of water where ships or other watercraft can anchor or dock

Source: Geographic Names Information System

Attributes: location, tidal or non-tidal, natural or artificially improved, name, type of vessel served, depth of water, size of area

Includes: artificial harbor, boat basin, boat harbor, fishing harbor, haven, inner harbor, island harbor, marina, natural harbor, refuge harbor, stranding harbor, tidal harbor

Haven: see Harbor
Hay Meadow: see Grassland
Hedge: see Barrier
Hedgerow: see Barrier
Homing Beacon: see Beacon
Horn Buoy: see Buoy
Hotspring: see Spring
Hulk: see Wreck
Ice Buoy: see Buoy
Inlet:

Definition: An opening of the sea into the land or of a lake into its shores

Source: Modified from a Dictionary of Geography, Monkhouse
Attributes: name, location, size, shape, width, depth, salinity
Includes: anse, arm, bay, bight, cove, gulf

Inner Harbor: see Harbor
Intermediate Grassland: see Grassland
Island Harbor: see Harbor
Junction Buoy: see Buoy
Jungle: see Woodland
Keg Buoy: see Buoy
Kill: see Stream
Landmark Beacon: see Beacon
Levee: see Embankment
Lift Bridge: see Bridge
Lighted Beacon: see Beacon
Lighted Buoy: see Buoy
Light Sound Buoy: see Buoy
Locality: see Municipality
Lode: see Artificial Watercourse
Long Grass Prairie: see Grassland
Malee Scrub: see Scrub
Mangrove: see Wetland
Mangrove Swamp: see Wetland
Maquis: see Scrub
Marais: see Wetland
Market Garden: see Crop Land
Marine Radiobeacon: see Beacon
Marsh: see Wetland
Mattress: see Revetment
Meadow: see Grassland
Meander: see Natural Watercourse
Midlatitude Grassland: see Grassland
Mineral Spring: see Spring
Mire: see Wetland
Moat: see Artificial Watercourse
Monsoon Forest: see Woodland
Mooring Buoy: see Buoy
Morass: see Wetland
Mott: see Woodland
Motte: see Woodland
Mulga: see Scrub
Mulga Scrub: see Scrub
Municipality:

Definition: A city, town or other district having powers of local self-government
Source: Canadian Council on Surveying and Mapping
Attributes: location, name, population, area
Includes: city, hamlet, locality, town, village

Muskeg: see Wetland

Narrows: see Natural Watercourse
Natural Harbor: see Harbor
Natural Watercourse:

Definition: A natural waterway through which water may or does run

Source: a new term, no source

Attributes: location, name, width, depth, volume, length, relationship to ground level

Includes: arroyo, barranca, bend, channel, cut off distributary, fairway, gulch, gully, meander, narrows, overflow channel pass, ravine, stream channel, wadi

Naval Port: see Port
Navigating Bridge: see Bridge
Obsequent Stream: see Stream
Obstruction Beacon: see Beacon
Obstruction Buoy: see Buoy
Omnidirectional Beacon: see Beacon
Orchard: see Crop Land
Outport: see Port
Overflow Channel: see Artificial Watercourse/Natural Watercourse
Paddy Field: see Crop Land
Palsa Bog: see Wetland
Pampas: see Grassland
Pass: see Natural Watercourse/Gap
Pasture: see Grassland
Peat Bog: see Wetland
Pile Beacon: see Beacon
Plain: see Grassland
Pocosin: see Wetland
Pontoon Bridge: see Bridge
Port:

Definition: A place provided with terminal and transfer facilities for loading and discharing cargo or passengers, usually located in a harbor

Source: Navigation Dictionary

Attributes: location, name, area, tidal or non-tidal, type of vessel served, water depth, salinity

Includes: canal port, entrepot, fishing port, naval port, outport, port of call, river port, seaport, terminal port

Port of Call: see Port
Post Beacon: see Beacon
Position Buoy: see Buoy
Prairie: see Grassland
Pup: see Stream
Puszta: see Grassland
Quaking Bog: see Wetland
Quagmire: see Wetland

Quarantine Buoy: see Buoy
Radar Beacon: see Beacon
Radar Buoy: see Buoy
Radar Responder Beacon: see Beacon
Radiobeacon Buoy: see Buoy
Raised Bog: see Wetland
Range: see Grassland
Ravine: see Natural Watercourse
Reforested Area: see Woodland
Refuge Harbor: see Harbor
Retaining Wall: see Revetment
Revetment:
Definition: A facing of stone, concrete, wood, etc. built to sustain an embankment
Source: Canadian Council on Surveying and Mapping
Attributes: location, length, construction material, height
Includes: mattress, retaining wall, riprap
Rio: see Stream
Riprap: see Revetment
River: see Stream
River Port: see Port
Rotating Beacon: see Beacon
Rotating Loop Beacon: see Beacon
Sagebrush: see Scrub
Salina: see Wetland
Salting: see Wetland
Salt Marsh: see Wetland
Salt Meadow: see Wetland
Savanna: see Grassland
Scrub:
Definition: A vegetation association in a semi-arid climate, or on poor sandy or stony soils, characterized by stunted trees, bushes, and brushwood. The scrub may be (I) Tropical and semi-desert type (Mulga, Spinifex, Chanaral, Acacia); (II) Warm temperature type (Maquis, Chaparral, Garigue, Malee, Brigalow, Sagebrush). The term is used loosely of any rough vegetation on heathland. The plants are mainly xerophilous in character, including cacti, throny aromatic shrubs, small gnarled evergreens, saltbushes, mesquite, creosite and sharp spiny grasses
Source: A Dictionary of Geography, Monkhouse
Attributes: location, area, predominant species, name, elevation
Includes: Acacia, brigalow, bush, chaparral, garigue, malee scrub, maquis, mulga, mulga scrub, sagebrush, spinifex
Sea Buoy: see Buoy

Seaport: see Port
Seawall: see Embankment
Seaway: see Artificial Watercourse
Seep: see Spring
Ship Canal: see Artificial Watercourse
Short Grass Prairie: see Grassland
Signal Bridge: see Bridge
Silva: see Woodland
Slash: see Wetland
Slough: see Wetland
Slue: see Wetland
Sono Buoy: see Buoy
Sound Buoy: see Buoy
Spar Buoy: see Buoy
Special Buoy: see Buoy
Spinifex: see Scrub
Spoil Ground Buoy: see Buoy
Spring:

Definition: A place where water issues from the ground naturally

Source: Modified from USGS

Attributes: location, name, force of flow, perennial/intermittent, temperature

Includes: hotspring, mineral spring, seep

Stand: see Woodland
Steppe: see Grassland
Stranding Harbor: see Harbor
Stream:

Definition: A natural body of water flowing on the land surface

Source: Modified from USGS

Attributes: location, name, relationship to ground level, width, depth, volume, length, perennial/intermittent, salinity, direction of flow, branch/parent, force of flow, tidal

Includes: anabranch, awawa, bayou, beck, braided stream, branch, brook, creek, fork, glacial stream, kill, obsequent stream, pup, rio, river, tideway, torrent

Stream Channel: see Natural Watercourse
String Bog: see Wetland
Submerged Ruins: see Wreck
Suspension Bridge: see Bridge
Swamp Forest: see Wetland
Swampland: see Wetland
Swing Bridge: see Bridge
Swinging Buoy: see Buoy
Taiga: see Woodland
Telegraph Buoy: see Buoy
Temperate Grassland: see Grassland

Terminal Port: see Port
Thermobuoy: see Buoy
Thickett: see Woodland
Thorn Forest: see Woodland
Tidal Flat: see Wetland
Tidal Harbor: see Harbor
Tidal Marsh: see Wetland
Tideway: see Stream
Topmark Buoy: see Buoy
Torrent: see Stream
Town: see Municipality
Transo Buoy: see Buoy
Tree Farm: see Woodland
Tree Plantation: see Woodland
Trestle Bridge: see Bridge
Tropical Grassland: see Grassland
Tropical Rain Forest: see Woodland
Truck Farm: see Crop Land
Truck Garden: see Crop Land
Trumpet Buoy: see Buoy
Trunk Buoy: see Buoy
Tulelands: see Wetland
Valley Bog: see Wetland
Veld: see Grassland
Veldt: see Grassland
Viaduct: see Artificial Watercourse/Bridge
Village: see Municipality
Vineyard: see Crop Land
Wadi: see Natural Watercourse
Wall: see Barrier
Warping Buoy: see Buoy
Watch Buoy: see Buoy
Weigh Bridge: see Bridge
Wetland:

Definition: A poorly drained land, fresh or saltwater, wooded or grass, possibly covered with open water

Source: Adapted from GNIS Documentation, AppendixB

Attributes: location, elevation, name, size of area, salinity, predominant plant species, tidal, seasonal depth of surface water

Includes: back marsh, backswamp, barrier flat, blanket bog, bog, cienega, dismal, everglade, fen, fenland, floating marsh, mangrove, mangrove swamp, marais, marsh, mire, morass, muskeg, palsa bog, peat bog, pocosin, quaking bog, quagmire, raised bog, salina, salting, salt marsh, salt meadow, slash, slough, slue, string bog, swampland, swamp forest, tidal flat, tidal marsh, tulelands, valley bog

Wooded Area: see Woodland

Woodland:

Definition: Land having a cover of trees and shrubs

Source: American Heritage Dictionaray

Attributes: location, elevation, area, predominant plant species, age of predominant growth, leaf type, evergreen/deciduous, percent tree cover, commercial/non-commercial, name

Includes: brake, brush, caatinga, coniferous forest, coniferous woodland, copse, deciduous forest, deciduous woodland, equatorial forest, equatorial rain forest, forest, jungle, grove, monsoon forest, mott, motte, reforested area, silva, stand, taiga, thickett, thorn forest, tree farm, tree plantation, tropical rain forest, wooded area

Wreck:

Definition: The ruined remains of a vessel which has been rendered useless, usually by violent action, as by the action of the sea and weather on a stranded or sunken vessel

Source: Modified from Navigation Dictionary

Attributes: location, size of wreck, relationship to water level

Includes: exposed wreck, hulk, submerged wreck

Wreck Buoy: see Buoy

Whistle Buoy: see Buoy

Winter Buoy: see Buoy

AN INTERIM PROPOSED STANDARD FOR TERMS AND DEFINITIONS

4.0 INTRODUCTION

A minimum requirement for digital cartographic data standards is a language which conveys completeness and understanding beyond any reasonable question. The emergence of digital cartographic applications technology, on the other hand, has appeared from a widely diverse array of institutions, governmental agencies and commercial firms, which do not necessarily talk to each other. The goal of the glossary is, therefore to bridge the communication gap by providing a standardized definition for a limited number of key terms which appear in these proposed standards.

The National Committee for Digital Cartographic Data Standards has focused considerable effort in selecting and defining appropriate terms in structuring these standards. The terms selection process was based primarily on the careful review of reports generated thus far by the NCDCDS leading to the Interim Proposed Standards.

Sections 4.1 to 4.3 contain definitions of fundamental cartographic objects. Section 4.4 contains terms and definitions used by the other three Working Groups in the specifications above.

4.1 DEFINITION OF 0-DIMENSIONAL CARTOGRAPHIC OBJECTS

4.1.1 • point - A 0-dimensional object that specifies geometric location. An optional set of coordinates specifes the location

4.1.2 * node - A 0-dimensional object that is a topological junction and specifies geometric location. A set of coordinates specifies the location.

4.2 DEFINITION OF 1-DIMENSIONAL CARTOGRAPHIC OBJECTS

4.2.1

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

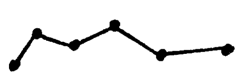
4.2.2

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

4.2.3

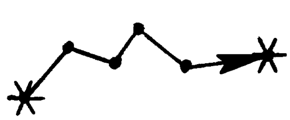
directed link - A link between two nodes with one direction specified.

4.2.4



string - A series of line segments strung together.

4.2.5



chain - A directed series of nonintersecting line segments strung together with nodes at each end of the string.

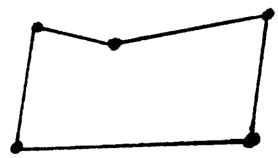
4.2.6

arc - A locus of points that forms a curve that is not closed.

4.3 DEFINITION OF 2-DIMENSIONAL CARTOGRAPHIC OBJECTS

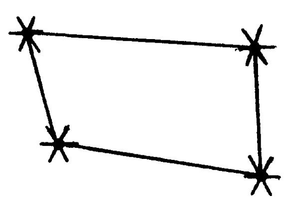
4.3.1 polygon - A 2-dimensional object that can be formed on a plane or other simple curved surface (e.g. sphere or ellipsoid) that can be formed in three ways:

4.3.1.1



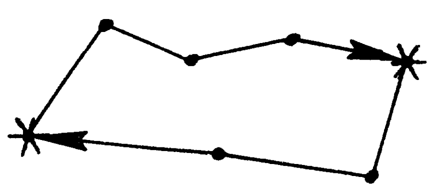
1) area bounded by a sequence of line segments with closure

4.3.1.2



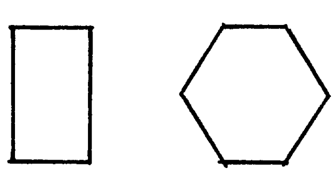
2) area bounded by a sequence of links or directed links with closure

4.3.1.3



3) area bounded by a chain(s) which has (have) closure

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



(paraphrased from the Manual of Remote Sensing)

4.4 GLOSSARY OF TERMS AND DEFINITIONS

4.4.1 accuracy - The closeness of results of observations, computations, or estimates to the true values or the values which are accepted as being true.

4.4.2 attribute - A defined characteristic of a feature. The only mandatory attribute shall be location.

4.4.3 attribute class - A specified group of attributes (e.g., those describing measure, serviceability, composition or structure).

4.4.4 attribute value - A specific quality or quantity assigned to an attribute.

4.4.5 audit trail - The path left by a transaction when it is processed. Record of transactions and transformations that permit the reconstruction of events in the lineage of a data base.

4.4.6 bias - Systematic distortion.

4.4.7 codes - A set of items, such as abbreviations or numbers, representing the members of another set.

4.4.8 completeness - Having all necessary or normal parts; containing all specified cartographic features. Every cartographic feature is contained in the data base; none are missing.

4.4.9 control point - Any station in a horizontal and vertical control system that is identified in the cartographic data and used for correlating the cartographic data with the horizontal and vertical control system.

4.4.10 data base management system - A special data processing system, or part of a data processing system, which aids in the storage, manipulation, reporting, management and control of data.

4.4.11 data definition language - A language that specifies the manner in which data is stored in a data base environment by a data base management system.

4.4.12 data interchange format - The procedures and/or rules used in the exchange of data between computer systems having different software and/or hardware.

4.4.13 data structure - A collection of data components that are constructed in a characteristic way.

4.4.14 datum - Any numerical or geometrical quantity or set of such quantities which may serve as a reference or base for other quantities.

4.4.15 digital cartographic analysis - The manipulation of cartographic data for the purpose of deriving information.

4.4.16 equivalence - Equality in value, significance, etc. As a property of a measurement scale; the ability to determine equality in value. As a property of a projection; the ability to represent areas in proportion to their true size.

4.4.17 error propagation - The accumulation of error from each operation or part of a system.

4.4.18 feature - A defined entity of interest that is not further subdivided.

4.4.19 feature class - A specified group of features (e.g.,hydrographic, land use, transportation).

4.4.20 geocoding - A system of abbreviation used in preparing geographic information for input to a computer, particularly political subdivisions.

4.4.21 image data - Digital data which form part of a digital model of an image.

4.4.22 interchange modeling - A conceptualization of the process of data interchange.

4.4.23 lineage - The record of the lineal descent and origin of the cartographic data. A narrative of the materials and procedures used to bring a cartographic data base to its current state.

4.4.24 location - Situation, position in space defined by a set of coordinates.

4.4.25 logical consistency - The degree to which cartographic features are accurately represented in the data structure and fulfill all the internal requirements of the data structure.

4.4.26 map scale - The ratio between a distance on a map and the corresponding distance on the earth.

4.4.27 map series - A group of map sheets having the same scale and cartographic specifications and collectively identified by the producer.

4.4.28 point (well-defined) - A position in the cartographic data that is well determined (for instance by a survey or can be plotted within the stated accuracy of the data).

4.4.29 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.

4.4.30 purity - A quantitative assessment of homogeneity or uniformity, particularly with respect to attribute accuracy.

4.4.31 quality - An essential or distinguishing characteristic necessary for the cartographic data to be fit-for-use.

4.4.32 reliability diagram - A diagram included in the cartographic data used to portray quality information about the data such as horizontal and vertical accuracies or zones sharing the same lineage.

4.4.33 resolution - The smallest unit which can be detected. Resolution provides a limit to precision and accuracy.

4.4.34 slivers - The misalignment of two cartographic features in the digital cartographic data when the two cartographic features occupy the same geographic location.

4.4.35 spatial relationships - An association of two or more elements in some dimension of space.

4.4.36 tessellation - A repeating pattern of either regular or irregular shapes.

4.4.37 topological data - data which is invariant under geometrical deformations and stretchings.

4.4.38 transformation - The computational process of converting a position from one coordinate system to another.

4.4.39 truth-in-labeling - The cartographic data must conform to the specifications in the label.

4.4.40 vector - A line segment.

SUPPORTING DOCUMENTATION

(p. 45 follows)

BACKGROUND AND ORGANIZATION FOR THE DEVELOPMENT OF AN INTERIM PROPOSED STANDARD: SUPPORTING DOCUMENTATION

by

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

0.0.1 INTRODUCTION

In the last decade the rise of computer-assisted processing in cartography and the use of cartographic databases have grown in a very dramatic way. In the early years agencies in the governmental sector and organizations in the private sector used these approaches on a largely experimental basis. Today one sees the same organizations and agencies using numerical processing and data bases on an ongoing day to day basis. This growth and development in the field of cartography is nothing short of remarkable. However this growth has been uneven and to a large extent not well coordinated in the civilian sector of cartography. It was recognized several years ago that if continued growth in the use of numerical methods and databases was to be sustained over a longer time period, that it would be necessary to develop standards for digital cartographic databases.

Each year millions and millions of dollars are being spent to reorganize, reformat, process, verify and check digital cartographic databases that one agency or organization obtains from another. In most cases this is a very time consuming process. If a comprehensive set of national digital cartographic data standards can be developed, then the entire field of cartography will benefit from this effort and at the same time save large amounts of time and money. To this end the National Committee for Digital Cartographic Data Standards has been established under the auspices of the American Congress on Surveying and Mapping, with the support of the U.S. Geological Survey.

This chapter reviews the background and organization of the Committee that has developed during the last three years of operation. The first year was spent defining the issues associated with developing such standards, the second year involved examining the alternatives available, and the third year has seen the development of an Interim Proposed Standard. Other sections in this report present and discuss the details of this Interim Proposed Standard.

0.0.2 RECOGNITION OF A NEED FOR COORDINATION AND STANDARDS

Explicit recognition of the need for coordination of effort in digital cartography and the need for standards has existed for at least a decade. In a 1973 report the Federal Mapping Task Force recommended a greater coordination of cartographic activities in the Federal sector of the profession and also recommended the establishment of a national digital cartographic database (Federal Mapping Task Force, 1973). More recently the National Research Council, an arm of the National Academy of Sciences, in 1980 in their recommendation for a multipurpose cadastre stated that:

> We recommend that technical studies continue to be sponsored by the federal government to identify consistent land information and display standards for use among and within federal agencies and between federal and state governments. These studies should rely on the authority of state governments to adopt the standards and organize the data collection in cooperation with the federal government to ensure compatibility on a national basis, delegating these functions to local governments where appropriate.

In 1981 a different National Research Council panel reviewed the original Federal Mapping Task Force recommendation and subsequently stated that:

> We recommend that the mapping, charting, geodesy, surveying and cadastral agencies of the federal government continue to sponsor cooperative programs, with state and local governments providing sufficient guidance to ensure conformance to national specifications and standards and thus to development of a fully integrated national information system.

and further stated that:

> We recommend that the geodetic and cartographic data bases be adequately supported, be readily accessible to all users, and even though serving different interests and needs, be made integral parts of a national mapping, charting, geodesy, surveying, and multipurpose cadastre information system.

A further statement on the procedures and standards for a multipurpose cadastre by yet another National Research Council (1983) stated that a lead agency be designated to:

> ...provide a structure for the formal recognition of procedures and standards for a multipurpose cadastre, as described above, and to oversee compliance with them by the federal establishment.

Other fiscal agencies in the Federal government have also recognized the need for more coordination of digital cartography and standards because of the fiscal efficiency that can be gained by better coordination of these activities. A recent General Accounting Office report (1982) recognized the duplicative nature of current computer-mapping programs in the Federal sector with a finding that in 1981 over a dozen major agencies spent in excess of $45 million on various kinds of digital mapping programs. This figure does not include any traditional hand cartography. After a thorough examination of the problem the General Accounting Office recommended that:

> ...the Director, OMB, issue a circular or other directive requiring the interagency coordination of computer mapping and preventing duplicative programs. The directive should create a rulemaking body to establish uniform standards for Federal computer mapping so that agencies can exchange data and the needs of map users are met at reasonable cost.

Their explicit recognition is intended to gain more efficiency in the use of government funds by achieving more cooperation and less overlap between agencies in this area. However one very direct implication from this recommendation is that cartographic standards must be developed to facilitate such cooperation. Just about the time that this OMB report was written, the Office of Management and Budget in April of 1983 (OMB,1983) issued a memorandum to coordinate the Federal digital cartographic program by establishing a Federal Interagency Coordinating Committee on Digital Cartography to oversee the process. The part of that mandate that is of interest here is that the group is charged to:

> Develop and adopt, for use by all Federal agencies, common Standards of content, format, accuracy for digital cartographic base data to increase interchangeability and enhance its potential for multiple use.

If there ever was any doubt that there is a real need to establish digital cartographic data standards that doubt has been dispelled once and for all.

Mandate for the National Committee for Digital Cartographic Data Standards

Although the tempo of recognition for the need for a national committee to recommend digital cartographic data standards has risen markedly rather recently, action to address the challenge directly has been going forward for several years. In 1980 a Memorandum of Understanding (MOU) was negotiated between the National Bureau of Standards and the U.S. Geological Survey which designated the Survey with the lead responsibility for developing, defining and maintaining data elements, and standards for earth science information systems. It should be recognized that digital cartographic standards are only one part of this overall mandate. Naturally the responsibility for the cartographic aspects of this mandate fell to the National Mapping Division of the Survey.

The National Mapping Division recognized early on that if a set of standards was to be developed that would really gain acceptance in the field of cartography, that not only the Federal agencies concerned with the result, but also the state and local agencies, the private sector and the research sector, mainly universities, must participate in this process. The National Mapping Division in all of its wisdom acted upon a recommendation from Moellering (1981) and encouraged the founding and organization of the National Committee for Digital Cartographic Data Standards to develop the standards for digital cartographic data. The Committee operates under the umbrella of the American Congress on Surveying and Mapping of which the American Cartographic Association is a member organization in the United States. The beauty of this arrangement is that the Committee can operate as an impartial and independent body to develop the needed digital cartographic standards in a setting that includes all segments of the profession.

0.0.3 ORGANIZATION AND WORK OF THE NATIONAL COMMITTEE

As a result of the establishment of the National Committee its primary goal has been defined as (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.

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

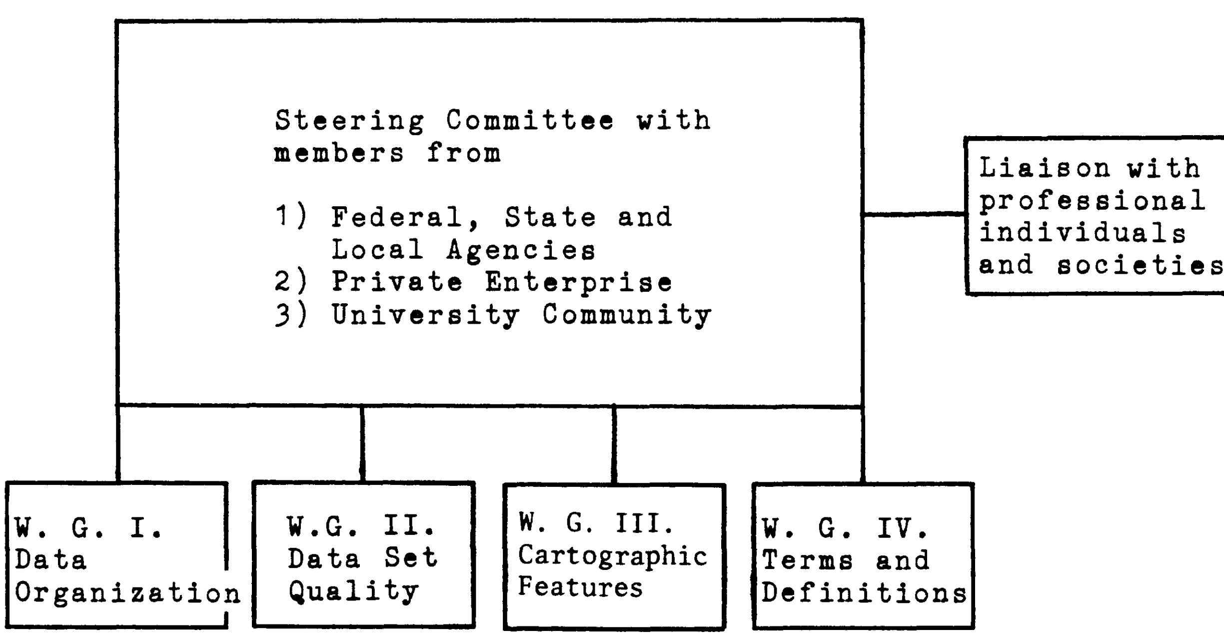


FIGURE 1. ORGANIZATION OF THE NATIONAL COMMITTEE FOR DIGITAL CARTOGRAPHIC DATA STANDARDS

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 listed in Section 0.1.

0.0.4 CONCEPTUAL CONSIDERATIONS

At the outset one must explore several theoretical issues which will provide the conceptual basis for developing the discussion pertaining to such standards. These theoretical issues are real and virtual maps, cartographic data levels, and the surface and deep structure of cartographic information.

Real and virtual maps

In recent years a number of cartographic products have been developed which have many of the characteristics of conventional maps, but are fundamentally different. For example, an image on a CRT display can look very much like a conventional map, but yet it is highly transient and can disappear with the push of a button. Data stored in computer memory can be easily converted into a map by plotting or CRT display, but it does not fit the conventional definition of a map. Moellering (1980, 1984) solved this conceptual problem by developing the notions of real and virtual maps. Two fundamental attributes which can be used to distinguish between classes of real and virtual maps are: 1) permanent tangible reality and 2) direct viewability. Figure 2 illustrates the situation by developing a four-class definition based on these two attributes. For example, a conventional real map has a permanent tangible reality and can be rolled up and carried around. It is also directly viewable in that the information in it can be directly read by a human viewer. Any cartographic product which lacks one or both of these two attributes is called a virtual map. In contrast, a CRT image is directly viewable, but it is highly transient, has no permanent reality, and therefore is a virtual map type I, meaning that it has many map-like characteristics, but is not a true real map. The data from which the CRT image is generated has neither a permanent tangible reality nor is it directly viewable. However it is directly convertable into a real map or into a CRT image. This is defined as a virtual map type III. A type II virtual map has a permanent tangible reality, but it is not directly viewable and examples of such maps may be as sophisticated as a laser data disk or as simple as a set of field data. Most of the type II maps can be converted into type III virtual maps. One should note that in all cases that virtual maps can be converted into real maps.

There are twelve different kinds of transformations between real and virtual maps as shown in figure 3. For cartographers and other spatial scientists those transformations define most operations on map data and can be expressed as transformations between map states t(s1→ S2). Therefore the task of digitizing cartographic information is t(R→V3) while entering numerical data tables associated with the map is t(V2→V3). Since most interactive analysis systems include a CRT map display and usually the capability for creating a hard copy map, transformations t(V3→R), t(V3→V1), t(V1→V3) and t(V1→R) are involved. In fact such a cartographic analysis and design system can be schematically depicted in terms of these transformations as shown in figure 4. Such a system has a raw data input with digitizing and data entry transformations. Notice that once inside the interactive system the information is always in a V1 or V3 virtual state which illustrated the highly manipulable nature of these two map states. Ease of manipulation is a primary advantage of these two states in such interactive systems. Since such a system uses man-machine interaction, the map display is read and interpreted by the operator of the CRT terminal which involves a transformation

DIRECTLY VIEWABLE AS A CARTOGRAPHIC IMAGE

PERMANENT TANGIBLE REALITY	YES	NO
YES	<u>REAL MAP</u> Conventional Sheet Map Globe Orthophoto Map Machine Drawn Map Computer Output Microfilm Block Diagram Plastic Relief Model	<u>VIRTUAL MAP-TYPE 2</u> Traditional Field Data Gazeteer Anaglyph Film Animation Hologram(stored) Fourier Transform(stored) Laser Disk Data
NO	<u>VIRTUAL MAP-TYPE 1</u> CRT Map Image a) refresh b) storage tube c) plasma panel Cognitive Map (two-dimensional image)	<u>VIRTUAL MAP-TYPE 3</u> Digital Memory(data) Magnetic Disk or Tape(data) Video Animation Digital Terrain Model Cognitive Map (relational geographic information)

FIGURE 2. THE FOUR CLASSES OF REAL AND VIRTUAL MAPS

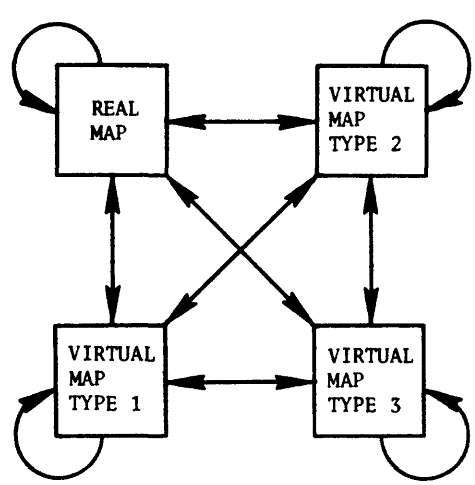


FIGURE 3. TRANSFORMATIONS BETWEEN REAL AND VIRTUAL MAPS

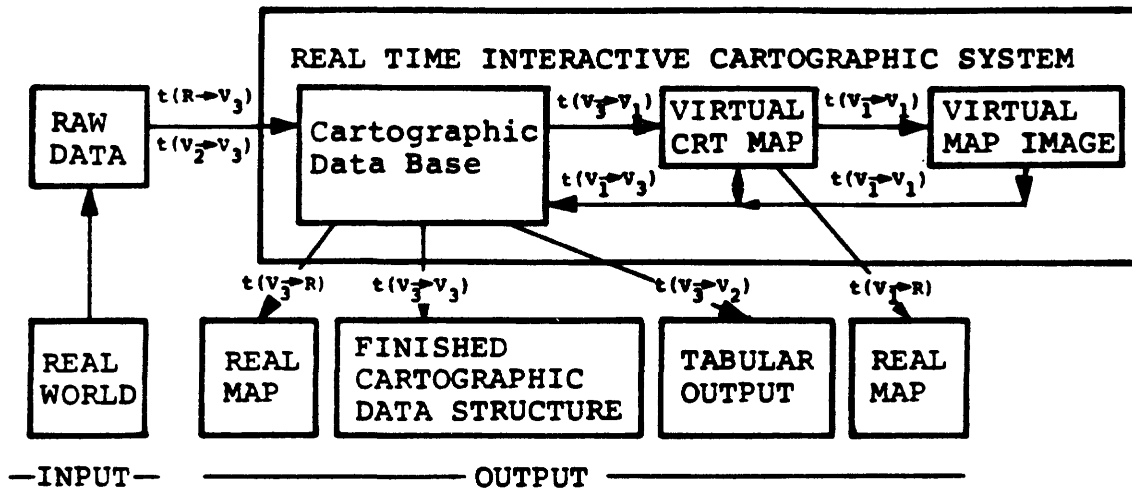


FIGURE 4. SIMPLIFIED DIAGRAM OF A REAL TIME INTERACTIVE CARTOGRAPHIC SYSTEM USING THE CONCEPT OF REAL AND VIRTUAL MAPS

from the CRT image to map image in the mind, t(V1→V1), a cognitive map. Inside of this interactive CRT system is where data processing, analysis and map display occur. At some point a solution to the problem is reached which needs to be documented or preserved. At that point hard copy output is generated in the form of real maps, t(V3→R), and other output. The advantage of this sort of transformational view of cartographic information is that it can be used as an aid in the conceptual design of such cartographic data analysis and display systems, as well as aid in the conceptual understanding of the nature of maps.

Cartographic data levels

Any computer-based system which analyzes spatial or cartographic data must have some way in which that data is organized, manipulated and subsequently managed. Most work on cartographic data structures has been directed towards specific implementations of data structures for specific systems. More recent work has moved in a direction of more generalized data structures as illustrated by the work of Peucker and Chrisman (1975), Haralick and Shapiro (1979) and the symposium organized by Schmidt (1977). Some of this work has moved in the direction of identifying fundamental topological aspects of cartographic data structure for specific data domains. Nyerges (1980) has identified six specific levels of cartographic data organization, which puts the data structure problem into perspective:

1) Data reality - the real world and Data pertaining to it concerning cartographic entities and relationships between them.

2) Information structure - a formal model that specifies the organization of information pertaining to a specific phenomenon. It includes data classes and relationships between them and acts as a skeleton for the canonical structure.

3) Canonical structure - a data model representing the inherent structure of a data set which is independent of specific applications and systems which manage such data.

4) Data structure - a logical data organization designed for a particular system in which specific relationships and links are implemented.

5) Storage structure - a specification of how a particular data structure is stored in data records in a particular system.

6) Machine encoding - the physical representation of how the structure is held in the physical devices of computer system hardware.

When one discusses the question of cartographic data structure, relationships or attributes, there must always be an awareness of these levels of cartographic information and which specific level is being addressed. An explicit recognition of these levels provides a clearer conceptual understanding of the specificity or generality of information being examined, and permits a clearer elucidation of the cartographic relationships being captured in the data structure.

Surface and deep structure

When one has a real map which presents a graphic image of some part of the real world, that graphic image is in essence an iconic replica of that segment of the spatial domain being portrayed. As such this real map provides the image which contains many explicit and implicit spatial relationships of interest to cartographers, geographers, and other spatial scientists. The graphic image is known as the surface structure, while spatial relationships are known as the deep structure of the map (Nyerges, 1980). When one converts this cartographic information into the digital data domain, it is clear that the concepts of deep and surface structure also apply as shown in figure 5. Early cartographic data structures were designed mainly for display purposes and were almost exclusively surface structure representations of cartographic reality. The situation frustrated many analytical cartographers and geographers because

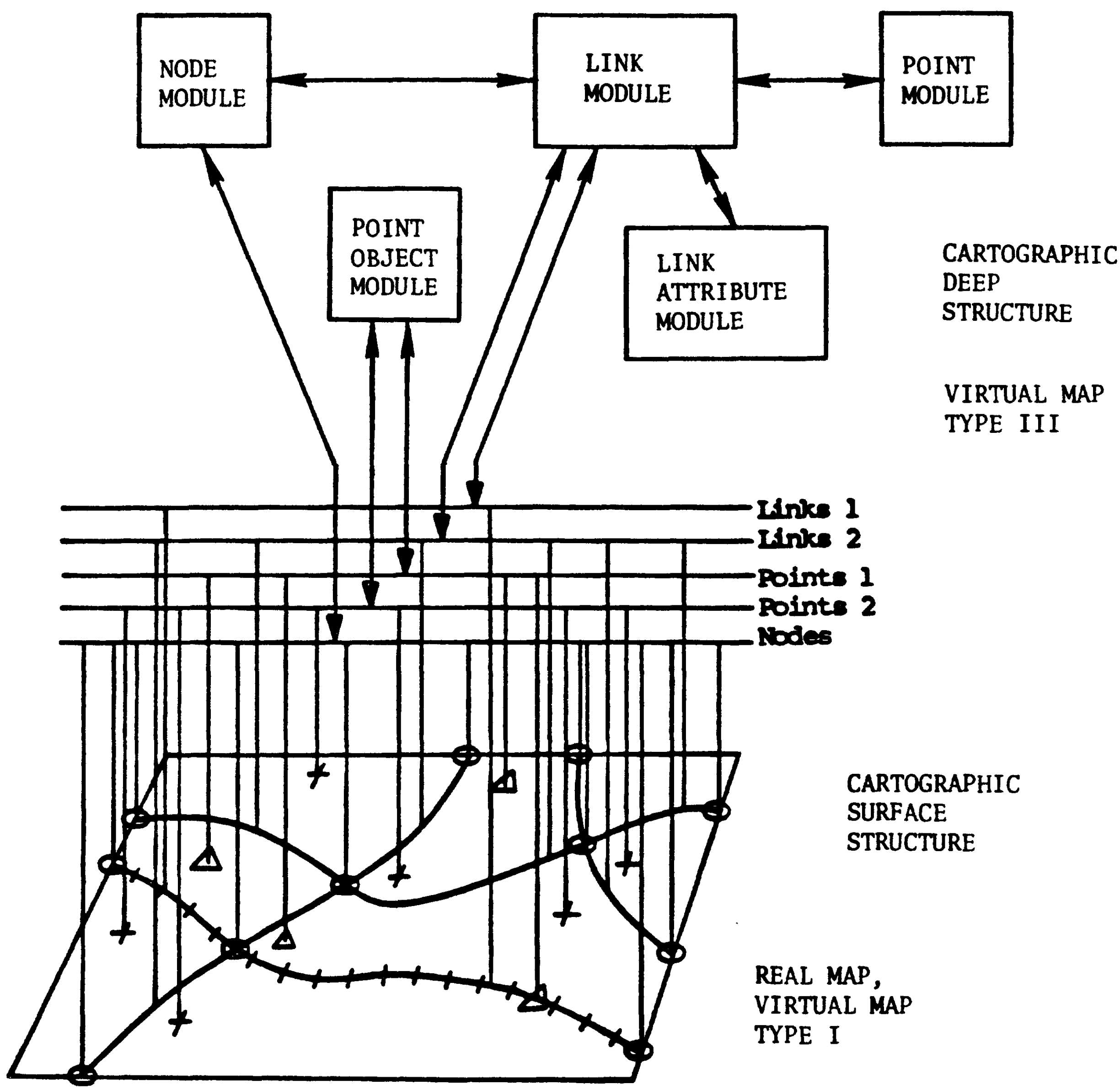


FIGURE 5. THE RELATIONSHIP BETWEEN DEEP AND SURFACE CARTOGRAPHIC STRUCTURE

such data structures could not conveniently be used for analytical purposes. The reason, of course, is that deep structure relationships necessary for spatial analysis were not preserved. Most modern cartographic data structures preserve some proportions of the deep and surface structures. Therefore the primary purpose of the use of a particular cartographic data structure will influence the proportions of deep and surface structure present in that implementation. It is interesting to note that there is a striking relationship between deep/ surface structure and real/virtual maps. It turns out that for the most part that surface structure representations are usually converted into real maps or virtual type I maps as CRT inages whereas deep cartographic structure is usually a type III virtual map and sometimes a type II virtual map. It is the combination of these two fundamental concepts along with an awareness of cartographic data levels that adds a critical insight into the problem of cartographic information organization in general and cartographic data structure specifically.

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 bringing 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.

0.0.4 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.

Three stages of the work have concerned the specifying of the issues and gaps in our knowledge, specifying possible alternative solutions for standards, and defining an Interim Proposed Standard. The general tasks for the WGs have been 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. Examine fidelity of graphical data, metric and topological
2. Examine coding reliability
3. Examine update and other temporal information
4. Examine linage of a data set
5. Examine checking procedures used by the producer to verify quality

III. Working Group on Cartographic Features

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

IV. Working Group on Terms and Definitions

1. Collect new terms defined by working groups
2. Define cartographic objects

More recent work by the WGs has refined these tasks and subsequently produced an Interim Proposed Standard. This work is briefely summarized below.

Working Group on Data Organization

The scope and goals of WG I have been 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 large data bases. The WG has identified terminology and definitions of terms currently being used in the area.

The proposed solution is to specify a superstructure that is widely machine readable which will contain a small number of exchange modules (defined formats) that can handle most kinds of cartographic data structures. These exchange modules are independent of one another, and therefore new data structures can be added to the system as the situation merits. The corpus of the superstructure and a few test exchange modules are presented in Section 1 of the Interim Proposed 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 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 is very informative to the user and indeed by very helpful in deciding whether a particular data set could successfully be used for a particular purpose.

The Working Group has developed the specifications for a quality report that should be provided to the potential user of a database. The five categories of information to be provided the user are: lineage, positional accuracy, attribute accuracy, logical consistency and completeness. These are discussed in Section 2 of the Interim Proposed Standards.

Working Group on Cartographic Features

The fundamental challenge of the Group has been to harmonize the feature definitions and coding schemes used by the major agencies in cartography. The recent work of the group indicates that such a scheme should 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. Recent work has involved gathering together a comprehensive list of such definitions and subsequently make a choice of the preferable definition that produces the best coverage for the entire set of definitions chosen for the Interim Proposed Standard, Section 3.

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. 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 been devised to handle the terms generated by the other Working Groups and a method for processing comments concerning the definition of these terms has been 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 completed. Recent work has involved making 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. This work is presented in Section 4 of The Interim Proposed Standard.

0.0.6 CYCLES OF DIGITAL CARTOGRAPHIC STANDARDS DEVELOPMENT

In order to develop effective digital cartographic standards the most efficient approach is to follow the strategy that is the same as one would the solving of any other scientific problem. Therefore one begins with the general considerations and progressively works down to the specific detailed problems and then back up to the general problem. As a result the solution process has been conceptualized into five basic cycles of work:

1) Define the fundamental issues involved
2) Define the alternatives to the problem
3) Formulate interim proposed standards
4) Test interim proposed standards
5) Generate final proposed standards.

At the end of each cycle a report has been written by the committee and circulated to the progession for thought, reflection and comment. Comments received by the Committee from concerned professionals have been integrated into the process quickly. It should be fairly clear that this incremental process is designed to minimize contrasting opinions at the end by integrating comments quickly into the process,

If one is to achieve an efficient solution to the problem of digital cartographic standards, one must begin by thoroughly addressing the issues. These issues were presented and discussed in Committee Report No. 3. Then follows an examination of the available alternatives to the solution of the problem. The tradeoffs must be weighed for each realistic alternative. These alternatives have been discussed in Report No. 4. This current report presents an Interim Proposed Standard which will be held as part of the AUTO-CARTO 7 meetings in Washington, D.C. in March, 1985. The review of the comments received from the profession will wrap up the work of cycle 3 in this standards process. The Spring of 1985 will see the beginning of cycle 4, the testing of the Interim Proposed Standard. It is clear that empirical testing is necessary in order to ascertain the efficiency of the Interim Proposed Standard. The results of these tests will be evaluated later in the cycle and any necessary adjustments will be made at that time.

0.0.7 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). 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 NDCDDS is concerned with cartographic activities in the entire profession at a more general scale, there are several areas of common interest. These areas have been explored and methods of cooperation and coordination between the two committees have been implemented. Fruitful results are expected.

0.0.8 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 6 shows a schematic diagram

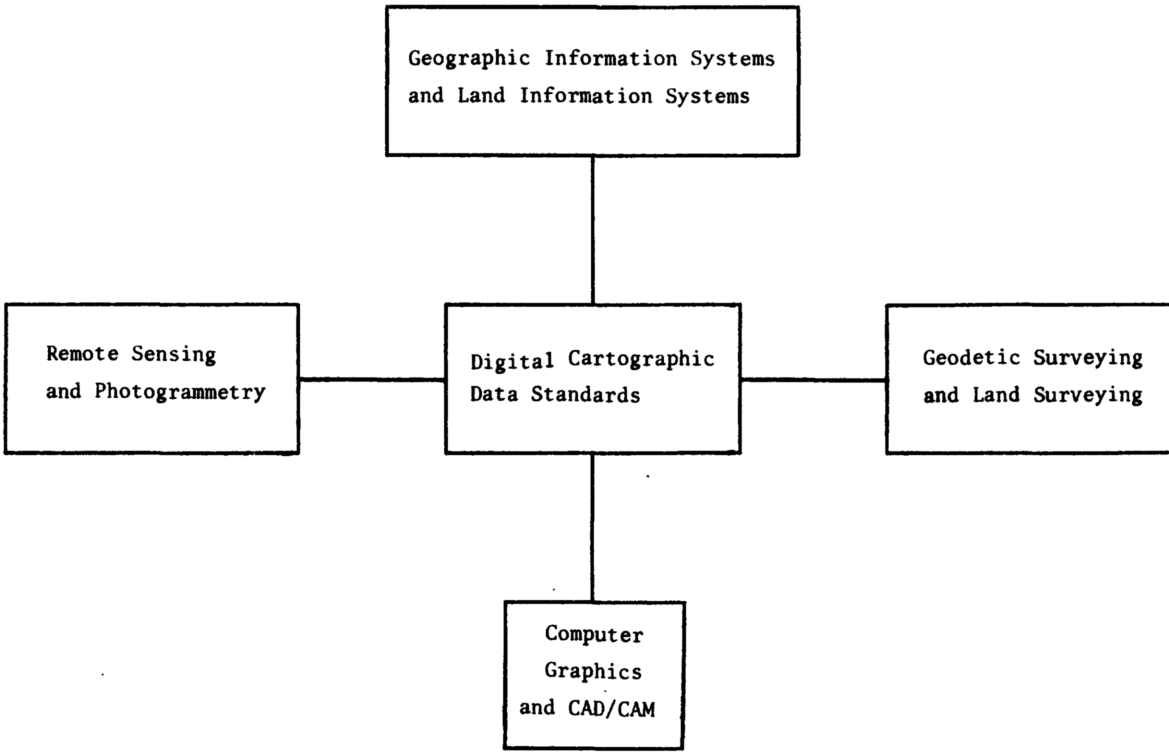


FIGURE 6. COGNATE AREAS WHICH INTERFACE TO DIGITAL CARTOGRAPHIC DATA STANDARDS

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 geen 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.

0.0.9 FUTURE WORK

The Committee is now in the process of finishing up this third cycle of defining an Interim Proposed Standard. This report presents that set of standards for the profession. All members of the cartographic community are invited to participate in this process by sending their written comments to the Committee. Standard comment forms may be 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 public hearing is being planned for the 1984 AUTO-CARTO 7 meetings in Washing, D.C. All members of the cartographic community are invited to participate in this hearing.

0.0.10 SUMMARY AND CONCLUSIONS

The Committee is now in the final stages of the cycle of defining the Interim Proposed Standard, with only the public hearing and comments remaining. The Committee will then move to the fourth cycle, that of testing the Interim Proposed Standard.

0.0.11 REFERENCES

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

Federal Mapping Task Force, 1973, Report of the Federal Mapping Task Force on Mapping, Charting, Geodesy and Surveying, Office of Management and Budget, Washington, U.S. Government Printing Office Report #2401-02426.

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.

Haralick, R., and L. Shapiro, 1979, "A Data Structure for a Spatial Information System", Proceedings of the International Symposium on Computer-Assisted Cartography, Vol II, pp. 291-303.

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

__________________, 1981, A Proposed Approach to Establish a Set of National Digital Cartographic Data Standards, Report to the National Mapping Division, U.S. Geological Survey.

__________________, 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 Meetings, Denver, pp. 201-212.

__________________, 1984, "Real Maps, Virtual Maps and Interactive Cartography", in G. Gaile and C. Willmott eds., Spatial Statistics and Models, Dordrecht, Holland: Diheidel Publishing Co., pp. 109-132.

National Research Council, 1980, Need for a Multipurpose Cadastre, Washington: National Academy Press.

__________________, 1981, Federal Surveying and Mapping: An Organizational Review, Washington: National Academy Press.

__________________, 1983, Procedures and Standards for a Multipurpose Cadastre, Washington: National Academy Press.

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.

Peucker, T., and N. Chrisman, 1975, "Cartographic Data Structures", American Cartographer, 2(1), pp. 55-59.

Schmidt, A., 1977, Papers of the First International Advanced Study Symposium on Topological Data Structures for Geographical Information Systems, Laboratory for Computer Graphics, Harvard University.

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

(p. 67 follows)

AN INTERIM PROPOSED STANDARD FOR DIGITAL CARTOGRAPHIC DATA ORGANIZATION: SUPPORTING DOCUMENTATION

Edited by: Timothy Nyerges

Working Group I on Data Organization:

Timothy Nyerges (Chair)	Northwest Cartography, Inc.
Donna Peuquet (Vice-Chair)	Univ. of California, Santa Barbara
Frederick Billingsley	Jet Propulsion Laboratory
A. Raymond Boyle	University of Saskatchewan
Hugh Calkins	State University of N. Y., Buffalo (Working Group IV Representative)
Robert Edwards	Oakridge National Laboratory
Robin Feagas	U. S. Geological Survey
William Liles	Xerox Corporation
David Pendelton	National Ocean Service
Daniel Rusco	Defense Mapping Agency

1.0 INTRODUCTION

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

Review of Cycle 1 and Cycle 2

During Cycle 1 three major issues in data organization were identified: terminology, modeling and data interchange. A decision was made to identify common as well as confusing terminology and to pass these terms on to the Working Group IV Terminology for clarification.

The modeling issue was discussed in regard to models of reality vs. models of maps vs. data models. This discussion carried over into Cycle 2. Working Group I decided not to try and standardize a single definition of a model for general cartographic use as this would "add concrete to confusion" in digital cartography. Models of reality can be defined as the salient charateristics of reality which are to be stored in a data base. A model of a map can be defined as the salient elements that appear on a map which are stored in a data base. A data model is a general description of specific sets of entities and the relatinships among the entities. An entity is a thing which exists and is distinguishable, i.e. we can determine one entity from another (Ullman, 1982). A data model is implemented using a data language. A data language is a set of procedures for defining and manipulating a data base (Kunii, 1983). Although the above definitions indicate the manner of use in the discussions of Working Group I, literature utilizing these terms is somewhat mixed and confusing.

Report 4 contained the development of a conceptual model for a digital cartographic data interchange process. A addressing a general problem this conceptual model includes the major factors involved with interchanging non-spatial data bases among diverse systems, including differences between:

1) source and target hardware environment,

2) structural organizations of data, or schemas, used by source and target systems,

3) source and target data base management software systems, and

4) data models upon which the source and target data base management systems (DBMS) are based, e.g. hierarchical, network and relational, etc.

The conceptual model was then extended to provide for the additional major factor unique to many spatial data systems:

5) the differences between methods of representing the topological relationships inherent in data of interest to spatial data analysts. In many cases, the topological structure is of primary importance.

The conceptual model has provided a useful framework for the analysis of the digital cartographic data interchange process in general, and focused efforts toward the definition and evaluation of interchange formats in particular. It has also provided a technical basis for the scope and extent of the complexities inherent in the exchange of this type of digital data.

Figure 1.1 is a diagram of the conceptual model. At the upper left is 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 spatial data. This source data base is to be transported from its resident hardware configuration to become the target data base under a different DBMS. The target DBMS is assumed to have a different data model. The target DBMS is resident on a different hardware architecture, and the transfer should result in no loss of data content or distortion of cartographic meaning.

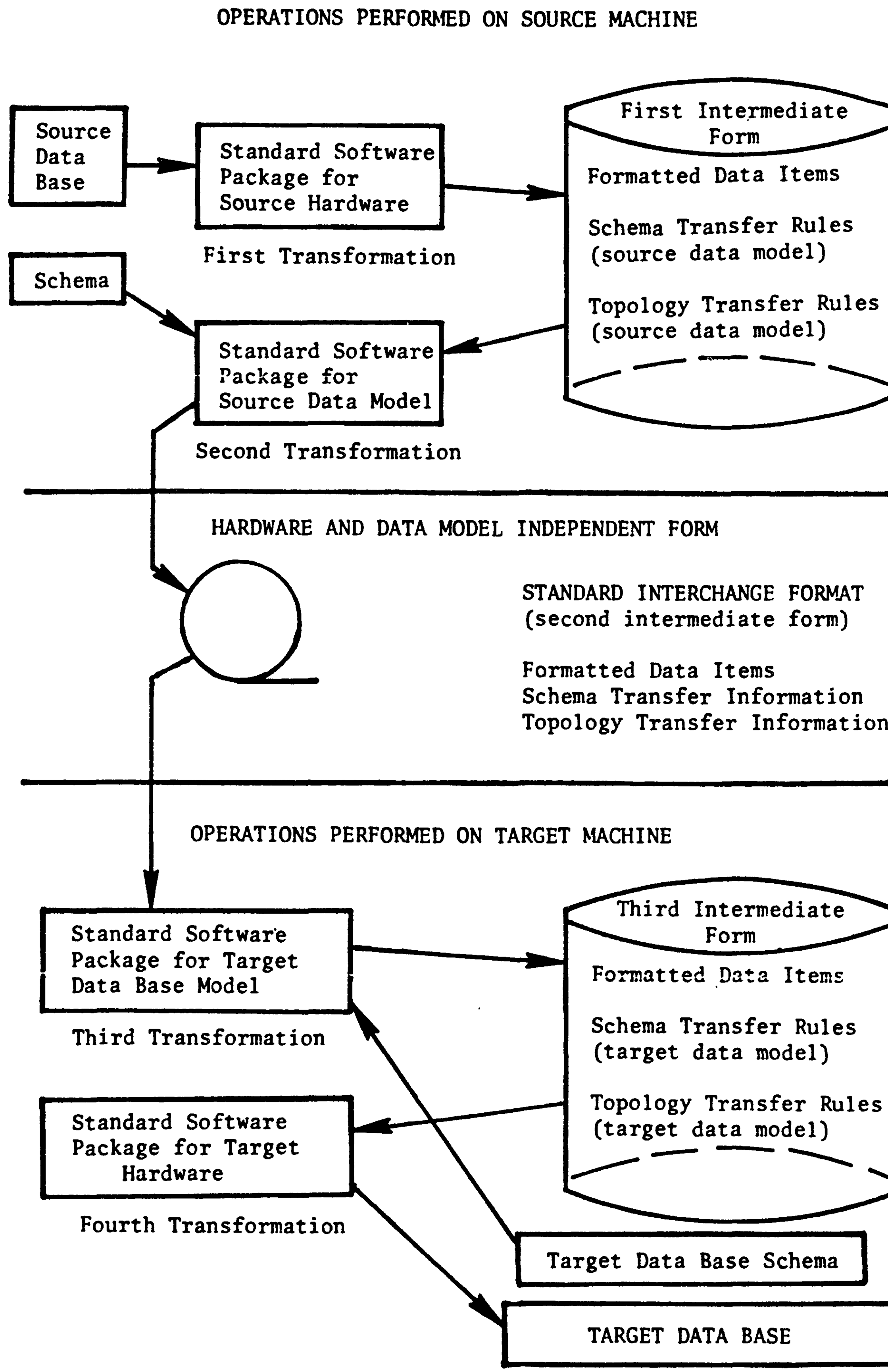


Figure 1.1 Digital Cartographic Data Interchange Process

The first and second transformations are implemented on the source machine in the form of standard software packages produced by the hardware and/or data base management system vendors. Notice that the second transformation is driven by the source data base schema, while the first transformation contains the source hardware dependent information and data model information, even if this information is proprietary.

The second intermediate form is the data base in standard interchange format and is independent of hardware and data model. The form is transferrable to any other system with third and fourth transformation software packages implemented for completing the data base transfer to the hardware architecture and data base management system of that target system.

The third and fourth transformations are implemented on the target machine in the form of standard software packages provided by the hardware and/or DBMS vendors. Notice that the third transformation is driven by the target data base schema, and the fourth transformation contains the target hardware and data model information, even if this information is proprietary.

The interchange process is completed at the lower right of the diagram with the transported target data base and/or set of files consisting of the data and topological information originally from the source data base. The data base interchange process is completed after the target data base has been brought up under the target data model and is successfully operated upon by the operators of that model.

As a result of this conceptual model of the interchange process, including the spatial data models, transformations, and forms involved, the following general conclusions can be reached:

1) Spatial data models and their implementaions as data base management systems must further evolve to a state suitable to allow the definition of a standard specification for each. This will provide a basis for the implementation and marketing of such systems by commercial vendors.

2) Commercially available data base management systems must be developed for the unique needs of spatial data management. Existing data base managements are not generally satisfactory for this type of data due to the need to effectively represent the multitude of topological relationships.

3) Standard intermediate forms for data and structure interchange must be developed.

4) A standard interchange format, or second intermediate form, capable of containing a range of complex relationships unique to spatial data must be developed and standardized,

5) The development and wide availability of standard software packages implementing the four transformations for standard data models across a variety of vendor supplied hardware and data base management systems would complete the move to generalized spatial data interchange, and

6) The general problem of data interchange among different data bases includes the networking of data bases. Gligor and Luckenbaugh, (Gligor and Luckenbaugh, 1984) present several good arguments for using structured protocols at the applications level to facilitate the interconncecting of heterogeneous data bases. These structured data transfer protocols are based on the International Standards Organization open network protocol standard (ISO/TC97/SC16, 1980).

Cycle 2 discussions resulted in identifying three alternatives for data interchange: 1) accept an existing application or organization specific format, 2) utilize a file definition approach that will allow an organization to describe the structure and data internally to the transmission similar to that proposed in ISO-DIS-8211 (ISO 8211, 1983), or 3) develop a new fixed format that would satisfy most cartographic applications.

For the purpose of generality and applicability across different organizations, the alternatives were narrowed to the last two as being the alternatives to pursue. These two alternatives are expanded as:

1) Define a data definition 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 subsantially new software at both the supplier and receiver facilities if data interchange software already exists. However, some modifications would be required.

2) Define a new format to work within the data definition superstrcture. 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 for coding various elements such as ANSI X3.43-1977 for local time of day. This structure would need to be maintained at a National level.

The recommendation was to pursue the first of these, while being sure that the second could follow as time and energies permit.

Several data interchange formats have been identified in Report 4 (Nyerges, 1984). These were discussed internally in committee meetings for their general suitability as per the conceptual model and the interchange objectives presented in the following section. Due to time limitations in the reporting cycle, a full report could not be produced detailing these discussions. In addition, early discussions showed that no single format could satisfy a majority of the objectives due to their limited scope of application.

1.Ø.1 General Digital Cartographic Data Interchange Objectives

The data interchange format must meet a defined set of objectives. The objectives must be stated clearly; then the details can be worked out. Without this, arguments over the details will have no ground for resolution. A suitable set of objectives for a general digital cartographic data interchange format is:

1) Provide an interchange format which is documented within the transfer media which will allow the receipient to read the data set to determine the basic logical and physical organizaton with minimal specific external information.

2) Provide a mechanism which will transfer raster, vector and text data in the same interchange format.*

3) Provide a format which will allow the inclusion of all necessary data such as: feature information, data quality, spatial and other data types, locational definitions, spatial and other relatonships, ancillary data, and map symbology.

4) Provide a format that will be expandable to include future types of catographic information.*

5) Adopt industry accepted standards to handle the various informaton required: e.g., Federal Information Processing Standards, American National Standards Institute standards, International Standards Organization standards or other labeling, coding ad transmission standards where applicable.

6) Provide an interchange format 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, disks and electronic transmission. The format shall allow storage on any of the available storage media.

7) Provide an interchange format that is computer independent to allow maximum transportability of data and permits flexibility in tradeoffs among transportability, storage overhead and processing efficiency.

8) Provide an interchange format that will allow transfer of logically different data sets on one physical medium or in onw equivalent electronic transmission. It must be possible to add incrementally onto the medium.

*These objectives are not included in the standard.

1.0.2 Data Organization Review

A general digital cartographic data interchange format must be able to support several types of data models. A data model consist of three components: a collection of object types, a collection of operators and a collection of general integrity rules (Codd, 1981). Codd formulated the concept of a data model within the context of the relational data base model (Codd, 1970).

Another definition for data model has been given by Ullman (1983) as a general description of specific sets of entities; leaving out the operators and rules as used by Codd. This specific set of entities would be an abstraction of the world. An entity is a thing which exists and is distinguishable; i.e. we can tell one entity from another. Thus, a chair, a person and a lake are each an entity (Ullman, 1982). An entity set is a class of entities that posses certain common characteristics. For example, lakes, mountains and desks are each entity sets. Relationships include such things as "left of", "less than" or "parent of". Both entities and relationships can have attributes or properties. These associate a specific value from a domain of values for that attribute with each entity in an entity set. For example, a lake may have attributes of size, elevation and suspended particulates among others.

The most basic characteristic of a data model (as defined by Ullman) is that it is an abstraction of reality. Each data model represents reality with a varying level of completeness. The more perfectly a model represents reality (i.e. the more completely all entities and relationships are incorporated) the more robust and flexible that model is in an application. However, the more precisely the model fits a single application, excluding entities and relations not required to deal with that application, the more efficient it will tend to be in storage space, speed and ease of use. The selection or design of a data model must, therefore, be based both on the nature of the phenomenon that the data represents and the specific manipulation processes which will be required to be performed on the data.

It is therefore, assumed that there can never be one single standard digital cartographic data format which is optimal in all cases. The ultimate objective of this standards process would be to derive a family of formats with at least one standard format for each major type of data model.

Many data model designers realize that in order to determine how a collection of data is to be represented in digital form, the data needs to be viewed at a number of discrete levels. These levels progress from reality, through user-oriented information structure or conceptual model, to the machine-oriented storage structure. There is, however, a lack of universal agreement as to how many levels of abstraction one should distinguish (Klinger, Fu and Kunii, 1977; Martin, 1975; Nyerges, 1980). These differences can in large part be attributed to design viewpoint. For the purpose of establishing a digital cartographic data format standard , two levels of abstraction are of interest. These are conceptual structure and data structure. The conceptual structure describes how the data are conceptually arranged in human terms, without machine or programming constraints. The data structure is a representation of the conceptual structures in terms of lists and arrays, and is designed to take the programming conventions and hardware into consideration.

In order to identify the minimum number and types of data formats which should be derived for a family of formats, the basic types of models on the conceptual level must be identified first.

Basic Types of Geographic Data Models

Two basic types of conceptual spatial data models have evolved for storing spatial images (including maps) in digital form, vector and tessellation models. In the vector type of data model, the basic logical unit in a geographical context corresponds to an entity in the world or a line on a map. This line may represent a river, street, area boundary or a segment of one of these. A series of x-y point locations along the line are recorded as descriptors of the entity. Points can be represented in a vector data organization as lines of zero length (i.e. one x-y location). With the tessellation type of organization, the logical unit is a location in space represented by a single cell or polygon in a mesh. The presence of specific entities are recorded as descriptors of each location.

Common usage has usually considered the tessellation type of data model to be based on a regular grid or rectangular pixel. However, this class encompasses more than the rectangular pixel; it also includes ay infinitely repeatable pattern of a regular polygon or polyhedron. This class can thus be referred to as a "regular tessellation".

Another class of spatial data models would be the hybrid of the vector and tessellation. This class would have characteristics of both classes.

Vector Data Model Subtypes

The two commonly known basic subtypes of vector data models are the spaghetti model and the topological model. There are a number of "classical" or prototypical examples of this type of model. These are discussed in Peucker and Chrisman (1975) and Peuquet (1984).

The spaghetti model is the simplest vector data model for geographic data. Each entity on the map becomes one logical record in the digital file, and is defined as strings of x-y coordinates. The lack of stored spatial relationships makes the spaghetti model unsuitable for analytical applications, although it is very efficient for reproducing the original map document as a vector plot.

The second subtype, and growing in acceptance, is the topological data model for geographic data. This type retains spatial relationships among entities by recording adjacency information explicitly. In addition, this topological information allows the spatial definitions of points, lines and polygon-type entities to be stored in a non-redundant manner.

Tessellation Data Model Subtypes

In tessellation of polygonal mesh models, the basic data unit is a unit of space for which entity information is recorded explicitly. Tessellation models can be further classified as regular, hierarchical and irregular tessellation models (Peuquet, 1984).

All three possible types of regular tessellations have been used as the basis of spatial data models. Each has different functional characterisitics which are based on the different geometries of the elemental unit (Ahuja, 1983). These three are square, triangular and hexagonal meshes.

Of these, the regular square or rectangular mesh has historically been the most widely used because 1) it was compatible with early pogramming language facilities and 2) it is compatible with a number of different types of hardware devices used for spatial data capture and output.

The primary advantage of the regular hexagonal mesh is that all neighboring cells of a given cell are equidistant from the centerpoint of the cell. Radial symmetry makes this model advantageous for radial search and retrieval functions. This is unlike the square mesh where diagonal neighbors are not the same distance away as neighbors in the four cardinal directions from the centerpoint.

A characteristic unique to all triangular tessellations, regular or irregular (i.e. where the triangles vary in size) is that the triangles do not all have the same orientation. This makes many procedures involving single-cell comparison operations which are simple to perform on the other two tessellations, much more complex. Nevertheless, this same characteristic gives triangular tssellations a unique advantage in representing terrain and other types of surface data.

1.1 Data Definition Sections

A data interchange shall consist of one or more files. A file is a collection of related records. A record is a collection of related data fields. Each data field stores an item of data. Each of these is treated as a logical unit.

The first record of a file must contain at least a record CORE section, and optionally a CORE EXTENSION section and DATA section. Coding in the CORE and in the CORE EXTENSION shall use the American Standard Coded Information Interchange (ASCII) 7-bit character set (ANSI X.34,1977) or where applicable, special control codes as specified in this document. Coding in the DATA section may be ASCII or Binary, as defined by the context of the interchange. This shall be noted in the Data Descriptive information of the CORE EXTENSION section.

When data interchange takes place through electronic transmission in an open networked environment (ISO/TC97/SC16, 1980) between machines, confusion between binary codes and special control shall be eliminated by using only ASCII characters in the CORE and CORE EXTENSION sections.

1.2 Data Definition Fields

The CORE shall consist of a series of field sets, each set composed of a two-character (decimal integer) size-field, a data field, and a unit separator (US). The data field may be subdivided to include a Tag subfield, an ASCII group separator (GS), and other data. Absence of the GS shall indicate no tag. In the last field set, the US shall be replaced by a record separator (RS).

(Note: Non-printable control characters 1/13 (ASCII GS), 1/14 (ASCII RS) and 1/15 (ASCII US) are represented in this document by GS, ";" (or RS) and "&" (or US) respectively.)

This logical format for data definition is recognized as the format of the file by assuming that this format is being used and finding the USs and the RSs in their proper places. If not found, then the data is defined with another format, and external documentation must be consulted.

1.2.1 Core Fields

The CORE section shall consist of specific field sets. The meaning of a field set is determined from the contents of a field set except for the first two field sets: Maintenance Authority and Format ID. In all cases, the first two field sets shall be the Maintenance Authority and Format ID (defined below).

A CORE field set may be constructed with or without tags which provide for the reference to an interpretation of the fields. When a tag is used in the field set, the field set shall be subdivided with the tag appearing in the first subdivision. In this case a field set would be composed of the following:

Data Length, Tag, ASCII GS, Data Field, US.

The Data Length shall indicate the length of the Tag, GS, Data Field and US by using a two-character (decimal integer) size. A GS missing from a field set shall indicate that no Tag is being used. In the absence of a tag, the order of field sets in the CORE mst be determined by reference to external documentation.

Alternatively, one field set can be designated as a tag list, eliminating the need for tags in other field sets. This tag list shall appear early in the CORE, immediately following the Format ID. Its data field would be subdivided with the first section being a TAGLIST tag. The ASCII GS, and variable length sub-fields, separated with commas, containing the series of CORE Tags would then follow in their respective order. This will allow the local implementor to use any desired Tag structure, including none at all. The CORE TAGLIST field set list is then considered to be an ordered list of CORE field sets. The field sets may include the tags defined in the TAGLIST as an option.

The CORE fields shall be specified in the following order (if no tags are used):

MAINTENANCE AUTHORITY

This field defines the authority under which the format is defined and maintained.

FORMAT IDENTIFICATION (FORMAT ID)

This field defines the specific format identification, including revisions as defined by the maintenance authority.

DATA START POSITION

This field defines the byte postion of the data in the record, relative to the first byte of the record.

RECORD TYPE AND SUBTYPE

A four byte code using ASCII characters.

DATA DEFINITION INDICATOR (DDI)

This field defines the style and location of Data Definition (DD) information pertaining to the data fields in the CORE EXTENSION. The DDI may take one of three forms, depending on the data defintion style to be used:

	DDI	Entry Map
No Data Definition At All	Zero	None, or all zero
Inline DD, Long Form	Terminator Symbol	Non-zero Integer
Inline DD, Short Form	Terminator Symbol	None, or all zero
DDR/DR Groups	Grouping Code	Non-zero Integer

The DDI Terminator Symbol shall be an ASCII printable symbol, but not), (or *, which have specific reserved meanings. The DDI Grouping Code shall be an ASCII alphanumeric character, but not zero.

RECORD NUMBER

The first record in each file shall be number "1".

RECORD LENGTH

The total logical record length, in bytes.

EXTERNAL AUTHORITY (EXTAUTH)

Under some conditions, the structure and meaning of the data fields can be determined by reference to known external documents. This field defines the source of those documents. This field shall apply to the entire file when found in the Volume Descriptor, or only to those data records carrying the same DDI, when found in a DDR.

EXTERNAL FORMAT (EXTFMT)

This field defines the document identification version as defined by the EXTAUTH. This field shall apply to the entire file when found in the Volume Descriptor, or only to those Data records carrying the same DDI, when found in a DDR.

PACKING FACTOR

Defines the division of physical records into logical records.

RECORD REPETITION INDICATOR

Allows subsequent data records to omit the CORE. This indicator shall contain the number of subsequent data records without a CORE or a CORE EXTENSION. The Data fields of those records shall be the same structure as those of the current record.

1.2.2 Core Extension Fields

The CORE EXTENSION section in a record shall be defined according to the type of record, Data Definition Record (DDR) or Data Record (DR). The major components and functions of these records are as follows:

Record	Component	Function
DDR	Leader	Identifies the DDR Contains the entry map (sizes of the tag, length, and position fields of the corresponding directory entries in this record)
	Directory	Gives tag, length, and position (relative to the start of the Data Description component) of each Data Descriptive field in this record.
	Data Descriptions	Structure of each corresponding Data Field in the DR.
DR	Leader	Identifies the DR Contains tge entry map (sizes of the tag, length, and position fields of the corresponding directory entries in this record)
	Directory	Gives tag, length, and position relative to start of the Data component) of each data field in this record.
	Data Fields	These fields have the format as described in the corresponding DDR Data Descriptions component.

The Tags described in the DDR and DR above relate the DDR Directory entries to the corresponding DR Directory entries. This implicitly links the data fields to their respective data descriptive fields. No tags appear in the DDR Data Descriptions or in the DR Data Fields.

The 8211 data definition capability presented here has been extended to include an ability to indicate the grouping and uses of various types of records, define records of various formats within one file, enclose and identify existing data in formats created by different organizations, and provide for transmission of binary data.

1.3 Different Forms of Data Definition

1.3.1 DDR Components with DR Components

When the optional DDR Components are included in the record, the fields will appear in the following order:

DATA DEFINITION RECORD (DDR) LEADER

Identifies a DDR. Contains an entry map and certain format information. See the 8211 defining document.

DIRECTORY

Multiple entries, one for each identified data field in the Data Description section, as follows (styled after 8211):

Data Description Field Tag	ASCII Chars of t length
Data Description Field Length	ASCII Integer of m digits
Data Description Field Position	ASCII Integer of n digits

DATA DESCRIPTION section of the DDR

Defines the structure of the DR data fields, and has a structure for each entry (corresponding to Directory entries), in accordance with 8211:

Field Control	Defines the type of field: integer, complex, etc.
Separator	RS/US
Field Name	Optional, a user supplied name
Label	Optional labels for subfields
Format Control	Optional Fortran Style (See Note 1) Format Designations needed

Note 1: Because the Fortran structure definitions do not include binary fields, the following designation for binary fields is used:

nZx...x(L|R)y...y

where,

n is the number of repetitions (ommitted if =1)
Z indicates a binary field
x is the field length in bits (decimal)
L|R indicates left or right justification of the bits
y is the number of significant bits (decimal)

1.3.2 Inline Structure

When a shorter form than the entire DDR/DR is desired, the structure shall include an Entry Map for the Long Form and not include and Entry Map for the short form. The Entry Map shall consist of:

	Field Size	
Size of directory Length Field	I1	Integer (=m length)
Size of Directory Position Field	I1	Integer (=n length)
Reserved	I1	ASCII zero
Size of Directory Tag Field	I1	Integer (=t length)

The fields shall be defined as:

Tag	Same tag as used in the Directory, followed by the US.
Length	Length of corresponding DR data field, followed by the US.
Position	Relative position of the corresponding DR data field in the data area (first position is zero), followed by the US.
Format Control	Fortran Style Format Designation (See Note 1), followed by the RS. A format label for a linear array is one in which a series of individual label subfields describes the components of a major data field. Fields for linear arrays shall use the chosen Terminator Symbol of the DDI to separate the sub-fields.

1.3.3 Short Form Inline Structure

"Format Control" is the only structuring used in the short form of the inline structure. For each data field, a Fortran Style Format Designation (See Note 1) shall be used, followed by a unit separator. The last Format Control Field shall be followed by a

Record Separator, instead of the unit separator. Linear array format control fields shall use the Data Definition Indicator terminal symbol as internal separators.

1.3.4 No Data definition At All

This form of structural definition requires the CORE as introductory information, but does not require any other information for definition.

1.4 Different Forms of Data Records

The Data Component shall consist of data fields in the same manner as described in Section 1.3 for each of the different structural definitions corresponding to CORE and CORE EXTENSION fields. The CORE EXTENSION is interpreted in the Data Component to be a "Data Record CORE EXTENSION".

In Data Records using the Inline Structure, Short Form Inline Structure, and/or No Data Defintion At All, the data fields shall be contiguous without unit separator or record separator terminators. Extenal format control definitions shall be the only means of encoding and decoding the data field values in the transmission when these definitions are used.

1.5 Types of Data Definitions in a Transmission

Files may be defined and transmitted in several versions, depending upon the robustness desired and the degree of field structure specification. These forms will use the components described in the standard in various combinations:

1.5.1 No Data Definition At All

This would be used when little flexibility in data interchange is needed by an organization. This would be coded as zero in the Data Definition Indicator. External documentation must be sought for all data record structures. (See Appendix A.1 of the standard for more details).

1.5.2 Data Definition (DD) included within the Data Record

The DD information would follow the core and precede the data fields. This structure might be used when each data record has a unique structure, and/or where the overhead of the full 8211 structure is undesireable. Thus, the shortened definition is appropriate. The core and a data definition structure are derived from the 8211 structure. (See Appendix A of the Proposed Standard for more details.)

1.5.3 Sets of Data Definition Records (DDR) and Data Records (DR)

Data Definition Records (DDR)

The 8211 definition requires only one DDR per data record group, with all fields tagged in the Directory and Data Descriptions. In the structure defined in the standard, multiple groups are allowed with all of the records in a record group having the same fomat. Thus, one such DDR would precede a set of data records, each of which has the structure as defined in the preceding DDR. This requires recognition coding in the DR which indicated that there is an external DDR. This will be indicated by a Grouping Code in the DDI of each DDR and DR, using the same symbol throughout the DDR-DR group.

The 8211 structure will start immediately after the core. Although there is some overlap between the purposes and information in the core and the 8211 DDR leader, the 8211 leader will be included in its entirety.

- When the 8211 DDR or DR core is used, the Core Entry Map is not needed; the 8211 entry Map (part of the leader) shall convey the required information. (See Appendix A in the Proposed Standard for more details.)

Data Records

The entire DR leader will be included (or not), as indicated in Par. 5.3.1.3 of the 8211 document. The DDR/DR repetitive characteristics would be coded in RP 6 of the Leader, using the coding of Par 5.3.1.3 of the 8211 document. However, note that the DR Base-Address-of-Data will be different in the DRs, depending on the repetitive (or not) inclusion of subsequent leader and directories. Therefore, the DR leader address-of-data field shall be defined to apply to the concurrent record, and the data start must be calculated by the software for subsequent records. Data locations as given in the Directory are relative to the start of the Data area (first position = zero). (See Appendix A of the Proposed Standard for more details.)

Data record structure is guided by several considerations: 1) It is desireable to enclose extant data sets without further modification. These normally will not have the 8211-style data field terminators. However, new data sets following the 8211 must have these terminators. 2) In the DDR/DR structure, the 8211 Record ID field must be present, with a tag Ø...1. Other structures may have different, untagged record ID structures. 3) It is necessary to use several differing (in structure) DRs in one file. An accompanying DDR must be available for each structure DR.

These can be accomodated by:

1) When enclosing data sets with inline DD and no tags, use the Core ad Core Extension, Entry Map, Directory, and Inline Structure, followed by the old data set verbatim.

2) Construct new data set per 8211, with DDRs and DRs, and data fields with terminators. Include a new record ID field preceding the data fields in the DR, and so indicate in the DDR by the tag Ø...1. This fulfills the required format for the DDR.

3) It will often be necessary to enclose several 8211-style DDR/DR sets within one transmission. The Data Definition Indicator includes information to indicate these groups of DDR/DR/DR... Within each set, the data records will have constant format.

The DR fields might have two field structures: Contents-terminator or Contents only. The former would be used with the external DDR to follow the 8211 structure. The latter is not 8211, but would be used with the more compact inline DD where the fields are in a pre-defined order and all are present or indicated with the unit separator 1/15.

It is to be recognized that the full 8211 structure is more robust than the inline or otherwise shortened versions, and is therefore preferred for new data sets.

1.6 Grouping the Records

The various records in a transmission serve a number of functions: Transaction Record, Data Definition Records, File Pointers, Data Record, Ancillary Records of various types, Feature Records, Attribute Records, etc. It is convenient to group these, with a Volume Directory Group containing a Volume Description and File Pointers, and perhaps Data Definition records for these. A separate file might contain Data Definition records for the records in this file, Ancillary Records (AR), Data records and others. A Null Volume Directory is defined as terminating the Volume.

A Volume might contain:

Superstructure
  Transaction Record

  Volume Descriptor
  DDR for the File Pointer
  File Pointer

Main File
  DDR for Data Records of Type A
  DDR for Data Records of Type B
  Text Record
  Data Records of Type A
  Data Records of Type B
  Ancillary Record, Security/Release
  Ancillary Record, Map Projection
    etc.

Superstrcture
  Null Volume Directory

1.6.1 Record Type Identification

Superstructure:

As indicated above, the superstructure will contain global information concerning the file as a whole. Suggested coding for the various Superstructure records is:

	Type	Sub-1
Transaction Record	S	T
Volume Descriptor	S	V
File Pointer	S	P
Null Volume Descriptor	S	N

1.6.2 Data Records

Records containing the data to be transmitted such as Imagery, Feature, Attribute and other records will be considered logically as data records. They wil require suitable treatment concerning data definition as previously discussed.

The standard formats include binary and ASCII encoded data. Binary fields are unsigned integer values. ASCII data is encoded in one seven-bit character per byte (eight bits, or octant), located in the low order bits Ø-6, with bit 7 equal to Ø. Numeric data fields represented in ASCII format are always right justified and the field is zero filled on the left. Character data within a field is always left justified and blank filled on the right. Zero fill will be binary zero in binary fields and

ASCII blank (to indicate no valid value) in ASCII fields. Blank fill always means ASCII blanks.

The Ancillary Records

In the concept of a complete transmission, there may be a group of records ancillary to the primary data which carry related information. Some examples may be calibration, release/security, mapping conversions, accuracy, data set history, audit trails, and others. These will be located as DATA records by the system, and need to be identifed by the Type-Subtype coding.

It is necessary that a suitable group of ancillary records be identified, together with the codes from which the desired set can be assembled for any particular task. The following are suggested. Most of these are applicable to various organizations and data types, and can be used directly. Many of the groups of conventions will need expansion to suit particular uses. The following Table lists some examples:

Type	Suggested Code			Derivation
	Type	Sub-1	Sub-2	
New Test Records	T			NCDCDS
Radiometric Calibration	D	C	R	Landsat
Data Security/Release	A	X	S	DMA SLF DSSG*
Data Set Params	A	X	P	DMA SLF DSPG*
Map Projection	A	X	M	DMA SLF DSMP*
Lineage (History)	A	L		NCDCDS
Registration Points Group	A	X	R	DMA SLF DSRG*
Control Point Data	A	C		Landsat
Logical Consistency	Q	L		NCDCDS
Completeness	Q	C		NCDCDS
Currency and Time Related	Q	T		NCDCDS
Positional Accuracy	Q	P		NCDCDS
Accuracy Subset Group	Q	X	A	DMA SLF DSAG*
Feature Record	D	X	F	DMA SLF FEA
Segment Reocrd	D	X	S	DMA SLF SEG
Trailer Record	A	T		Landsat
Data Records	D			
Keyword List	A	K		

X - Used to signify DMA Standard Linear Format (SLF) Groups
* - SLF uses these in one record (Defense mapping Agency, 1983)

1.7 Preliminary Investigations

This section reports on preliminary investigation using the data definition structure described in the standard.

1.7.1 Context of Investigations

To be meaningful, the translation of extant formats should be performed in the context of the interchange process described in Section 1.0. This model itemizes four separate transformations; two implemented on the source machine and two on the target machine. The first and last transformations contain hardware dependent and DBMS information which may be proprietary. In our preliminary investigations, we are not concerned with these two transformations.

The second transformation, driven by the source data base schema, produces the standard interchange format, and the third transformation, driven by the target data base schema, reads the standard interchange format. It is these two transformations which should be investigated.

All transformations, however, are seen to be implemented in the form of standard software packages produced by the hardware and/or data base management system vendors. For data interchange to occur, these software packages must be in place on both the source and target systems. The software which would perform the two transformations of interest does not exist. And the complexity of the interchange process and the apparent infinite diversity of data sources and targets, suggests an extensive software design and development effort. This effort can not be part of the preliminary investigations.

1.7.2 Proposed Methodology

The methodology we propose for our preliminary investigations will not, by necessity, result in data files translated to or from an actual digital data set formatted per the proposed standard. We can, however, make some initial assumptions about the software which would perform the translations. Based upon these assumptions, we can itemize the steps one would take to use the translation software. And then, as far as possible, we should actually perform the steps for data available in a few existing formats. In this manner, we might be able to see if the structures carried by these formats can be mapped into and out of the proposed interchange format structures.

Initial Assumptions

As indicated above, the transformations of concern are driven by the source and target data base schemas. It is possible, and even desirable in many cases to have a separate software package for each data base schema to be transferred. We should assume here, however, that the software is very general, knowing nothing about any particular schema to be transferred. The software will know the proposed standard interchange format schema. But no schema-conversion intelligence can be assumed. We will be required to give precise instructions to the software, expressing a given exising format schema in terms of the strctures and objects available in the proposed standard interchange format. The actual format and syntax of these instructions can only be determined after a detailed software design is completed. But we should assume they must be sufficient to define the logical objects within the interchange format, and load the objects with tag-length-position.

Developing Schemas for Existing Formats

The first step in a preliminary investigation of a given data set is to develop a diagram corresponding to the data model of the data set. Then a diagram corresponding to the way the data will be stored in the interchange format is necessary. This will provide us with a tool to translate one structure into another. An optimal translation would be one where each object in the data structure of the source is translated to an object in the intermediate form. The records could be typed as discussed previously to reflect the source and target data sets.

The tagged record and field approach to defining a data set in the 8211 data definition can handle "elementary data, vectors, arrays, and hierarchies. User structures such as sequential, hierarchical, relational and indices can be converted into the interchange format. Network structures can be interchanged but additional pre-processing and post-processing is necessary to preserve logical linkages." Annex A to ISO/DIS 8211 provides guideines for mapping these various structures into the interchange standard.

If hierarchical structures are to be transferred, the linkages may be expressed in the schema in the form of "tag pairs" or special data fields. A network data base, however, should be resolved into its hierarchical components and then transferred; or resolved further into its relational counterparts and then transferred.

Providing Data Definition Information

Once a schema has been developed, defining the objects as a data model for interchange, the internal structure of the files, can be given. The internal structure of each field type should be defined in terms of data types (character, implicit floating point, explicit floating point, explicit point scaled, character mode bit string, bit field, and mixed types) available to the interchange format. Some data fields may require format control and others may not depending on the implementation of the interchange software.

Translating the Schema and Data Definition Information

The details of translating the schema and data definition information should be left to the software. An implementation of the software for the ISO/DIS 8211 format has been prepared by Oakridge National Lab personnel (ORNL/CSD/TM-207, October, 1983). This could provide a starting point for many implementations.

1.7.3 Department of Energy (DOE) Geographic Exchange Standard

This is an example of an embodiment of the ISO/DIS 8211 standard as experimented within DOE (Department of Energy, 1978) for data interchange. The coding of the information in this example is similar to the proposed standard, but is not entirely the same. A goup of examples for record coding appears in Appendix of this document.

The DOE document defines several data structures, such as grid, point lists, point neighbors, area segments, area chains, polygons, chained polygons, triangles, Thiessen polygons, neighborhoods, linear segments, chains, and spaghetti files. It provides tags for each and the DOE structure used.

It illustrates the use of the Data Definition Files and the Data Files of the ISO standard. One of the illustrations provided is as follows:

Simple Point List Example

```
DDF
     TAG 000 0000=:;&NWS POINT FILE7001PTNPTNPTS;
         001 0000=:;&RECORD KEY:
         PTN 0000=:;&REGION NAME:
         PTS 2400=:;&&POINTS IN DEGREES&*NAME!NEGATIVE WEST
                      LONGITUDE!LATITUDE&(A,2R)=

DF
     TAG 001 1;
     PTN SOUTHEAST;
     PTS AGS&-80.48&31.50&JAX&-79.85&27.47&...=
```

The first four characters following the tag number are the "structure code". For tag 000, the structure code is meaningless, so 0000 is used. The implementation requires the terminators to be defined next, hence "=:;&". In tag 000, "NWS POINT FILE" is the name of the file. "001PTNPTNPTS" signifies that the tag PTN is owned by tag 001, and tag PTS is owned by tag PTN. An "&" separates the file label from the hierarchy information.

Tag PTN consists of simple alphanumeric data which is named "REGION NAME". Data for tag PTS consists of a mixed array as indicated by the structure code "2400". Data elements are named by the label subfield elements, "NAME", "NEGATIVE WEST LONGITUDE", "LATITUDE". The asterisk "*" is required to indicate missing row information. The format subfield (A,2R) indicates delimited elements in the pattern of an alphanumeric element followed by two explicit point elements. (The various specific codes used by DOE and their meanings are defined in the DOE embodiment document, 1978).

Each field (DDF TAG entry) is terminated with a ";" and the file is terminated with an "=". Since both the DDF and DF are one-record files, ":" is not used. (This would be used in multi-record files.)

The DF contains corresponding tags where tag 001 gives the record key of "1". The "REGION NAME is SOUTHEAST". Data elements are separated by the unit terminator "&" as required by the standard when the element field widths are not specified in the DDF format subfield.

Thus the entries shown are:

Name	Negative West Longitude (degrees)	Latitude (degrees)
AGS (Augusta)	-80.48	31.50
JAX (Jacksonville)	-79.85	27.47

Other examples of more complex structures are given as well. The document gives tags to essentially all exchange entities which might be encountered; and more could be defined by a particular user, if required. For the ones in the example, PTN means point name, the name of a collection of points, and PTS are points, with "A,2N" signifying the pointname, "geoxy".

1.8 Conclusions

The conceptual model of the interchange process provides a convenient method of describing the stages of the interchange process.

The objectives for a general digital cartographic interchange format provide an operational checklist of objectives to make sure the format is accomplishing a goal of general cartographic applicability.

A general format which will include extant formats is practical because this allows continued use of different interchange formats as well as support the exchange of existing internal formats.

The envelop structure and data definition features permit the definition of the structure to be included in the interchange of the data. This eliminates the need for reverting to ecternal documentation for every field in the interchange format. Tje overhead which permits the self-definition of structures, makes the format more complicated than formats that are defined externally. The self-definition, however, promotes generality. The proposed standard can be applied to many different data sets, and thus supports a general exchange of data.

REFERENCES

Ahuja, N. (1983). "On Approaches to Polygonal Decomposition for Hierarchical Image Decomposition," Computer Vision, Graphics and Image Processing, Vol. 24, pp. 200-214.

ANSI X3.4-1977 (1977). American National Standards Institute, "American National Standard Code for Information Interchange", 1977.

Brooks, A., Hammerling and McNeely (1983). "User's Guide for an IBM PL/I Implementation of ISO/DIS 8211 Information Processing," October, 1983.

Codd, E. F. (1970). "A Relational Model of Data for Large Shared Data Banks", Communications of the ACM, Vol. 13, pp. 378-387.

Codd, E. F. (1981). "Data Models in Database Management," Proceedings, Workshop on Data Abstraction, Databases and Conceptual Modelling.

Defense Mapping Agency (1983). "Standard Linear Format for Digital Cartographic Feature Data," Defense Mapping Aerospace Center, St. Louis, MO. May, 1983.

Department of Energy (1978). "Interlaboratory Working Group on Data Exchange - A Geographic data Exchange Primer", Battelle Pacific Northwest Lab.

Gligor, V. and G. Lukenbaugh, (1984). "Interconnecting Heterogeneous Database Management Systems", Computer, January, 1984, pp. 33-43.

ISO 646-1973(E). International Standards Organization, "7-Bit Coded Character Set for Information Processing Interchange."

ISO/DIS/8211 (1983). International Standards Organization, "Specification for a Data descriptive file for Information Interchange."

ISO/TC97/SC16/537 (1980). International Standards Organization, "Open Systems Interconnection, Basic Reference Model," December, 1980.

Klinger, A., K. S. Fu, and T. L. Kunii (1977). Data Structures, Computer Graphics and Pattern Recognition, Academic Press : New York.

Kunii, H. (1983). "Graph Data Language, A High Level Access-Path Oriented Language," unpublished dissertation, Department of Computer Science, University of Texas at Austin.

Martin, J. (1975). *Computer Data-Base Organization*, Prentice Hall, Inc. : Englewood Cliffs.

Nyerges, T. (1980). "Modeling the Structure of Cartographic Informaton for Query Processing," unpublished dissertation, Department of Geography, The Ohio State University.

Nyerges, T. (1984). "Digital Cartographic Data Standards: Alternatives for Data Organization" in *Digital Cartographic Data Sandards: Examining the Alternatives* ed. H. Moellering, Issues in Digital Cartographic Data Standards Report 4, Numerical Cartography Lab, Department of Geography, Ohio State University, Columbus, Ohio.

Peucker, T. and N. Chrisman (1975). "Cartographic Data Structures," *American Cartographer*, Vol. 1, No. 2, pp. 55-69.

Peuquet, D. (1984). "A Conceptual Framework and Comparison of Spatial Data Models," unpublished report, Remote Sensing Unit, Department of Geography, University of California at Santa Barbara.

Ullman, J. (1983). *Principles of Database Systems*, Computer Science Press : Rockville, MD.

ILLUSTRATIONS OF RECORD FORMS

TRANSACTION RECORD

A Transaction Record will normally be the first record of a transmission. It will be illustrated as a record containing a single undifferentiated data field.

A Transaction Record with 360 bytes total length, M=5, N=4, m=2, n=3, t=3 and one text field might be coded with the full INLINE STRUCTURE data definition as:

Field		Rel Pos'n	Format	Contents
DCDS CORE				
Size of Maintenance Auth Field	M'	0 - 1	I2	05
Maintenance Authority		2 - 5	A4	DCDS
Unit Separator		6	A1	&
Size of Format ID Field	N'	7 - 8	I2	05
FormatID		9 - 12	A4	5BR3
Unit Separator		13	A1	&
Size of Data Start Position	S	14 - 15	I2	04
Data Start Position		16 - 18	I3	097
Unit Separator		19	A1	&
Size of Record Type/Subtype	T	20 - 21	I2	05
Record Type and Subtype		22 - 25	A4	STbb
Unit Separator		26	A1	&
Size of Data Definition Indicator	D	27 - 28	I2	02
Data Definition Indicator (DDI)		29	A1	!
Unit Separator		30	A1	&
Size of Record Number Field	L	31 - 32	I2	02
Record Number		33	I1	1
Unit Separator		34	A1	&
Size of Record Length Field	K	35 - 36	I2	04
Record Length		37 - 39	I3	360
Unit Separator		40	A1	&
Size of EXAUTH Field	A	41 - 42	I2	01
Unit Separator		43	A1	&
Size of EXTFMT Field	F	44 - 45	I2	01
Unit Separator		46	A1	&
Size of Entry Map Field	E = 0 or 4	47 - 48	I2	05
Entry Map		49 - 52	4I1	2303
Unit Separator		53	A1	&
Size of Packing Field	P	54 - 55	I2	01
Unit Separator		56	A1	&
Size of Record Repetition Factor	R	57 - 58	I2	01
Record Separator		59	A1	;
Error Correcting Code (Example only)		60 - 63	A4	AAAA
Record Separator		64	A1	;
CORE EXTENSION				
Size of Directory Field		65 - 66	I2	09
Tag (size = t)		67 - 69	A3	TXT
Length (size = m)		70 - 71	I2	20

Position (size = n)	72 - 74	I3	086
Field Terminator (Unit Separator)	75	A1	&
Size of Inline Structure Field	76 - 77	I2	19
Tag	78 - 81	A4	TXT&
Length	82 - 85	A4	262&
Relative Position	86 - 89	A4	000&
Format Control	90 - 95	A6	A(262)
Terminator (Record Separator)	96	A1	;
DATA FIELD			
Text	97 - 358	A262	ttt...ttt
Terminator	359	A1	;

ILLUSTRATION OF ENCLOSING A PREVIOUS FORMAT

Consider the 360 byte Landsat TM Volume Directory (first portion only for brevity):

Field	Local Rel Pos'n	Format	Contents * Indicates Decimal to be Binary Coded
CCT Core 0 - 15 C16			
Record Number	0 - 3	C4	1*
Record Type	4 - 7	C4	90H 90H 12H 12H
Record Length	8 - 11	C4	360*
ASCII/EBCDIC Flag	12	A1	A
Reserved	13 - 15	A3	bbb
The Data A344			
Superstructure Format Doc	16 - 27	A12	Mixed ASCII
Revision Number	28 - 29	I2	Decimal
Revision Letter	30 - 31	A2	ASCII Alpha
Software Release Number	32 - 43	A12	Mixed ASCII
Tape/Mission ID	44 - 59	A16	Mixed ASCII
Logical Volume ID	60 - 75	A16	Mixed ASCII
(Remainder, for brevity)	76 - 359	A284	Mixed ASCII

CASE 1. No DD

Landsat Volume Descriptor enclosed and identified, with full DCDS Core:

	Local RP	Global RP	Format	Contents
DCDS CORE				
Size of Local ControlAuth Field	M'		I2	05
Local Control Authority			A4	DCDS
Unit Separator			A1	&
Size of Local Format ID Field	N'		I2	05
Local Format ID			A4	5BR3
Unit Separator			A1	&

Size of Data Start Position	S		I2	04
Data Start Position			I3	080
Unit Separator			A1	&
Size of Record Type/Subtype	T		I2	05
Record Type and Subtype			A4	SVbb
Unit Separator			A1	&
Size of Data Definition Indicator	D		I2	02
Data Definition Indicator (DDI)			A1	0
Unit Separator			A1	&
Size of Record Number Field	L		I2	03
Record Number			I1	02
Unit Separator			A1	&
Size of Record Length Field	K		I2	04
Record Length			I3	441
Unit Separator			A1	&
Size of EXAUTH Field	A		I2	08
EXAUTH Field			A7	LANDSAT
Unit Separator			A1	&
Size of EXTFMT Field	F		I2	12
EXTFMTField			A11	CCT-CCB-002
Unit Separator			A1	&
Size of Entry Map Field	E = 0 or 4		I2	01
Unit Separator			A1	&
Size of Packing Field	P		A2	01
Unit Separator			A1	&
Size of Record Repetition Factor	R		I2	01
Record Separator			A1	;
Error Correcting Code (Example Only)			A4	AAAA
Record Separator			A1	;
THE DATA		* Decimal Number, binary coded		
CCT Core C16				
Record Number	0 - 3	80 - 83	C4	1*
Record Type	4 - 7	84 - 87	C4	90H 90H 12H 12H
Record Length	8 - 11	88 - 91	C4	360*
ASCII/EBCDIC Flag	12	92	A1	A
Reserved	13 - 15	93 - 95	A3	bbb
The Data A344				
Superstructure Format Doc	16 - 27	96 - 107	A12	Mixed ASCII
Revision Number	28 - 29	108 - 109	I2	Decimal
Revision Letter	30 - 31	110 - 111	A2	ASCII Alpha
Software Release Number	32 - 43	112 - 123	A12	Mixed ASCII
Tape/Mission ID	44 - 59	124 - 139	A16	Mixed ASCII
Logical Volume ID	60 - 75	140 - 155	A16	Mixed ASCII
(Remainder, for brevity)	76 - 359	156 - 439	A284	Mixed ASCII
Terminator		440	A1	;

<u>CASE 1A. No DD. Minimum Core Form</u>

	Local RP	Global RP	Format	Contents
DCDS CORE				

Size of Local ControlAuth Field	M'	I2	05
Local Control Authority		A4	DCDS
Unit Separator		A1	&
Size of Local Format ID Field	N'	I2	05
Local Format ID		A4	5BR3
Unit Separator		A1	&
Size of Taglist Field		I2	22
Taglist		A22	TAGLIST, EXAUTH,EXTFMT
Unit Separator		A1	&
Size of EXAUTH Field	A	I2	08
EXAUTH Field		A7	LANDSAT
Unit Separator		A1	&
Size of EXTFMT Field	F	I2	12
EXTFMTField		A11	CCT-CCB-002
Record Separator		A1	;
Error Correcting Code (Example Only)		A4	AAAA
Record Separator		A1	;

ILLUSTRATIONS OF RECORD FORMS

TRANSACTION RECORD

A Transaction Record will normally be the first record of a transmission. It will be illustrated as a record containing a single undifferentiated data field.

A Transaction Record with 360 bytes total length, M=5, N=4, m=2, n=3, t=3 and one text field might be coded with the full INLINE STRUCTURE data definition as:

Field		Rel Pos'n	Format	Contents
DCDS CORE				
Size of Maintenance Auth Field	M'	0 - 1	I2	05
Maintenance Authority		2 - 5	A4	DCDS
Unit Separator		6	A1	&
Size of Format ID Field	N'	7 - 8	I2	05
FormatID		9 - 12	A4	5BR3
Unit Separator		13	A1	&
Size of Data Start Position	S	14 - 15	I2	04
Data Start Position		16 - 18	I3	097
Unit Separator		19	A1	&
Size of Record Type/Subtype	T	20 - 21	I2	05
Record Type and Subtype		22 - 25	A4	STbb
Unit Separator		26	A1	&
Size of Data Definition Indicator	D	27 - 28	I2	02
Data Definition Indicator (DDI)		29	A1	!
Unit Separator		30	A1	&
Size of Record Number Field	L	31 - 32	I2	02
Record Number		33	I1	1
Unit Separator		34	A1	&
Size of Record Length Field	K	35 - 36	I2	04
Record Length		37 - 39	I3	360
Unit Separator		40	A1	&
Size of EXAUTH Field	A	41 - 42	I2	01
Unit Separator		43	A1	&
Size of EXTFMT Field	F	44 - 45	I2	01
Unit Separator		46	A1	&
Size of Entry Map Field	E = 0 or 4	47 - 48	I2	05
Entry Map		49 - 52	4I1	2303
Unit Separator		53	A1	&
Size of Packing Field	P	54 - 55	I2	01
Unit Separator		56	A1	&
Size of Record Repetition Factor	R	57 - 58	I2	01
Record Separator		59	A1	;
Error Correcting Code (Example only)		60 - 63	A4	AAAA
Record Separator		64	A1	;

CORE EXTENSION			
Size of Directory Field	65 - 66	I2	09
Tag (size = t)	67 - 69	A3	TXT
Length (size = m)	70 - 71	I2	20
Position (size = n)	72 - 74	I3	086
Field Terminator (Unit Separator)	75	A1	&
Size of Inline Structure Field	76 - 77	I2	19
Tag	78 - 81	A4	TXT&
Length	82 - 85	A4	262&
Relative Position	86 - 89	A4	000&
Format Control	90 - 95	A6	A(262)
Terminator (Record Separator)	96	A1	;
DATA FIELD			
Text	97 - 358	A262	ttt...ttt
Terminator	359	A1	;

ILLUSTRATION OF ENCLOSING A PREVIOUS FORMAT

Consider the 360 byte Landsat TM Volume Directory (first portion only for brevity):

Field	Local Rel Pos'n	Format	Contents * Indicates Decimal to be Binary Coded
CCT Core 0 - 15 C16			
Record Number	0 - 3	C4	1*
Record Type	4 - 7	C4	90H 90H 12H 12H
Record Length	8 - 11	C4	360*
ASCII/EBCDIC Flag	12	A1	A
Reserved	13 - 15	A3	bbb
The Data A344			
Superstructure Format Doc	16 - 27	A12	Mixed ASCII
Revision Number	28 - 29	I2	Decimal
Revision Letter	30 - 31	A2	ASCII Alpha
Software Release Number	32 - 43	A12	Mixed ASCII
Tape/Mission ID	44 - 59	A16	Mixed ASCII
Logical Volume ID	60 - 75	A16	Mixed ASCII
(Remainder, for brevity)	76 - 359	A284	Mixed ASCII

CASE 1. No DD

Landsat Volume Descriptor enclosed and identified, with full DCDS Core:

	Local RP	Global RP	Format	Contents
DCDS CORE				
Size of Local ControlAuth Field	M'		I2	05
Local Control Authority			A4	DCDS
Unit Separator			A1	&
Size of Local Format ID Field	N'		I2	05
Local Format ID			A4	5BR3
Unit Separator			A1	&
Size of Data Start Position	S		I2	04
Data Start Position			I3	080
Unit Separator			A1	&
Size of Record Type/Subtype	T		I2	05
Record Type and Subtype			A4	SVbb
Unit Separator			A1	&
Size of Data Definition Indicator	D		I2	02
Data Definition Indicator (DDI)			A1	0
Unit Separator			A1	&
Size of Record Number Field	L		I2	03
Record Number			I1	02
Unit Separator			A1	&
Size of Record Length Field	K		I2	04
Record Length			I3	441
Unit Separator			A1	&
Size of EXAUTH Field	A		I2	08
EXAUTH Field			A7	LANDSAT
Unit Separator			A1	&
Size of EXTFMT Field	F		I2	12
EXTFMTField			A11	CCT-CCB-002
Unit Separator			A1	&
Size of Entry Map Field	E = 0 or 4		I2	01
Unit Separator			A1	&
Size of Packing Field	P		A2	01
Unit Separator			A1	&
Size of Record Repetition Factor	R		I2	01
Record Separator			A1	;
Error Correcting Code (Example Only)			A4	AAAA
Record Separator			A1	;

THE DATA		* Decimal Number, binary coded		
CCT Core C16				
Record Number	0 - 3	80 - 83	C4	1*
Record Type	4 - 7	84 - 87	C4	90H 90H 12H 12H
Record Length	8 - 11	88 - 91	C4	360*
ASCII/EBCDIC Flag	12	92	A1	A
Reserved	13 - 15	93 - 95	A3	bbb
The Data A344				
Superstructure Format Doc	16 - 27	96 - 107	A12	Mixed ASCII
Revision Number	28 - 29	108 - 109	I2	Decimal

Revision Letter	30 - 31	110 - 111	A2	ASCII Alpha
Software Release Number	32 - 43	112 - 123	A12	Mixed ASCII
Tape/Mission ID	44 - 59	124 - 139	A16	Mixed ASCII
Logical Volume ID	60 - 75	140 - 155	A16	Mixed ASCII
(Remainder, for brevity)	76 - 359	156 - 439	A284	Mixed ASCII
Terminator		440	A1	;

CASE 1A. No DD. Minimum Core Form

	Local RP	Global RP	Format	Contents
DCDS CORE				
Size of Local ControlAuth Field	M'		I2	05
Local Control Authority			A4	DCDS
Unit Separator			A1	&
Size of Local Format ID Field	N'		I2	05
Local Format ID			A4	5BR3
Unit Separator			A1	&
Size of Taglist Field			I2	22
Taglist			A22	TAGLIST, EXAUTH,EXTFMT
Unit Separator			A1	&
Size of EXAUTH Field	A		I2	08
EXAUTH Field			A7	LANDSAT
Unit Separator			A1	&
Size of EXTFMT Field	F		I2	12
EXTFMTField			A11	CCT-CCB-002
Record Separator			A1	;
Error Correcting Code (Example Only)			A4	AAAA
Record Separator			A1	;
THE DATA				
CCT Core C16				
Record Number	0 - 3	67 - 70	C4	1*
Record Type	4 - 7	71 - 74	C4	90H 90H 12H 12H
Record Length	8 - 11	75 - 78	C4	360*
ASCII/EBCDIC Flag	12	79	A1	A
Reserved	13 - 15	80 - 82	A3	bbb
The Data A344				
Superstructure Format Doc	16 - 27	83 - 94	A12	Mixed ASCII
Revision Number	28 - 29	95 - 96	I2	Decimal
Revision Letter	30 - 31	97 - 98	A2	ASCII Alpha
Software Release Number	32 - 43	99 - 110	A12	Mixed ASCII
Tape/Mission ID	44 - 59	111 - 126	A16	Mixed ASCII
Logical Volume ID	60 - 75	127 - 142	A16	Mixed ASCII
(Remainder, for brevity)	76 - 359	143 - 426	A284	Mixed ASCII
Terminator		427	A1	;

CASE 2. IN-LINE DD

INLINE DATA DEFINITION, NO TAGS IN DATA FIELDS

The forms illustrated below depend on to what depth it is desired to locate fields by the tags. Major components not coded.

INLINE DATA DEFINITION FORM 1 - SHORT FORM, FULL CORE

If the shortest in-line DD is desired in a data record, the coding (for M=5, N=4, field separator = &) with the SHORT FORM STRUCTURE would be:

(Record coded is the Landsat Volume Descriptor):

Field	Local Rel Posn	Global Rel Posn	Format	Contents
DCDS CORE				
Size of Local ControlAuth Field	M'		I2	05
Local Control Authority			A4	DCDS
Unit Separator			A1	&
Size of Local Format ID Field	N'		I2	05
Local Format ID			A4	5BR3
Unit Separator			A1	&
Size of Data Start Position	S		I2	04
Data Start Position			I3	121
Unit Separator			A1	&
Size of Record Type/Subtype	T		I2	05
Record Type and Subtype			A4	SVbb
Unit Separator			A1	&
Size of Data Definition Indicator	D		I2	02
Data Definition Indicator (DDI)			A1	!
Unit Separator			A1	&
Size of Record Number Field	L		I2	02
Record Number			I1	2
Unit Separator			A1	&
Size of Record Length Field	K		I2	04
Record Length			I3	482
Unit Separator			A1	&
Size of EXTAUTH Field	A		I2	08
EXTAUTH Field			A7	LANDSAT
Unit Separator			A1	&
Size of EXTFMT Field	F		I2	12
EXTFMTField			A11	CCT-CCB-002
Unit Separator			A1	&
Size of Entry Map Field	E = 1 OR 5		I2	01
Unit Separator			A1	&
Size of Packing Field	P		A2	01
Unit Separator			A1	&
Size of Record Repetition Factor	R		I2	01
Record Separator			A1	;
Error Correcting Code (Example Only)			A4	AAAA
Record Separator		79	A1	;

CORE EXTENSION				
SHORT FORM STRUCTURE A41				
SS Format Contr Field Length		80 - 85	A6	A(12)&
Rev Number Field Length		86 - 90	A5	I(2)&
Rev Letter Field Length		91 - 95	A5	A(2)&
Software Rel Field Length		96 - 101	A6	A(12)&
Mission ID Field Length		102 - 107	A6	A(16)&
Logical Vol ID Field Length		108 - 113	A6	A(16)&
RemainderFieldLength		114 - 120	A7	A(285);
THE DATA A360				
CCT Core C16				
Record Number	0 - 3	121 - 124	C4	1*
Record Type	4 - 7		C4	90H 90H 12H 12H
Record Length	8 - 11		C4	360*
ASCII/EBCDIC Flag	12		A1	A
Reserved	13 - 15		A3	bbb
CCT Data A344				
Superstructure Format Doc	16 - 27		A12	MIXED ASCII
Revision Number	28 - 29		I2	Decimal
Revision Letter	30 - 31		A2	ASCII Alpha
Software Release Number	32 - 43		A12	Mixed ASCII
Tape/Mission ID	44 - 59		A16	Mixed ASCII
Logical Volume ID	60 - 75		A16	Mixed ASCII
(Remainder, for brevity)	76 - 359		A284	Mixed ASCII
Terminator		481	A1	;

It can be seen that the contents are the same as the original data, moved over to give room for the core and extension.

Note that Fortran coding convention would allow the Structure description field, representing the Mission ID and Logical Volume to be coded: 2A(16)& as they have the same logical structure at this level of coding. This will move some of the following fields.

INLINE DD FORM 2 - MULTILINE IMAGE, coded with INLINE STRUCTURE, FULL CORE

255 lines, 256 pixels/line. Pixel, line, and band sequence in an ancillary record. 8-bit pixels, one per byte.

LINE 1

Field	Local Rel Posn	Global Rel Posn	Format	Contents
DCDS CORE				
Size of Local ControlAuth Field	M'	0 - 1	I2	05
Local Control Authority			A4	DCDS
Unit Separator			A1	&
Size of Local Format ID Field	N'		I2	05
Local Format ID			A4	5BR3
Unit Separator			A1	&
Size of Data Start Position	S		I2	04
Data Start Position			I3	097
Unit Separator			A1	&
Size of Record Type/Subtype	T		I2	05
Record Type and Subtype			A4	DIbb
Unit Separator			A1	&
Size of Data Definition Indicator	D		I2	02
Data Definition Indicator (DDI)			A1	!
Unit Separator			A1	&
Size of Record Number Field	L		I2	02
Record Number			I1	2
Unit Separator			A1	&
Size of Record Length Field	K		I2	04
Record Length			I3	353
Unit Separator			A1	&
Size of EXTAUTH Field	A		I2	01
Unit Separator			A1	&
Size of EXTFMT Field	F		I2	01
Unit Separator			A1	&
Size of Entry Map Field	E = 1 or 5		I2	05
Entry Map	m,n,0,t		4I1	2304
Unit Separator			A1	&
Size of Packing Field	P		I2	01
Unit Separator			A1	&
Size of Record Repetition Factor	R		I2	04
Record Repetition Factor			I3	254
Record Separator			A1	;
Error Correcting Code (Example Only)			A4	AAAA
Record Separator		67	A1	;
CORE EXTENSION				
DIRECTORY A10				
Imagery Data Field				
Tag (size = t)		68 - 71	A4	IMGY
Length (size = m)			I2	19
Position (size = n)			I3	078
Field Terminator		77	A1	;
INLINE STRUCTURE A19				

Imagery Data Field			
Tag	78 - 82	A5	IMGY&
Length		A4	256&
Relative Position		A2	0&
Format Control	89 - 96	A8	256Z8R8;
THE DATA			
Imagery Data (Pixels)	97 - 351	Z256	Pixels
Terminator	352	A1	;

LINES 2 - 255 (SEPARATE RECORDS, 1 LINE EACH):

DATA			
Imagery	0 - 255	Z256	Pixels
Terminator	256	A1	;

INLINE DD - MULTILINE IMAGE, PACKED, WITH REPETITION, AND SHORT FORM INLINE STRUCTURE, MINIMUM CORE

255 lines, 256 pixels/line, packed 5 lines per block.

RECORD 1, LINES 1-5

Field	Local Rel Posn	Global Rel Posn	Format	Contents
DCDS CORE				
Size of Local ControlAuth Field	M'	0 - 1	I2	05
Local Control Authority			A4	DCDS
Unit Separator			A1	&
Size of Local Format ID Field	N'		I2	05
Local Format ID			A4	5BR3
Unit Separator			A1	&
Size of Taglist Field			I2	20
Taglist			A19	TAGLIST, RLEN,RECREP
Unit Separator			A1	&
Size of Record Length Field	K		I2	05
Record Length			I4	1345
Unit Separator			A1	&
Size of Record Repetition Factor	R		I2	04
Record Repetition Factor			I3	050
Record Separator			A1	;
Error Correcting Code (Example Only)			A4	AAAA
Record Separator		53	A1	;
CORE EXTENSION				
SHORT FORM STRUCTURE				
Image Line Structure		54 - 64	A11	5(256Z8R8);
DATA				
Image Line 1		65 - +	Z256	Pixels
Image Line 2			Z256	Pixels

Image Line 3	etc	Z256	Pixels
Image Line 4		Z256	Pixels
Image Line 5		Z256	Pixels
Terminator	1345	A1	;

NEXT RECORD, Lines 6 - 10 :

DATA

Image Line 6
Image Line 7
Image Line 8 etc
Image Line 9
Image Line 10
Terminator

Data Records repeated for a total of 255 lines

SLF RECORD WITH INLINE DATA DEFINITION - SHORT FORM

DMA SLF Data Set Identifier (DSI) Record

Field		Global Rel Posn	Format	Contents
DCDS CORE				
Size of Local ControlAuth Field	M'	0 - 1	I2	05
Local Control Authority			A4	DCDS
Unit Separator			A1	&
Size of Local Format ID Field	N'		I2	05
Local Format ID			A4	5BR3
Unit Separator			A1	&
Size of Data Start Position	S		I2	04
Data Start Position			I3	###
Unit Separator			A1	&
Size of Record Type/Subtype	T		I2	05
Record Type and Subtype			A4	DXIb
Unit Separator			A1	&
Size of Data Definitiiton Indicator	D		I2	02
Data Definition Indicator (DDI)			A1	!
Unit Separator			A1	&
Size of Record Number Field	L		I2	02
Record Number			I1	2
Unit Separator			A1	&
Size of Record Length Field	K		I2	##
Record Length			IK-1	####
Unit Separator			A1	&
Size of EXTAUTH Field	A		I2	04
External Authority			A3	DMA
Unit Separator			A1	&
Size of EXTFMT Field	F		I2	12
External Format			A11	SLF-16/5/84
Unit Separator			A1	&
Size of Entry Map Field	E = 1 or 5		I2	01

Unit Separator		A1	&
Size of Packing Field	P	I2	01
Unit Separator		A1	&
Size of Record Repetition Factor	R	I2	01
Record Separator		A1	;
Error Correcting Code (Example Only)		A4	AAAA
Record Separator		A1	;

CORE EXTENSION

SHORT FORM STRUCTURE

Data Set Identification Group (DSIG)	(A4!A5!A20!I3!A4!A4!A40)&
Security Group (DSSG)	(A4!A1!A2!A6!A21!A40)&
Data Set Parameter Group (DSPG)	(A4!A3!A3!S5!A3!A3!A3!S5!A4!A4! (I2!I2!I4!a1)!(I3!I2!I4!A1)!I10! I10!2((I2!I2!I4!A1)!(I3!I2!I4!A1))! 5I6!A40)&
Map Projection Group (DSMP)	(A4!A2!4(A10)!I9!A40)&
History Group (DSHG)	(A4!3I1!A15!2I2!A3!A8!2A10!A2!5I4! 5I1!4(2I2)!A40)&
Variable Field Address Group (DSVG)	(A4!2A5!A40)&
Registration Points Group (DSRG)	(A4!I3(I6!(2I2!I4!A1)!(I3!I2!I4!A1)! I8!3I6))&
Accuracy Subset Group (DSAG)	(A4!I2(3I4!2I2!(2I2!I4!A1)! (I3!I2I4!A1)));

DATA (650 + 51*n + 35*m)

DSIG	[According to the Structure] Lengths:	80
DSSG		74
DSPG		194
DSMP		95
DSHG		140
DSVG		54
DSRG	[n is in posn 5-7]	7 + n*51
DSAG	[m is in posn 5-6]	6 + m*35

Variable, depending upon the actual data

CASE 3 - DATA DEFINITION RECORD / DATA RECORD SET

DDR: M=5 N=4 m=2 n=3 t=3

The record to be coded will be the DMA SLF DIR. Only coding of the data groups DSIG, DSSG, DSPG, DSMP will be shown for illustration. Each data field will be treated as a compound field.

Field	Local Rel Pos'n	Global Rel Pos'n	Format	Contents
DCDS CORE A83				
Size of Local ControlAuth Field	M'		I2	05
Local Control Authority			A4	DCDS
Unit Separator			A1	&
Size of Local Format ID Field	N'		I2	05
Local Format ID			A2	5BR3
Unit Separator			A1	&
Size of Data Start Position	S		I2	04
Data Start Position			I3	154
Unit Separator			A1	&
Size of Record Type/Subtype	T		I2	05
Record Type and Subtype			A4	DXIb
Unit Separator			A1	&
Size of Data Definition Indicator	D		I2	02
Data Definition Indicator (DDI)			A1	A
Unit Separator			A1	&
Size of Record Number Field	L		I2	02
Record Number			I1	4
Unit Separator			A1	&
Size of Record Length Field	K		I2	04
Record Length			I3	418
Unit Separator			A1	&
Size of EXTAUTH Field	A		I2	08
EXTAUTH Field			A7	DMA SLF
Unit Separator			A1	&
Size of EXTFMT Field	F		I2	12
EXTFMTField			A11	SLF-16/5/84
Unit Separator			A1	&
Size of Entry Map Field	E = 1 or 5		I2	01
Unit Separator			A1	&
Size of Packing Field	P		I2	01
Unit Separator			A1	&
Size of Record Repetition Factor	R		I2	01
Record Separator			A1	;
Error Correcting Code (Example only for Illus)			A4	AAAA
Record Separator		78	A1	;
CORE EXTENSION				
8211 DDR LEADER A24				
Record Length	0 - 4	79 -	I5	00418
Implementation Level	5		I1	1
Leader ID	6		A1	L
Code Extension	7		A1	b

Reserved		8		A1	b
Application Indicator		9		A1	b
Field Control Length		10 - 11	-	A2	06
Addr of Data Descrip		12 - 16	-	I5	00154
Character Set Indicator		17 - 19	-	A3	bbb
Entry Map					
Size of Dir'y Field-Length	m	20		I1	3
Size of Dir'yField-Pos'n	n	21		I1	3
Reserved		22		I1	0
Size of Dir'y Field-Tag	t	23	102	I1	4

DDR DIRECTORY

Record ID Field					
Tag	size=t		103 -	A4	0001
Length	size=m		-	I3	016
Position	size=n		-	I3	000
DSIG Field					
Tag	size=t		-	A4	DSIG
Length	size=m		-	I3	043
Position	size=n		-	I3	016
DSSG Field					
Tag	size=t		-	A4	DSSG
Length	size=m		-	I3	040
Position	size=n		-	I3	059
DSPG Field					
Tag	size=t		-	A4	DSPG
Length	size=m		-	I3	125
Position	size=n		-	I3	099
DSMP Field					
Tag	size=t		-	A4	DSMP
Length	size=m		-	I3	040
Position	size=n		-	I3	124
Field Terminator			153	A1	;

DATA DESCRIPTION

Rec. ID Field Structure	0 - 15	154 -	A16	0000;&RECO RD ID;
DSIG Structure	16 - 58	-	A43	1000;&DSI G GROUP&&(A 4!A5!A20! I3!A4!A4!A4 0);
DSSG Structure	59 - 98	-	A40	1000;&DSS G GROUP&&(A 4!A1!A2!A 6!A21!A40);
DSPG Structure	99 - 123	-	A125	1000;&DSP G GROUP&&(A 4!A3!A3!S 5!A3!A3!A3! S5!A4!A4! (I2!I2!I4!A 1)!(I3!I2 !I4!A1)!I10

				!I10!2((I 2!I2!I4!A1) !(I3!I2!I 4!A1))!5I6! A40);
DSMP Structure	124 - 163	- 417	A40	1000;&DSM P GROUP&&(A 4!A2!4(A1 0)!I9!A40);

THE DATA RECORD

As the DMA SLF is an already-defined structure, it could be treated as an olddataset by reference to it.

However, if it were treated as a new data set with each data field treated as a vector character field, it could be coded, in accordance with the DDR above:

Field	Local Rel Pos'n	Global Rel Pos'n	Format	Contents
DCDS CORE 0 - 11 A16				
Size of Local ControlAuth Field	M'		I2	05
Local Control Authority			A4	DCDS
Unit Separator			A1	&
Size of Local Format ID Field	N'		I2	05
Local Format ID			A4	5BR3
Unit Separator			A1	&
Size of Data Start Position	S		I2	04
Data Start Position			I3	154
Unit Separator			A1	&
Size of Record Type/Subtype	T		I2	05
Record Type and Subtype			A4	DXIb
Unit Separator			A1	&
Size of Data Definition Indicator	D		I2	02
DataDefinitionIndicator (DDI)			A1	A
Unit Separator			A1	&
Size of Record Number Field	L		I2	02
Record Number			I1	5
Unit Separator			A1	&
Size of Record Length Field	K		I2	04
Record Length			I3	621
Unit Separator			A1	&
Size of EXTAUTH Field	A		I2	08
EXTAUTH Field			A7	DMA SLF
Unit Separator			A1	&
Size of EXTFMT Field	F		I2	12
EXTFMTField			A11	SLF-16/5/84
Unit Separator			A1	&
Size of Entry Map Field	E = 1 or 5		I2	01
Unit Separator			A1	&

Size of Packing Field	P			I2	01
Unit Separator				A1	&
Size of Record Repetition Factor	R			I2	01
Record Separator				A1	;
Error Correcting Code (Example only for Illus)				A4	AAAA
Record Separator			78	A1	;
CORE EXTENSION					
8211 DR LEADER A24					
Record Length		0 - 4	79 -	I5	00621
Reserved		5		A1	b
Leader ID		6		A1	D
Reserved		7 - 11		A5	bbbbb
Address of Data Start		12 - 16	-	I5	00154
Reserved		17 - 19	-	A3	bbb
Entry Map					
Size of Dir'y Field-Length	m	20		I1	3
Size of Dir'yField-Pos'n	n	21		I1	3
Reserved		22		I1	0
Size of Dir'y Field-Tag	t	23	102	I1	4
DR DIRECTORY					
Record ID Field					
Tag size=t			103 -	A4	0001
Length size=m			-	I3	020
Position size=n			-	I3	000
DSIG Field					
Tag size=t			-	A4	DSIG
Length size=m			-	I3	081
Position size=n			-	I3	020
DSSG Field					
Tag size=t			-	A4	DSSG
Length size=m			-	I3	075
Position size=n			-	I3	101
DSPG Field					
Tag size=t			-	A4	DSPG
Length size=m			-	I3	195
Position size=n			-	I3	176
DSMP Field					
Tag size=t			-	A4	DSMP
Length size=m			-	I3	096
Position size=n			-	I3	371
Field Terminator			153	A1	;
Data Area					
RecordIDField		0 - 18	154 -	A19	Data Set ID
Terminator		19		A1	;
DSIG Structure		20 - 99	-	A80	Data per SLF
Terminator		100		A1	;
DSSG Structure		101 - 174	-	A74	Data per SLF
Terminator		175		A1	;
DSPG Structure		176 - 369	-	A194	Data per SLF
Terminator		370		A1	;
DSMP Structure		371 - 465	-	A95	Data per SLF
Terminator		466	620	A1	;

(p. 113 follows)

AN INTERIM PROPOSED STANDARD FOR DIGITAL CARTOGRAPHIC DATA QUALITY: SUPPORTING DOCUMENTATION

prepared by N. Chrisman

Working Group II on Data Set Quality:

Nicholas Chrisman (chair)	University of Wisconsin
Charles Poeppelmeier (vice-chair)	Defense Mapping Agency
Wallace Crisco	Bureau of Land Management
John Davis	Kansas Geological Survey
Gunther Greulich	Survey Engineering of Boston
George Johnson	National Ocean Service
David Meixler	Bureau of the Census
Dean Merchant	The Ohio State University
George Rosenfield	U.S. Geological Survey
John Stout	Petroleum Information Corp.

GENERAL BACKGROUND

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 Phase 3 of our deliberations: developing an interim proposed standard. This work is based on previous phases, reported earlier (Chrisman, 1983; 1984). Comments on this report are requested from all interested parties. A public discussion will occur at the ACSM Annual Meeting in March 1985.

We find "quality" to be a wide-ranging concern which can cover any issue affecting the use of cartographic data. 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. These standards place a substantial responsibility on the user to evaluate the quality report to ensure fitness for the particular application.

2.0.2 Levels of testing

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.

Deduction
Deductive estimates of quality are not presumed to derive directly from sources of higher accuracy. Each instrument or procedure in a sequence of operations is expected to have a separate calibration study that yields an error distribution. A quality statement for a particular product is made by propagating the separate error effects. Extrapolation from calibration tests (or deductions of possible error magnitudes) to their combination in a particular application can be either realistic or misleading depending upon whether the components are truly independent. Some form of deductive logic is required for any test procedure which is not exhaustive. It is possible to conduct detailed tests for limited areas (or for selected sheets) and use deductive logic to extrapolate to the rest of the data. Such tests must be well-designed to be effective and the sampling design must be explicitly described. In many instances, exhaustive testing is impractical so deductive extrapolation is the only feasible alternative.

Internal evidence
A test based on internal evidence requires some form of redundancy in the data. 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, permiting 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. However, a weaker test should not be completely disregarded. For other aspects, particularly issues of data structure and logical consistency, there is no more rigorous test.

Comparison to source
A test by comparison to source examines 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.

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.

2.1 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 (digital data base). This information is required for a user to evaluate fitness, and it is required by a producer to maintain and update the data.

Digital data bases are very easily modified, particularly with current graphic work stations. Recently computer scientists have paid attention to the problem of "versioning" or how to record the modifications and different states of a data base. Currently 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 audit trail provides valuable information to the user on currency, procedures applied and sources used.

A lineage report shall contain a record of all known processes applied to the data and the source materials used. As a practical matter, it must be recognized that narrative reports of cartographic products can vary enormously in terms of detail and complexity. Some descriptive reports for NOS shoreline maps run to 26 or more pages, while many other maps only have a few cryptic marginal notes. Using a fully compliant quality report it should be possible to trace the data back to its ultimate origin (raw form). A producer can note that information for certain processes is unknown, or that sources are unknown.

A single method of describing lineage cannot be specified, due to variations in data types and structures. Lineage can be identified by appending a code to each data record or data file, as appropriate. A complete description of such codes must be a part of the quality report.

Producers of digital cartographic data shall provide users with a description of the source material from which data are derived and the methods of derivation, including all mathematical transformations involved in producing the final digital files. The description must include the vintage of the source material. Files with mixed sources or different transformations must be described at sufficient detail to identify the actual source for each element of the file. In these cases, either a lineage code on each record or a reliability overlay will be required.

In some cases, particularly for files with a single graphic source and no complex transformations, it may be possible to test the lineage statement by use of the tests of positional accuracy based on comparison to source (see 2.2.3). With multiple sources, complex transformations or other complications testing may not be feasible. In most cases, the lineage report can be assembled during the construction of the data base and should not require testing.

Reliability overlay

Quality records, including lineage and other accuracy estimates have the potential of becoming as detailed as the smallest units of the data base. Generalizations about map sheets or map series will not be sufficient for all users. A revival and extension of the reliability diagram is an alternative to complete disaggregation. A reliability overlay, registered as an overlay in a standard multi-layer data base, describes zones sharing common sources, dates and standards. These attributes would be stored as other thematic data are. The overlay could also include points (such as geodetic control) or lines (such as flight lines or traverses). As updates and modifications occur, the reliability overlay must be modified.

Documentation of transformation algorithms

In the generation of digital cartographic data bases, the correctness of transformation software is crucial. The creation of standard coordinate manipulation packages, tested by authories, and placed in the public domain, would greatly advance the profession. Until such software becomes available, there is a need to provide full information to a user on the nature of the algorithms applied to the coordinates delivered.

The standard requires a description of the transformations used, or the use of separate documentation. The later approach is clearly more efficient, in that the software is likely to be used for many different applications. In either case, the quality report must include the specific parameters applied to the specific situation.

An originator of software shall specify the nature of computational steps taken to preserve precision of coordinates through all potential transformations. The documentation should provide a set of sample computations including numerical values for coefficients to confirm the equivalence of the transformations. These specifications are intended to create a rather rigorous standard, because the current situation allows data to be described as "UTM coordinates" when there are substantially different numerical approaches actually used.

2.2 POSITIONAL ACCURACY

The Working Group recommends standards by adopting existing ones, and by modifying them for the needs of digital cartography.

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.

Geodetic standards

Ultimately the use of coordinates for geographic features implies the use of latitude and longitude with the quasi-arbitrary origin based on the Meridian of Greenwich and the Equator. All coordinates used for the transfer of digital cartographic data should bear a known (and expressed) mathematical relationship to latitude and longitude. Typically this relationship is expressed with reference to a standard datum.

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.

The Federal Geodetic Control Committee (FGCC) (1974) has established a set of standards for geodetic control. This standard is well-accepted in the United States and is consistent with similar standards elsewhere. Changing technology for geodetic surveying may alter many constraints on control. The standard is in the process of redefinition 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.

Quality of control for all digital cartographic data bases should be expressed in terms of the FGCC standards. The class of each control point used to construct a product should be noted in the lineage report. Control points, in this context, refer to horizontal triangulation stations or vertical benchmarks. If the control comes from the National Geodetic Reference Network, the points in that data base shall be identified. If 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.

Along with the standards for geodetic control, standard reference ellipsoids define horizontal datum(s) and standard geoids define vertical datum(s). Support for these standards would seem obvious if a datum were enshrined forever. The current horizontal standard is dated 1927 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 is not yet available. The vertical datum will also be changed.

The conversion of the horizontal datum will alter the definition of state plane coordinates and other projections used in digital cartography. The methods used to convert the coordinates should become part of the lineage report. Our working group supports the use of the new datum and the continual conversion of digital data bases to reflect improved control.

National Map Accuracy Standards

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. The formulation that 90% of tested points should lie within a specified tolerance does not yield an acceptable measure of the distribution 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 NMAS statement (THIS MAP COMPLIES...) does not constitute an adequate description of error nor an adequate lineage report. A more complete lineage report and description of accuracy should be included with all new cartographic products, in addition to the NMAS statements where appropriate. It is recognized, however, that the NMAS statement may constitute the sole available accuracy report for certain existing products.

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.

Alternative standards for map accuracy for large scale maps 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 test procedures and their numerical treatment are more germane to digital cartographic data standards than a particular classification scheme. We believe these concepts of testing can be extended to small scale maps.

2.2.1 Deductive estimates of positional accuracy

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. Rigorous calibration tests for all equipment are required. The explicit assumptions of error propagation between procedures should be stated. 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.

2.2.2 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.

2.2.3 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. Furthermore, the method used to register the check plot with the original is crucial to the validity of the test.

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.

2.2.4 Independent source 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. 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. The standard test for positional accuracy of digital cartographic data should be conducted using the rules prescribed by the ASP draft standard. The test can be performed on horizontal coordinates or it can include vertical information as well. The test report should indicate the results of the positional tests and the number and location of the test points. A statement of compliance to the ASP standard is inadequate by itself.

The ASP standard continues to define statistical tests to determine if the data supports statements related to specific thresholds. We do not suggest that producers complete this analysis, since we do not suggest any fixed thresholds. The ASP procedures produce an estimate of the bias (the mean deviation) and precision (reported as standard error) in each coordinate.

Reporting accuracy for less defined features

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). 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.

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).

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. 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.

The standard requires documentation of methods used to make coordinate transformations. This documentation and other information on the steps that produced the data are included in the lineage section of the quality report.

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. Achievement of "sub-pixel" accuracy in registration is not to be confused with a statement of the total error in positions, due to the roundoff implicit in cellular representation. 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.

2.3 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. These standards consider all the other attributes of a feature other than position, which must logically include the identity of the feature itself. Attribute in this section includes "feature code" and "geocode".

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 (Stevens, 1948). 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 that levels of measurement do not complicate the section on positional accuracy above. 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 attribute accuracy. Accuracy assessment for continuous measures shall be performed using numerical methods similar to those used for positional accuracy. The positional accuracy tests are designed to treat three dimensional coordinates, and the third dimension need not be physiographic. Certain spatial effects can make the problem more complex than an exercise in elementary statistics.

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. 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.

Categorical attributes

Categorical attributes are very common in digital cartography, particularly nominal attributes such as feature codes, geocodes, land use types, parcel identifiers, etc. 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.

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.

2.3.1 deductive estimates

Deduction about errors in attributes is probably more difficult than it is for positional error. Error in classification is dependent on the circumstances of the area and subject matter as much as the technology applied. Still, any estimate can provide a user with useful information to evaluate fitness for use. Even a guess, if based on the informed experience of professional personnel, can be worth recording. Of course, any guess or deduction should be carefully explained as such and not oversold.

Internal evidence
An attribute test can exploit particular characteristics of the classification scheme. A count of unlikely neighbors (obtained from internal inspection) could be used as an indicator of overall attribute error. Such tests are considered below under the topic of consistency.

Comparison to source
Attributes can be checked against a source document. In most production situations, this process leads to immediate corrections, not to statistics on accuracy.

2.3.2 Independent sources of higher accuracy
In remote sensing applications it is common to test the classification results against a source of higher accuracy (often called ground truth). The ground truth data is usually obtained for a sample of points. Sampling strategies can be devised, using the classical alternatives discussed in Berry and Baker (1968; Rosenfield and others, 1982). 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/areas should be reported as a component to the reliability overlay.

Usually the percentage correct is reported to summarize the results. This figure has the advantage of being 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.

Any study of attribute error should generate a misclassification matrix. This matrix shows the ground truth data arrayed in columns and the tested classification arrayed in rows. The percentage correct merely summarizes the diagonal of this matrix. The rest of the cells show which 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. As a minimum, a categorical attribute accuracy study shall report the misclassification matrix in count form to permit subsequent analysis.

2.3.3 Tests based on polygon overlay
Some accuracy studies can be carried out by exhaustive polygon overlay instead of point samples. Polygon overlay has the advantage of covering substantial areas, instead of picking isolated points, thus avoiding some problems of map scale. The overlay can be used to check positional differences as well as classification accuracy (as discussed above).

If polygon overlay is applied, it is desireable to know if one map is imputed to be of higher accuracy. Still, such a test can provide useful answers even if the two maps are taken as repetitions with the same accuracy.

Once the overlay is computed, the analysis of classification error follows the same structure as applied to point samples. The main difference is that overlay generates frequencies in the form of areas. Areas, being in arbitrary units, cannot be subjected to statistical manipulations that can be applied to carefully constructed point samples.

Purity and scale-related errors

A major problem with cartographic attributes is the interpretation of map 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 a polygon is examined too closely, its homogeneity 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 have small inclusions 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. An attribute accuracy test should not 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 and 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, establishing the link between purity and attribute accuracy. Any study of purity or other attribute accuracies shall be referenced in the quality report.

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.

2.4 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.

The standard does not intend to legislate specific data structures, but some are more rigorous in the quality delivered to a user. Logical consistency has been a major concern that lead some groups to adopt more complex data structures. The set of properties which can be verified in a topological system, for example, offer a substantial amount of quality information to a user. Less complex data structures will produce a less complete statement about logical consistency.

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 shall describe relationships which can be checked, given the data structure applied. The report should detail the tests performed and the results of the tests.

Permissible values

One class of tests for permissible values pertains to virtually any type of data. A test for permissible values assures that attributes are in the set of given codes or that positions are in the study area. This is a weak test, but it does trap gross blunders, 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.

Point data

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

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. 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, or flatness in lakes, or downhill drainage. 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.

A 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 may not catch on, but it does demonstrate that innovations can occur. A workable standard must be able to accomodate data structures which have not been designed yet.

Unstructured digitized lines
A large amount of digital cartographic data has been generated to serve the purposes of display - reproducing traditional products. Such a file is judged only by the products it produces; there is no logic internally that can verify that the encoding is complete and consistent.

The basic questions that apply to a graphic product can be summarized in the following list:

- Do any lines intersect only where intended?
- Are any lines entered twice?
- Are all the areas completely described?
- Are there any undershoots or overshoots?
- Are any polygons too small, or any lines too close?

Together, these questions concern logical consistency. Given unstructured digitized lines, it may be difficult to provide useful answers for the quality report. Deduction based on experience with other files does not seem useful to report the logical consistency of this type of data. The next section on topological data structures develops the tests based on internal evidence which could apply to more complex data structures.

A graphic comparison between the product and the original document mainly tests positional accuracy. It could be possible to use this technique, in a rather laborious way, to answer some of the questions posed above. Automated tests of logical consistency are possible for other data structures and thus seem much more attractive.

Vector data, topological data structure
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 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. The objects incident at each node or polygon must form a closed ring, and 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 checks proposed by White verify consistency of topological structure for cases, like DIME, where the topology was entered manually.

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.

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. When using a topological data structure, the quality report shall contain a description of the consistency checks actually applied. The report can either say that all consistency errors were corrected, or detail the remaining errors by case.

Topological encoding has become increasingly common in geographic information systems. In most systems, topologically coded data is not entered manually as the DIME files were. Turning raw lines into chains involves a process of intersection of the lines, but this check was required anyway. The intersection process not only detects intersections as potential coding mistakes, it actively inserts a new node into the data base.

Using this approach, polygon information cames from an independent source, a visual centroid with a tag. This separate source preserves the duality of DIME, but with some modifications. 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 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.

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. Systems vary widely and should be documented.

2.5 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. Completeness concerns the relationship between the whole data base and the phenomenon it is meant to model. This abstract concern has lead to two specific concerns: geocodes and relationships to a universe.

Use of standard geocodes

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 such as the FIPS codes for states, counties, and places should be adopted as appropriate. In addition, the quality report shall indicate the relationship between the geocoding scheme and the objects in the data base. Selection criteria, such as minimum width and minimum area shall be described. For example, does this file encode all census tracts within a city once and only once?

Another quality issue is the meaning of geocodes 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 geocodes should be specified so that parallel creations of conflicting codes are avoided. The toughest problems are historical consistency and maintenance into the future.

Relation of identifiers to universe

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.

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.

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. For use in geographic information systems, features should be included for consistent, defensible reasons. The quality report shall describe the nature of the objects portrayed in the data base. Relationships between these objects and a more abstract set of such entities should be discussed. In particular, relationships of hierarchy and exhaustiveness should be indicated. Selection criteria, such as minimum width and minimum area, shall be described. Methods for testing completeness have not developed far enough to be included in a standard at this time.

CURRENCY AND TIME-RELATED EFFECTS

The temporal component is a crucial element in most user's assessment of fitness. Temporal effects are often difficult to distinguish from the other components of quality mentioned above. In the report for a data base, these potential confusions should be recognized. Because each issue ties into one of the sections of the report, there is no separate section for currency.

Temporal effects on lineage

The section on lineage stressed the importance of time information. Dates of source material and subsequent modifications must be provided in the lineage report. The temporal information can be associated with entire files, but frequently, due to multiple sources, differential updating and other complications, there will be a need for more detail. A reliability overlay can assist in capturing such information without attaching a date to each object.

The date of source material is not the only date required. It would be most useful to a user to find a period of validity. This range of dates would be adjusted to reflect the speed of change in the objects mapped.

Temporal effects on positional accuracy

Tests of positional accuracy based on internal evidence or comparison to source will not be affected by temporal effects. By contrast, tests based on independent sources of higher accuracy depend on repeated measurement of the same object. These tests are superior only if the objects have not moved with time. Again, the degree of this effect varies with the processes at work. Unfortunately, given a test result, it is difficult to determine how much of the positional error should be attributed to real change instead of mapping errors. Tests based on monumented geodetic markers or property corners in areas without active earth movements may be reliable over many years, while other features such as streams, wetlands and land cover may change relatively quickly. Tests of positional accuracy based on independent sources must be designed to avoid confusing change with positional error.

Temporal effects on attribute accuracy
The temporal problems mentioned for position apply to attribute accuracy. Changes in attributes are less likely to be linked to earthquakes and other physical processes. Human impacts frequently alter the classification of objects. Whether physically or culturally caused, change can be confused with misclassification. As with position, tests based on independent sources of higher accuracy can be affected if the two sources are not simultaneous.

Quality testing should not trail on after map production. The easiest way to eliminate the temporal problem is to perform quality checking at the time of data collection.

Temporal effects on consistency
Logical consistency is an internal property of fidelity to an ideal data structure. In general, this property is not affected by changes in the real world. Problems can occur with changes to the data base. The process of updating and modifications can destroy the integrity of a data structure. Any test of logical consistency must be accompanied by a date on which it was performed. Any subsequent modifications should be flagged as unchecked, or subject to the same tests.

Temporal effects on completeness
The temporal effects on attribute accuracy can lead to effects on completeness. As new objects appear or disappear, the universe will change. The date of the definitions used should be a part of the quality report. In particular, geocoding schemes should create a new code for new objects and retire old codes. Examples include census tracts and many parcel numbering systems. By using a specific code for each object in time and space, the temporal information is automatically provided. Too many systems attach an identifier to entities which change over time and space. Some examples include counties and cities. Additional time information is needed in these cases to evaluate completeness.

REFERENCES

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

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

Chrisman, N.R. 1984, Alternatives for Specifying Quality Standards for Digital Cartographic Data: in Report 4, NCDCDS, Columbus OH

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 Information Systems, IGU, Amherst NY, p. 64-69

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

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

Federal Geodetic Control Committee, 1974, Classification, Standards of Accuracy, and General Specifications of Geodetic Control Surveys: Washington DC, GPO 1980-0-333-276

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

Merchant, D. 1982a, Spatial Accuracy Standards for Large-Scale Line Maps: paper presented at ACSM-ASP Annual Convention, 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

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

Stevens, S.S. 1946, On the theory of scales of measurement: Science, vol. 103, p.677-680

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

AN INTERIM PROPOSED STANDARD FOR DIGITAL CARTOGRAPHIC FEATURES: SUPPORTING DOCUMENTATION

Edited by Warren Schmidt

Working Group III on Features:

Warren Schmidt (chair)	Digital Mapping Unlimited
Robert Rugg (vice-chair)	Virginia Commonwealth University
Mary Clawson	IIT Research Institute
Beth Driver	Technology Service Corporation
Richard Hogan	National Ocean Survey
Leslie Kemp	Defense Mapping Agency
Joel Morrison	U.S. Geological Survey
Fred Tamm-Daniels	Tennessee Valley Authority
William Hess (observer)	Central Intelligence Agency
Robert Jacober (observer)	U.S. Air Force
Roger Payne (observer)	U.S. Geological Survey

3 CARTOGRAPHIC FEATURES

The standards recommend a comprehensive set of cartographic features and attributes representing the universe of geographical information used in digital mapping. The exhaustive list of features will encompass the present content requirement for data bases used in support of topographic, aeronautical, hydrographic and thematic mapping. The attribute listing represents probable types, conditions, categories and values which qualify, identify and quantify a given feature. The standards will also act as a cartographic data dictionary by providing common definitions for cartographic features. To facilitate query and exchange of cartographic data bases, a standard recommended alphanumeric identification code will be provided for each feature and attribute.

3.1 Descriptive Model

Many classification schemes were examined and the following five-part scheme was initially proposed:

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:

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

- o Attribute - a defined characteristic of a feature.

- o Attribute Value - a specific quality or quantity assigned to an attribute.

To accommodate users, the deleted classes were reintroduced as user options and defined as follows:

- o Feature Class - a specified group of features (e.g., hydrographic, land use, transportation)

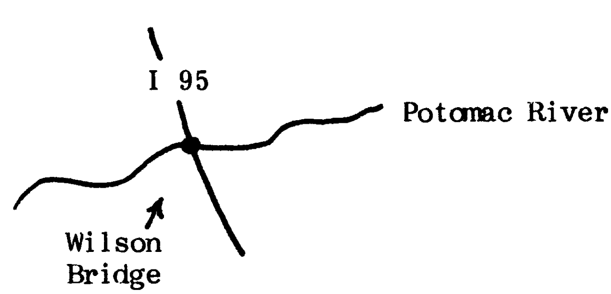
- o Attribute Class - a specified group of attributes (e.g., those describing measure, serviceability, composition, or structure.

3.1.1 Relationships

Relationships exist within the data base both between different features and an attribute and a feature. These associations are captured in four ways:

- o topology in the data base
- o the use of attributes or attribute values
- o the use of the optional feature classes to group features
- o logical structure or coding of the data set.

The first three devices are typically used to capture feature-to-feature relationships as the following example illustrates:



- o Topology identifies the point as part of three separate features--Route I95, the Wilson Bridge, and the Potomac River. These three features could then be related by virtue of colocation.

- Attributes are used in each of the following feature descriptions to relate the bridge with the road and stream without a detailed determination of colocation.

  Feature: road
  Attribute: location
  Value: coordinate string
  Attribute: name
  Value: Rt. 495
  Value: Capital Beltway
  Attribute: width
  Value: 6 lane
  Value: 60m
  Attribute: surface material
  Value: concrete
  Feature: stream
  Attribute: location
  Value: coordinate string
  Attribute: name
  Value: Potomac River
  Attribute: length
  Value: 383 miles
  Attribute: perennial

  Feature: bridge
  Attribute: location
  Value: coordinate string
  Attribute: name
  Value: Wilson Bridge
  Attribute: transportation mode
  Value: highway
  Attribute: road name
  Value: Rt. 495
  Attribute: stream
  Value: Potomac River

- Another optional means of establishing feature to feature relationships is by using feature classes, which are of interest as a collection of similar features. A common example of this is to code all transportaion features as a class. In this case, roads, bridges, overpasses and tunnels are then related because they all have the same feature class code. Other typical feature classes may include hydrography, hypsography, vegetation, and demarcation.

Attribute to feature relationships will be clearly established by the logical structure of the data base. Specifically, every feature in the data base will be related to a minimum of one attribute which is location. The structure will allow for an unlimited number of attributes. For example:

Feature: buoy
Attribute: location
Value:38^{o}45'N-76^{o}30'W
Attribute: color
Value: black
Attribute: color pattern
Value: solid
Attribute: shape
Value: cylindrical
Attribute: beacon type
Value: radar reflector
Attribute: light characteristic
Value: flashing
Attribute: light color
Value: green
Attribute:period of light (seconds)
Value: 3

Feature: municipality
Attribute: location
Value:25^{o}47'-80^{o}13'W
Attribute: name
Value: Miami
Attribute: population
Value: 852,705
Attribute: elevation
Value: 3 meters
Attribute: area
Value: 37 square miles
o
o
o
o
Attribute: Hispanic population
Value: 210,410

3.2 Cartographic Feature Definitions

A comprehensive list of features and attributes is being prepared at Virginia Commonwealth University (VCU) by five students under the supervision of Professor Robert Rugg of the VCU Department of Urban Studies and Planning. The students are Oona Przygocki (Coordinator), Carol Byrnes, Jack Green, Donna Kennon, and Shawn Smith. This work has been funded by an allocation to the ACSM National Committee for Digital Cartographic Data Standards by the U.S. Geological Survey, supplemented by a grant specifically for feature definition work allocated to VCU by the Defense Mapping Agency. A sample of the features definitions and a description of the ongoing effort appear in Appendix 3A. A sample of the attribute definitions follows this Supporting Documentation in Appendix 3B.

3.2.1 Maintenance

It is recommended that maintenance of the standard list of features, attributes, and attribute values be provided by a national body which will rule on all additions and changes to the standard. A standard will not exist for long without this clearly established national body. There are two reasons for maintenance. First, it is not possible or wise for the original standards body to attempt to provide codes for all features and attributes that might be used on maps. These features are best identified within the specialized disciplines that use them. However, if the standard is to remain useful, it will be necessary to determine that any additional set of features and attributes developed conforms to the scheme of the original list. The national body will provide that determination as well as arbitrate when proposed new lists conflict with existing or other proposed lists. Changes in mapping technology may bring new feature or attribute requirements. For example, radar reflectivity can be an important attribute of features on modern navigational charts but could not have been anticipated prior to the advent of radar.

The National Committee on Digital Cartographic Data Standards cannot establish this kind of national body but recommends that the US Board on Geographic Names or a body modeled after it be used to maintain the standard feature and attribute list. A group made up of representatives from federal mapping agencies will provide a reasonable cross-section of mapping disciplines and will lend authority to the feature codes within the federal government, the major producer of base maps within this country. If such a body cannot be organized within the federal government, the responsibility for maintenance of the standard list of features and attributes will fall upon the American Congress of Surveying and Mapping.

3.3 Cartographic Feature Codes

The assignment of codes to the list of features and attributes has been intentionally delayed pending the completion of definitions. The codes will not impose a structure upon the features and attributes but are supplied for retrieval and updating. For testing, preliminary codes will be assigned.

3.4 Alternatives

The following alternatives provide additional background both on the topics discussed above and on other issues encountered.

3.4.1 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.

3.4.2. 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.

3.4.3. 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 does not 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.

3.4.4 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 encourage compatibility and eventual adoption of the system. Attributes should be multiple and appropriately describe feature characteristics.

3.4.5 Relationships Between Features

A single spatial entity can be described by more than one feature. In one case, the relationship between these features is an aspect of definition. An example would be "island" which can be defined as a feature in its own right or as the area bounded by the feature "shoreline." The Working Group concluded that both island and shoreline are valid features. Another case is two dissimilar features sharing the same space. A river, which also serves as a boundary, is an example of a collocational relationship. Both types can be handled by either attributes or data structure. The Working Group would like to avoid the complexity of embedding these relationships in the definitions and prefers the data structure approach. Either way, it is necessary to recognize these occurences and accommodate them.

3.4.6 "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.

3.4.7 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 presented an alternative to be considered. Different levels of minimum content have been suggested, but that raised questions of scale independence and who decides. It was decided that the only mandatory attribute shall be location. Given the variety of map products and applications, any attempt to specify minimal attributes as standard would be impractical. It may be appropriate for users, however, to specify required attributes for certain products.

3.4.8 Completeness

The issue of "completeness" of a feature set 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.

3.4.9 "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.

3.4.10 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 lists, the thinking of the Working Group is that the conversion to other coding schemes would be in one direction from standard to non-standard but not the reverse.

3.5 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, Draft, I-Standards for the Classification of Topographic Features, II-Standards for the Quality Evaluation of Digital Topographic Data, III-EDP Standards Applied to Digital Topographic Data, Ottawa: Topographical Survey Division, Surveys and Mapping Branch.

Canada Statistic, 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.

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

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 International Cartography (Auto-Carto Six), abstract p. 247, Ottawa, The Canadian Surveyor, Summer 1984, Ottawa: The Canadian Institute of Surveying, pp. 91-98.

______________, 1984, "Alternatives for Cartographic Features: A Progress Report-Working Group III," in H. Moellering, ed., Digital Cartographic Data Standards: Examining the Alternatives, Issues in Digital Cartographic Data Standards, Report No. 4, Columbus, Ohio: National Committee for Digital Cartographic Data Standards, pp. 33-42.

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 Agriculture, Forest Service, 1980, Cartographic Specification's and Symbol's Handbook, Item 86-C, Washington: Government Printing Office.

U.S. Department of Commerce, Coast and Geodetic, Survey, 1963, Nautical Chart Manual, 6th Edition, Pub. 83-1, Washington, Government Printing Office.

______________, National Oceanic and Atmospheric Administration, 1978, Coastal Mapping Handbook, Washington, Government Printing Office.

______________, National Ocean Survey, 1979, Nautical Chart Manual, NOSTM No. 1 (NDM 75-01), 7th Ed, Washington, Government Printing Office.

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 Distri t 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 Defense, Defense Mapping Agency and U.S. Department of Commerce, National Ocean Survey, 1979, _Chart No. 1, U.S. Nautical Chart Symbols and Abbreviations,_ 7th Edition, Washington: Defense Mapping Agency/HTC.

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-G _Land Use and Land Cover Digital Data;_ 895-F _Geographic Names Information System;_ 895-G _Digital Line Graph Attribute Coding Standards,_ Reston, VA: USGS.

______________________________, 1983, _Feature Categories of The Geographic Names Information System,_ Reston, VA: USGS unpublished.

U.S. Department of Transportation, 1974, _A Survey of National Geocoding Systems,_ Report No. DOT-TSC-OST-74-26, Washington: Government Printing Office.

U.S. Naval Oceanographic Office, 1966, _Glossary of O eanographic Terms,_ Sp-35, 2nd Ed., Washington: Government Printing Office.

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

Wood, B., 1984, "Topographic Feature Coding," Master's Thesis, Department of Geography, University of Ottawa, Ottawa.

Wood, B. and Douglas, D., 1983, "Cartographic Feature Coding," Sixth International Symposium on Automated Cartography (AutoCarto Six), Ottawa: pp. 616-626.

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

Appendix 3B
SAMPLE ATTRIBUTE DEFINITIONS

ARTIFICIAL WATERCOURSE

Buoyed: marked with buoys used as navigation aids
Charted Depth: the recorded distance from the tidal datum to the bottom surface at a point, using an assumed velocity of sound in waters of 800 fathoms per second (U.S.) and with no velocity or slope corrections made
Commercial Shipping: travel or traffic by water vessels carrying commercial goods
Composition: the specified mixture or combination of one or more elements or ingredients
Covered: having something placed over or about another thing
Depth: the vertical measurement downward from the surface; for water features, the vertical distance from the plane of the hydrographic datam to the bed of the sea, lake or river
Drainage: the act, process or mode of draining or drawing water from a land surface
Flood control: control or drainage of a rising and overflowing body of water
Hydroelectric power: production of electricity by water power
Irrigation: the supplying of water by artificial means to land for agricultural purposes
Length: the longer or longest dimension of a feature
Lighted: marked with lights used as aids to navigation, or to general night time use
Location: situation, position in space
Name: a word or phrase that constitutes the distinctive designation of a person or thing
Navigable: deep enough and wide enough to afford passage to ships; capable of being steered
Passenger transportation: the conveyance of human passengers
Recreation: refreshment of one's mind or body after labor through diverting activity; play
Relationship to ground level: the occupation of space in relation to ground level (Values: above ground, below ground, at ground level)
Shape: spatial form
Slope: a rate of rise or fall of a quantity against a horizontal distance, expressed as a ratio, decimal, fraction, percentage, or the tangent of the angle of inclination

Volume: space occupied or cubic capacity as measured in cubic units

Waterage: the movement of goods or merchandise (such as logs) by water

Water body connection: a watercourse which acts as a link between two larger bodies of water

Water supply: the conveyance of available water for human use, or the storage of water intended for conveyance

Width: the measurement taken at right angles to the length; breadth; the measurement of the extent of something from side to side

STREAM

Branch/parent: relationship between a main stream and one of its tributaries

Direction of flow: the line or course of movement of water or lava shown by the position of one point relative to another without reference to the distance between them. The direction is usually indicated in terms of its angular distance from a reference direction.

Force of flow: the strength or energy exerted by the movement of water or lava

Salinity: (salty/brackish/fresh) the proportion of dissolved salts in pure water. Brackish water is slightly saline with a salt content less than that of sea water, sometimes defined as 15-30 parts of salt per thousand.

Perennial/intermittent: present at all seasons of the year vs. occurring or appearing in interrupted sequence

Tidal: the alternating rise and fall of water level caused by the astronomic tide-producing forces

AN INTERIM PROPOSED STANDARD FOR TERMS AND DEFINITIONS: SUPPORTING DOCUMENTATION

Edited by Harold Moellering

Working Group IV on Terms and Definitions is composed of:

Dean T. Edson. Co-Chairman	E-Quad Associates
Harold Moellering, Co-Chairman	Ohio State University
Erich Frey, Vice-Chairman	National Ocean Service
Frank Beck	U.S.Geological Survey
Mark Monmonier	Syracuse University
Hugh Calkins	State University of NY

SUPPORTING DISCUSSION FOR THE DEFINITION OF FUNDAMENTAL CARTOGRAPHIC OBJECTS

4.0 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 examination of 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. 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 section 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, Youngman (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.

Because cartographic work takes place with coordinate systems that are planar as well as nonplanar, such as latitude, longitude, these objects must be viable in both settings. Efforts have been made to define these primitive and simple cartographic objects such that they will be valid in situations where spherical or elliptical coordinate systems are being employed, as well as the more straightforward planar case.

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.

4.1 DISCUSSION OF 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. The specification proposed here has just two classes and is essentially alternative 0-2 as specified in committee Report No. 4 (Moellering, 1984). These two classes are geometry only and topology with coordinates. The point is most used in surface structure representations and nontopological substructure components in more complex cartographic objects. The node (a 0-cell) is defined to serve as a punctiform topological object that explicitely recognizes connectivity, but it also serves as a basic component for more elaborate compound objects. For the small percentage of applications that require topology only and where coordinates are not used, a truncated node is used where the coordinates are not defined.

4.2 DISCUSSION OF 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 continous lines utilize discrete elements when processed by digital systems. The specifications of the terms line segment, link, directed link, string and chain are rather straightforward and flow rather naturally from the mix of terms in current use in the profession. The term arc is where most of the ambiguity is centered. Here an arc is defined as a locus of points that forms a curve that is not closed. This specification was chosen because it 1) harmonizes with the concept in Euclidean geometry, 2) it harmonizes with the new IGES CAD/CAM standard (Parks, 1984), 3) it can be used to define objects specified by sets of various kinds of polynomial equations (e.g. polynomial expansion, Fourier series, etc.), and 4) it harmonizes with the other linear objects in the specification. It was concluded that the other use of the term arc as a topologically conceptual analog of the link was ambiguous and inappropriate in view of reasons 1 - 4 above. Please refer to the discussion in Report No. 4 (Moellering, 1984, pp. 230).

4.3 DISCUSSION OF 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. An alternative form is to define a separate primitive called a pixel. The specification of the 2-dimensional cartographic objects implies that here they are either polygons or pixels. Polygons have been defined in several ways with different kinds of objects as shown in the definitions. A 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.

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, Vol. 1, Harvard Papers on Geographic Information Systems, Reading, MA: Addison-Wesley.

Biggs, R., 1978, "Cartographic Data Structures for Thematic Applications of Automated Cartography", in G. Dutton, ed., Data Structures: Hierarchial and Overlaid, Vol. 4, Harvard Papers on Geographic Information Systems, Reading, MA: Addison-Wesley.

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, Vol. 4, Harvard Papers on Geographic Information Systems, Reading, MA: Addison-Wesley.

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., 1967, "A Computer Representation of Plane Region Boundaries", The Australian Computer Journal, Vol. 1, pp. 44-50.

________, 1978, "The Structural and Algorithmic Basis of a Geographic Data Base", in G. Dutton, ed., Data Structures: Hierarchial and Overlaid, Vol. 4, Harvard Papers on Geographic Information Systems, Reading, MA: Addison-Wesley.

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, Vol. 7, Harvard Papers on Geographic Information Systems, Reading, MA: Addison-Wesley.

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, Vol. 4, Harvard Papers on Geographic Information Systems, Reading, MA: Addison-Wesley.

Faugeras, O., M. Hebert, P. Mussi, and J. Boissonnat, 1984, "Polyhedral Approximation of 3-D Objects without Holes", Computer Vision, Graphics, and Image Processing, 25(2), pp. 169-183.

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, Vol. 7, Harvard Papers on Geographic Information Systems, Reading, MA: Addison-Wesley.

__________, 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.

Harary, 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.

Moellering, H., ed., 1984, Digital Cartographic Data Standards: Examining the Alternatives, Issues in Digital Cartographic Data Standards, Report No. 4, Columbus, Ohio: National Committee for Digital Cartographic Data Standards.

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.

Parks, C., 1984, "IGES Update, A Report from the Electrical Subcommittee", Computer Graphics World, 7(3), pp. 43-50.

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.

Reeves, W., 1983, "Particle Systems-A Technique for Modeling a Class of Fuzzy Objects", ACM Transactions on Graphics, 2(2), pp. 91-108.

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.

Wojcik, Z., 1984, "An Approach to the Recognition of Contours and Line-Shaped Objects", Computer Vision, Graphics, and Image Processing, 25(2), pp. 184-204.

Youngmann, C., 1978, "A Linguistic Approach to Map Description", in G. Dutton, ed., Spatial Semantics: Understanding and Interacting with Map Data, Vol. 7, Harvard Papers on Geographic Information Systems, Reading, MA: Addison-Wesley.

General Comments on the Work of the National Committee for Digital Cartographic Data Standards, Cycle 3

NAME: ______________________ WORK PHONE: ____________________

ADDRESS: ______________________ EVENING PHONE: ____________________

____

____

____

Your Comments Please:

SIGNED: ______________________________

DATE: ______________________________

Cut Here

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

Please
Place
Postage
Here

Prof. Harold Moellering, Chairman
Nat. Com. Dig. Cart. Data Stds.
Numerical Cartography Laboratory
The Ohio State University
158 Derby Hall
154 North Oval Mall
Columbus, Ohio 43210-1318

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