Natural Resources Canada
Geomatics Canada

Ressources naturelles Canada
Géomatique Canada

CanImage
(Landsat 7 Orthoimages at the 1:50,000 Scale)

Metadata Format

Centre for Topographic Information
Customer Support Group
2144 King Street West, Suite 010
Sherbrooke (Quebec) J1J 2E8
1 800 661-2638 (Canada and the United States)
http://www.ctis.NRCan.gc.ca

October 2003

Canada

Table of contents

1- OBJECTIVE

The information on a CanImage data set is transferred in two distinct files: One describing topographic information and the other describing the metadata. The present document describes the transfer format of the data set metadata for CanImage product.

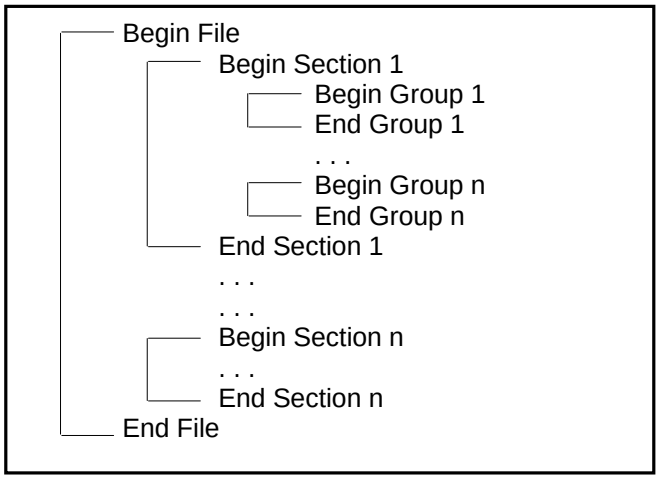
2- INFORMATION STRUCTURE

Metadata are classified into information sections. Each section has key words making it possible to determine the format and the nature of the specified value. The data are organized so as to enable the format's evolution and to ease its reading.

3- FORMAT GENERAL DESCRIPTION

3.1- File Structure

The file is structured in information sections. One section may repeat the same information group more than once (example: POLYGON section). The beginning and end of files, sections and information groups are delimited by key words (BEGIN, END). The following is the general structure of a metadata file:



3.2- Line Structure

Data are in lines and conform to the following format:

- comment column 1;
- key word columns 2 to 15;
- separator column 16;
- value columns 17 to 80 inclusively.

Information is supplied in lines by combining the key word, separator and value. Lines have a maximum length of 80 characters. The *Value* field is ended with a carriage return to position 81 or before. The «blank» characters inserted left of the *Value* are not interpreted. However, the first «non-blank» character met indicates the beginning of the value. The next characters have to conform to the format determined for this particular value. A chain of characters of «0» in length (carriage return to position 17) or a chain of «blank» characters is considered empty. Information requiring more than one line shall conform to the same format (key word, separator and value).

3.2.1- Comments – Column 1

Lines with an exclamation mark « ! » in column number 1 are considered comment lines and are not interpreted.

3.2.2- Key Words – Columns 2 to 15

The key word field contains a defined chain of characters for the purpose of identifying the nature of the specified value. The key words meet two (2) objectives: either they delimit the structure of the file or they identify in a unique fashion the specified values.

3.2.2.1- Key Words and Structure related values

The key words BEGIN and END are used to delimit the file and its sections. If more than one information group are associated to a section, they are also delimited by the key words BEGIN and END. The beginning and end of the file are combined with the FILE value. The beginning and end of a section are combined with the section name followed by the *SECTION suffix (e.g. DATA_SET_*SECTION). Finally, the beginning and end of a group of information are combined with the group name.

3.2.2.2- Key Words and Specific values

Each key word defines the nature of the value field. The key words are unique and must conform to the designated spelling.

3.2.2.3- Structure Example

```
 BEGIN          FILE
!Beginning of the data set section
 BEGIN          DATA_SET_SECTION
 ...
 END            DATA_SET_SECTION
!Beginning of the polygon section
 BEGIN          POLYGON_SECTION
 NB_POLYGONS    2
!First polygon
 BEGIN          POLYGON
 ...
 END            POLYGON
!Second polygon
 BEGIN          POLYGON
 ...
 END            POLYGON
!END of the polygon section
 END            POLYGON_SECTION
 END            FILE
```

3.3- Separator – Column 16

Column 16 is used as separator between key words and values. The «blank» character (« ») is used as separator.

3.4- Values – Columns 17 to 80 inclusively

The *Value* field contains the information to be transmitted. This field must conform to the format designated by each key word. Field lengths must be adhered to. The carriage return will be used to indicate the end of the chain of characters.

4- FORMAT DETAILED DESCRIPTION

The detailed description specifies the values and the format to be used for transferring metadata. The metadata file format is subjected to the following rules.

- **Key words:** The key word id used to clearly identify the given value and it is invariable (example: PROVINCE).

- **Value type and length:** The second information provides the numerical (N) or alphanumerical (A) type and its maximum length (examples: A(2) for two alphanumerical characters, N(2) for integers under 100, N(7.3) for a number with a maximum accuracy of three digits or decimals).

- **Description length:** The third information provides the description field maximum length (example: A(15) for 15 alphanumerical characters). This field contains a brief description that will facilitate consulting the metadata file. This description is separated from the value by a «blank» character and included between parentheses. The given length includes the parentheses. The description may be given in either English or French. Several key words have no description and are identified as having a «0» description length. The description is always optional.

- **Number of lines:** The fourth information provides the maximum number of lines that can be used for this key word (example: 4L for 4 lines maximum).

Example:

PROVINCE	**A(2)**	**A(27)**	**4L**

4.1- TERRITORY Section

The TERRITORY Section initiates the metadata file. It contains the metadata linked to the territory represented. The data must comply with the following format and order.

BEGIN	**A(30)**	**A(0)**	**1L**

The key word BEGIN is associated with the FILE value to indicate the file's beginning.

BEGIN	**A(30)**	**A(0)**	**1L**

The BEGIN key word is associated with the TERRITORY_SECTION value to indicate the beginning of the section.

NTS	**A(6)**	**A(0)**	**1L**

Territory number according to National Topographic System (NTS) sectioning for Canada at 1:50 000 scale. The NTS number format takes the form «999A99».

DATA_SET_NAME	**A(30)**	**A(0)**	**1L**

Name associated to the data set. The name of a territory is the one used for the map sheet (topographic map) corresponding to the NTS number.

PROVINCE	**A(2)**	**A(27)**	**4L**

Provinces (or Territories) ordered by decreasing order of the percentage of territory they occupy. More than one province (or Territory) may be supplied by repeating the key word and by giving different values.

ZONE_NUMBER	**N(2)**	**A(0)**	**1L**

Main UTM (Universal Transverse Mercator) zone inside of which lies the data set territory, in whole or in major part. This zone is also used as the reference to the cartographic coordinate system for the entire data set. Possible value of the UTM zone is included between 7 and 23.

PCT_OF_LAND	**N(3)**	**A(0)**	**1L**

Value included between 1 and 100 which indicates, in percentage, the surface of the territory covered by mainland (i.e. in Canada excluding only the waterbodies of coastal areas, of the Great Lakes region and other major waterbodies). The reliability degree of this value is within ± 5 %.

END	**A(30)**	**A(0)**	**1L**

The key word END is associated to the TERRITORY_SECTION value to indicate the end of the section.

4.2- DATA_SET Section

The DATA_SET section contains metadata linked to a specific data set. Each data set, identified by its NTS number and its edition/version number, has data set metadata. The data must conform to the following format and order.

BEGIN	**A(30)**	**A(0)**	**1L**

The key word BEGIN is associated to the DATA_SET_SECTION value to indicate the section's beginning.

EDITION_VERSIO	**A(5)**	**A(0)**	**1L**

Edition and version number of the data set in the «edition.version» form and «99.09» format (e.g. 1.01).

SPEC	**A(6)**	**A(15)**	**1L**

Version number of the product Standard «*CanImage: Standards and Specificatiosn»* which the data conform to (e.g. 1.0).

DATE_AVAILABLE	**A(10)**	**A(0)**	**1L**

Date of the availability of the data set. The date is expressed in the following format: YYYY/MM/DD.

MOSAIC	**A(1)**	**A(3)**	**1L**

This field indicates if the data set originate from the creation of a mosaic. The data set is a mosaic when more than one orthoimage (source data) are used during its creation. Value of one (1) indicates it is a mosaic while a zero (0) value means it isn't one.

SYSTEM_COORD	**A(3)**	**A(35)**	**1L**

This information specifies the coordinate system associated with the data set (e.g. UTM (UTM Projection)). From this information one can deduce the unit of measurement applying to the coordinates of the 4 corners delineating the data set, which corresponds to the data set minimum bounding box. When the value is «UTM» coordinates are expressed in «metres» and when the value is «GEO» they are then expressed in «decimal degrees».

CORNER_NW	**N(12.7) or N(12.3)**	**A(0)**	**1L**

Coordinates of the **Northwest** corner of the data set. Coordinate accuracy varies according to the coordinate system (SYSTEM_COORD) used. The accuracy is N(12.7) or «9999.9999999» for Geographical coordinates (GEO) while it is N(12.3) or «99999999.999» for coordinates in the UTM Projection. Regarding this field's value, the coordinate in the East-West direction is written first; it is followed by a «blank» character, and finally by the coordinate in the North-South direction. The value of this field can be of two types, for example:

- `-85.0000000 49.5000000` for SYSTEM_COORD = GEO (Geographic (Longitude/Latitude));
- `644810.000 5486058.000` for SYSTEM_COORD = UTM (UTM Projection)

CORNER_NE	**N(12.7) or N(12.3)**	**A(0)**	**1L**

Coordinates of the **Northeast** corner of the data set. Coordinate accuracy varies according to the coordinate system (SYSTEM_COORD) used. The accuracy is N(12.7) or «9999.9999999» for Geographical coordinates (GEO) while it is N(12.3) or «99999999.999» for coordinates in the UTM Projection. Regarding this field's value, the coordinate in the East-West direction is written first; it is followed by a «blank» character, and finally by the coordinate in the North-South direction.

CORNER_SE	**N(12.7) or N(12.3)**	**A(0)**	**1L**

Coordinates of the **Southeast** corner of the data set. The coordinate accuracy varies according to the coordinate system (SYSTEM_COORD) used. The accuracy is N(12.7) or «9999.9999999» for Geographical coordinates (GEO) while it is N(12.3) or «99999999.999» for coordinates in the UTM Projection. Regarding this field's value, the coordinate in the East-West direction is written first; it is followed by a «blank» character, and finally by the coordinate in the North-South direction.

CORNER_SW	**N(12.7) or N(12.3)**	**A(0)**	**1L**

Coordinates of the **Southwest** corner of the data set. The coordinate accuracy varies according to the coordinate system (SYSTEM_COORD) used. The accuracy is N(12.7) or «9999.9999999» for Geographical coordinates (GEO) while it is N(12.3) or «99999999.999» for coordinates in the UTM Projection. Regarding this field's value, the coordinate in the East-West direction is written first; it is followed by a «blank» character, and finally by the coordinate in the North-South direction.

NB_LINES	**N(5)**	**A(0)**	**1L**

Number of lines in the data set (e.g. 1855).

NB_COLUMNS	**N(5)**	**A(0)**	**1L**

Number of columns in the data set (e.g. 3710).

PCT_CLOUDS	**N(3)**	**A(13)**	**1L**

Approximate percentage of the cloud cover over the whole territory covered by the data set, calculated according to the coordinate system (SYSTEM_COORD). The value of this field is approximate and rather corresponds to a range (e.g. 10 (5-14.999 %)).

PCT_ICE	**N(3)**	**A(13)**	**1L**

Approximate percentage of the ice cover to the ground over the whole territory covered by the data set, calculated according to the coordinate system (SYSTEM_COORD). The value of this field is approximate and rather corresponds to a range (e.g. 0 (0-4.999 %)).

COMMENT	**A(64)**	**A(0)**	**8L**

Notes and remarks recorded by the team responsible for producing the data set. This field may be empty.

END	**A(30)**	**A(0)**	**1L**

The key word END is associated to the DATA_SET_SECTION value to indicate the end of the section.

4.3- POLYGON Section

The POLYGON section contains the metadata localized by polygon. The polygons identify the data sources used to generate the data set. They are enumerated by descending order according to the size of the territory covered (e.g. source orthoimage covering the largest part of the data set appears first in the POLYGON section). The group POLYGON is repeated as many times as there are metadata polygons. Key word COORDINATES is present as often as there are vertices delineating the polygon. The same scene (source orthoimage) may appear more than once in the data set but it will then cover different areas (polygons). Key word NB_POLYGONS indicates the number of polygons present in the data set. The data must comply with the following format and order.

BEGIN	**A(30)**	**A(0)**	**1L**

The key word BEGIN is associated to the POLYGON_SECTION value to indicate the beginning of the section.

NB_POLYGONS	**N(5)**	**A(0)**	**1L**

This field provides the number of polygons included in this section.

BEGIN	**A(30)**	**A(0)**	**1L**

The key word BEGIN is associated to the POLYGON value to indicate the beginning of the information group.

NO_POLYGON	**N(6)**	**A(0)**	**1L**

Sequence number identifying in a unique fashion each polygon in the data set (format 000009).

ID_SCENE	**A(6)**	**A(0)**	**1L**

Identifier of the source orthoimage. Identifier corresponds to the merging of the satellite scene track and frame numbers and takes the following form: <track><frame> (format : 099099).

EDITION_VERSIO	**A(5)**	**A(0)**	**1L**

Edition and version number of the source orthoimage in the «edition.version» form and format 99.09 (e.g. 1.01).

ACQUIS_DATE	**A(10)**	**A(0)**	**1L**

Date of the acquisition (capture) of the source orthoimage. The date is expressed in the following format: YYYY/MM/DD.

PRECISION	**N(5)**	**A(0)**	**1L**

Planimetric accuracy of the source orthoimage with a confidence level of 90 %. The precision is the degree of correspondence of the geometric data versus geodetic foundation (horizontal reference system). It is calculated by analyzing all the parameters applied to the geometric correction of the source data (for more details, see *CanImage – Standards and Specifications*). The meter is the unit of measurement of this value.

PCT_NTS	**N(7.3)**	**A(0)**	**1L**

Percentage of coverage of the data set by the source orthoimage polygon. This percentage is expressed according to the format: 999.999 (e.g. 99.999). Summation of all polygons' coverage within the data set makes 100%.

REF_CORNER_NTS	**N(1)**	**A(3)**	**1L**

Indicates if the data set reference corners coordinates (CORNER_NW, CORNER_NE, etc.) have been extracted from this source orthoimage (i.e. from this polygon). Corresponds to the source orthoimage covering the largest part of the data set. A value of one (1) indicates it is a reference while a zero (0) value means it isn't one.

NB_COORD	**N(5)**	**A(0)**	**1L**

Number of coordinates delineating the polygon.

SYSTEM_COORD	**A(3)**	**A(35)**	**1L**

This information indicates the coordinate system associated to the coordinates delineating the metadata polygon (e.g. GEO (Geographic (Longitude/Latitude))). From this information one can deduce the unit of measurement applying to the coordinates of the polygon. When the value is «UTM» coordinates are expressed in «metres» and when the value is «GEO» they are then expressed in «decimal degrees».

COORDINATES	**N(12.7) OR N(12.3)**	**A(0)**	**nL (where n = NB_COORD)**

Coordinates of a polygon vertex (point). Coordinate accuracy varies according to the coordinate system (SYSTEM_COORD) used. The accuracy is N(12.7) or «9999.9999999» for Geographical coordinates (GEO) while it is N(12.3) or «99999999.999» for coordinates in the UTM Projection. Regarding this field's value, the coordinate in the East-West direction is written first; it is followed by a «blank» character, and finally by the coordinate in the North-South direction. The polygon always represents a simple area. Coordinates of the polygon vertices are listed sequentially and coordinates of the first vertex are repeated at the end of the list in order to geometrically close the polygon. The value of this field can be of two types, for example:

- `-85.0000000 49.5000000` for SYSTEM_COORD = GEO (Geographic (Longitude/Latitude));
- `681798.384 5486058.000` for SYSTEM_COORD = UTM (UTM Projection).

END	**A(30)**	**A(0)**	**1L**

The key word END is associated to the POLYGON value to indicate the end of the information group. The information group is repeated as often as there are polygons.

END	**A(30)**	**A(0)**	**1L**

The key word END is associated to the POLYGON_SECTION value to indicate the end of the section. This line appears only once after the last information group has been written.

END	**A(30)**	**A(0)**	**1L**

The key word END is associated to the FILE value to indicate the end of the file. This is the file's last line.

5- EXAMPLE OF A DATA SET IN GEOGRAPHIC COORDINATES (LONG./LAT.)

```
 BEGIN          FILE
!
 BEGIN          TERRITORY_SECTION
 NTS            042F07
 DATA_SET_NAME  NAGAGAMISIS LAKE
 PROVINCE       ON (Ontario)
 ZONE_NUMBER    16
 PCT_OF_LAND    90
 END            TERRITORY_SECTION
!
 BEGIN          DATA_SET_SECTION
 EDITION_VERSIO 1.00
 SPEC           1.0 (Standards 1.0)
 DATE_AVAILABLE 2002/01/22
 MOSAIC         0 (No)
 SYSTEM_COORD   GEO (Geographic (Longitude/Latitude))
 CORNER_NW      -85.0000000 49.5000000
 CORNER_NE      -84.5000000 49.5000000
 CORNER_SE      -84.5000000 49.2500000
 CORNER_SW      -85.0000000 49.2500000
 NB_LINES       1855
 NB_COLUMNS     3710
 PCT_CLOUDS     10 (5-14.999 %)
 PCT_ICE        0 (0-4.999 %)
 COMMENT
 END            DATA_SET_SECTION
!
 BEGIN          POLYGON_SECTION
 NB_POLYGONS    1
!
 BEGIN          POLYGON
 NO_POLYGON     000001
 ID_SCENE       023026
 EDITION_VERSIO 1.0
 ACQUIS_DATE    2000/10/10
 PRECISION      15
 PCT_NTS        100
 REF_CORNER_NTS 1 (Yes)
!
 NB_COORD       5
 SYSTEM_COORD   GEO (Geographic (Longitude/Latitude))
 COORDINATES    -85.0000000 49.5000000
 COORDINATES    -84.5000000 49.5000000
 COORDINATES    -84.5000000 49.2500000
 COORDINATES    -85.0000000 49.2500000
 COORDINATES    -85.0000000 49.5000000
 END            POLYGON
!
 END            POLYGON_SECTION
!
 END            FILE
```

6- EXAMPLE OF A DATA SET (MOSAIC) IN UTM PROJECTION

```
 BEGIN          FILE
!
 BEGIN          TERRITORY_SECTION
 NTS            042F07
 DATA_SET_NAME  NAGAGAMISIS LAKE
 PROVINCE       ON (Ontario)
 ZONE_NUMBER    16
 PCT_OF_LAND    90
 END            TERRITORY_SECTION
!
 BEGIN          DATA_SET_SECTION
 EDITION_VERSIO 1.00
 SPEC           1.0 (Standards 1.0)
 DATE_AVAILABLE 2002/01/22
 MOSAIC         1 (Yes)
 SYSTEM_COORD   UTM (UTM Projection)
 CORNER_NW      644810.000 5486058.000
 CORNER_NE      681935.000 5486058.000
 CORNER_SE      681935.000 5457168.000
 CORNER_SW      644810.000 5457168.000
 NB_LINES       1926
 NB_COLUMNS     2475
 PCT_CLOUDS     10 (5-14.999 %)
 PCT_ICE        0 (0-4.999 %)
 COMMENT
 END            DATA_SET_SECTION
!
 BEGIN          POLYGON_SECTION
 NB_POLYGONS    2
!
 BEGIN          POLYGON
 NO_POLYGON     000001
 ID_SCENE       023026
 EDITION_VERSIO 1.0
 ACQUIS_DATE    2000/10/10
 PRECISION      15
 PCT_NTS        99.999
 REF_CORNER_NTS 1 (Yes)
!
 NB_COORD       6
 SYSTEM_COORD   UTM (UTM Projection)
 COORDINATES    644810.000 5486058.000
 COORDINATES    644810.000 5457168.000
 COORDINATES    681935.000 5457168.000
 COORDINATES    681935.000 5486030.101
 COORDINATES    681798.384 5486058.000
 COORDINATES    644810.000 5486058.000
 END            POLYGON
!
 BEGIN          POLYGON
 NO_POLYGON     000002
 ID_SCENE       023025
 EDITION_VERSIO 1.0
 ACQUIS_DATE    2000/10/10
 PRECISION      21
 PCT_NTS        .001
 REF_CORNER_NTS 0 (No)
!
```

```
 NB_COORD        4
 SYSTEM_COORD    UTM (UTM Projection)
 COORDINATES     681935.000 5486058.000
 COORDINATES     681798.384 5486058.000
 COORDINATES     681935.000 5486030.101
 COORDINATES     681935.000 5486058.000
 END             POLYGON
!
 END             POLYGON_SECTION
!
 END             FILE
```

APPENDIX A - Domain values and authorized combinations

Some fields must respect predefined domain values. When an entry contains more than one distinct parts, the domain of each part is listed along with its authorized combinations for the entry. The domain values will evolve to better respond to the change of the production environment and describe the phenomena we want to keep and classify.

TERRITOIRY Section

NTS

- Any NTS number at the 1:50,000 scale valid for Canadian territory.

PROVINCE

- AB Alberta
- BC British Columbia
- FR France
- GL Greenland
- MB Manitoba
- NB New Brunswick
- NF Newfoundland
- NS Nova Scotia
- NT Northwest Territories
- NU Nunavut
- ON Ontario
- PE Prince Edward Island
- PQ Quebec
- SK Saskatchewan
- US United States
- YT Yukon Territory

ZONE_NUMBER

- Whole or integer value included in the domain [7,23].

PCT_OF_LAND

- Value included in the domain [0,100].

DATA_SET Section

SPEC

- 1.0 Standards 1.0

MOSAIC

- 0 No
- 1 Yes

SYSTEM_COORD

- GEO Geographic (Longitude/Latitude)
- UTM UTM Projection

PCT_CLOUDS

- Value included in the domain [0 10 20 30 40 50 60 70 80 90 100].

PCT_ICE

- Value included in the domain [0 10 20 30 40 50 60 70 80 90 100].

POLYGONS Section

PRECISION

- Whole or integer value included in the domain [1,999].

PCT_NTS

- Value included in the domain [0.001,100].

REF_CORNER_NTS

- 0 No
- 1 Yes