RESEARCH NOTES
12
Canadian Forest Service
Laurentian Forestry Centre

Importance of Metadata in Forest Research:
Example of the Lac Édouard Experimental Forest

Claude Delisle and Louis Archambault

Introduction

A metadata describes other data in paper or electronic format without disclosing the raw data. The purpose of a metadata is to provide information on data in order to better assess their quality and validity and make appropriate use of them. A metadata usually focuses on the history of data, from the time they are measured until their entry in a computer, as well as the accuracy and location of the measured data. Metadata are crucial to any organization because they make data comprehensible and shareable for future users. At the Laurentian Forestry Centre of the Canadian Forest Service (CFS), a metadata was created in 2002 to compile information relative to forest inventories and silvicultural activities conducted at the Lac Édouard Experimental Forest.

The objective of this research note is to briefly explain what a metadata is and demonstrate its importance by using the example of the Lac Édouard Experimental Forest.

Lac Édouard Experimental Forest

The Lac Édouard Experimental Forest is located within the confines of the La Mauricie National Park and covers an area of almost 15 km^2. The Experimental Forest was established in 1918 by the Conservation Commission, the Laurentide Pulp and Paper Company and the federal Department of Agriculture. It is therefore one of the oldest sample plot networks in eastern North America. The plot was created in order to study growth, mortality and post-harvest regeneration. In 1925, a network of 321 sample plots was established. The sample plots measuring two square chains (809.4 m^2) were distributed systematically at intervals measuring 10 chains (210 m) by 10 chains between the plots. In 1936, the sample plot network was remeasured and, while the spacing between the cruise lines was maintained, the plots were resized to one square chain. The sample plot network was measured again in 1946. An experimental partial cut was carried out between 1950 and 1957. Following these cuts, measurements were taken in the summers of 1956 and 1957. Some sample plots were added to bring the total number up to 401. Ten years later, in 1967, the sample plots were measured again. The sample plot network was then abandoned for about 30 years until Université Laval's Department of Forestry and Geomatics resumed monitoring of the sample plot network in 1994. In 2001, the CFS worked together with the Department to locate and measure the plots. The research work was successful since 201 plots had been found and remeasured by September 2005. The history of the Lac Édouard Experimental Forest and the measurements taken there over the years are described in a metadata.

The Lac Édouard Experimental Forest metadata

Data gathering at the Lac Édouard Experimental Forest in 1955. (Photo: L.E. Ray)

Although there are several international standards for metadata geographic information, such as those of the International Organization for Standardization (ISO), the Federal Geographic Data Committee (FGDC) and the European Committee for Standardization (CEN), none of them is recognized as an official international standard and set up to describe research sample plots. Nowadays, various user communities make up for it by using a wide variety of metadata standards. The Lac Édouard Experimental Forest metadata was modelled on these international standards in order to meet the specific need of keeping structured information on this sample plot network for long-term use. This metadata, whose contents are available in French only, includes extensive information to help understand a body of data and attempts to answer the following questions:

- What does the data table contain?
- How is it organized?
- Who entered the data and in what context?
- What were the devices used to take measurements?
- What types of processing did the data undergo (screening, sorting, transformation, etc.)?
- What is the quality of the data (degree of accuracy and possible errors)?

Data storage

A key aspect of the metadata is how the data tables are organized and described. All of the available data from the Lac Édouard Experimental Forest since its creation have been stored in a Microsoft Access relational database. Because the measurement data are varied, they are stored in specific tables, including tables containing forest inventory data, allometric measurements, positioning of stems and saplings in the sample plots, regeneration measurements along with light measurements, pedological data, soil mineral analyses and topographical data. To make the tables easier to understand, each field bears a descriptive name for the measured variable. For example, the "Tige" (Stem) table contains dendrometric data and is made up of the following fields:

Tige													
ParcId	TigeId	TigeNum	TigeEssId	TigeEtatId	DHP_pouce	DHP_mm	Hauteur_m	Mesure_imp	Mesure_dcm	Age_1m	Age_dhp	Age_souche	TigeEtageId
2772	186524	25	41	1	16	406	22.60	Non	Non	-99	174	-99	5

Table 1 includes a description of each of these fields.

Table 1.

Description of the "Tige" table fields

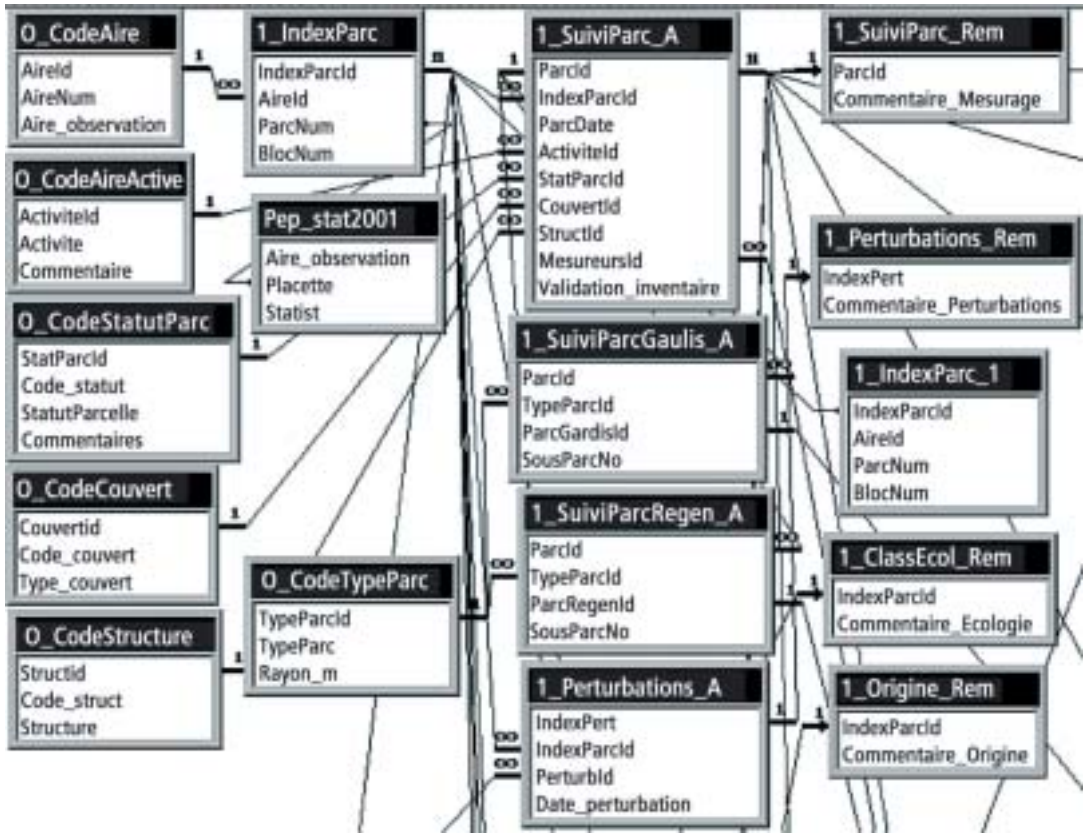
Field Name*	Data Type	Description
ParcId	Numeric	Specific number of the sample plot
TigeId	AutoNumber	Specific number of the measured stem
TigeNum	Numeric	Number of the stem in the sample plot
TigeEssId	Numeric	Species code
TigeEtatId	Numeric	Stem status code
DHP_pouce	Numeric	DBH in inches
DHP_mm	Numeric	DBH in millimetres
Hauteur_m	Numeric	Stem height in metres
Mesure_imp	Yes/No	If tree was measured in 1-inch classes
Mesure_clcm	Yes/No	If tree was measured in 2-cm classes
Age_1m	Numeric	Age measured at 1 m
Age_dhp	Numeric	Age measured at DBH
Age_souche	Numeric	Age measured at stump height
TigeEtageId	Numeric	Stem storey code

* Field names are given in French only, as the contents of the metadata are in French.

There are over 100 tables in the overall database. The "Tige" table alone contains nearly 116,000 records. In all, the database contains more than 84 Mb. The tables are joined by permanent relationships. Some of these links are shown in Figure 1.

Figure 1.

Representation of some of the relationships between tables in the Access database.



The development of the Lac Édouard Experimental Forest metadata required considerable effort to recover the historical data and compile the information accumulated over the past 75 years. Over five months of work was required to develop this metadata containing over 135 pages. However, the annual updates only require a few days of work. The relational database took more than four months to develop, not counting the time required to enter the data.

Because the creation of a metadata requires an amount of effort that many might find discouraging, particularly because of the large quantity of detailed information required for certain standards, it becomes essential to stick to key information that best describes the data. In that regard, the drafting of a form can make it easier to enter this key information and make the task less onerous.

Because software applications change and sometimes even disappear over time, it is recommended that the metadata and data tables be saved in text format so that they can always be accessed in the coming years. The digital maps, graphs and photographs included with the metadata should be saved in an uncompressed bitmap format. The metadata and data should be stored in the same place. Lastly, given the rapid changes in storage media and peripherals, it is necessary to ensure that the records are transferred to these new storage media.

Conclusion

Any organization or individual involved in and allocating funding to data gathering should create a metadata so that other people can easily reuse the data. This is particularly important in forest research because the impact of natural and human disturbances on ecosystems can only be assessed over the long term.

References

Delisle, C. 2005. Métadonnée de la Forêt expérimentale du Lac Édouard, Parc de la Mauricie. 137 p. (Available in French only, contact the author to obtain a copy).

Federal Geographic Data Committee. 1998. Content Standard for Digital Geospatial Metadata (FGDC-STD-001-1998). http://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata (7 February 2006).

This Canadian Forest Service publication is part of a series that aims to distribute the results of forest research in a concise and timely manner.

Please send your comments and suggestions to:
Pamela Cheers, Head, Publications
Laurentian Forestry Centre
1055 du P.E.P.S., P.O. Box 10380
Stn. Sainte-Foy, Québec,
Quebec, Canada G1V 4C7
Tel.: (418) 648-5253 Fax: (418) 648-3354
E-mail: pcheers@nrcan.gc.ca

For more information:

Claude Delisle
CFS, LFC, 1055 du P.E.P.S., P.O. Box 10380
Stn. Sainte-Foy, Québec, Quebec, Canada G1V 4C7
Tel.: (418) 648-4918
Fax: (418) 648-5849
E-mail: claude.delisle@nrcan.gc.ca

This publication is also available in electronic format on the LFC Web site at: www.cfl.cfs.nrcan.gc.ca

Cette publication est également disponible en français.

Catalogue number Fo113-2/12-2006E
ISBN 0-662-42808-0
ISSN 1209-0379