

CONSIDERATIONS REGARDING THE USE OF SHAPEFILE DATA FORMATS

SIGEOM geomatics data are available in Esri's FGDB and Shapefile formats. While a Shapefile is an open, easy-to-use data exchange format that can store geographic and descriptive informations (<http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>), its inherent limitations must be considered. Despite the time and effort involved in producing consistent data in this format, they remain incomplete. Shapefile format does not allow to fully translate the complexity of the relational data model.

Indeed, annotations, attribute relationships, topology relationships, attribute domains, coordinate precision and other capabilities are not supported by this data format.

Furthermore, Shapefiles make use of the dBase file format (.dbf file) to store attribute tables. This data format, developed in the 1980s, is now obsolete because data representation improvements have been made since then, such as the Unicode standard, to support most of today's writing systems. This explains why Shapefiles are not optimal for storing information in a language other than English.

Also, shapefiles do have certain limitations with attributes. The main issues are:

- Null values cannot be stored;
- Numbers are rounded up;
- Unicode character strings are not supported;
- Field names longer than 10 characters are not allowed;
- The maximum record length is only 4,000 bytes;
- The maximum number of fields is only 255, and so on.

It is therefore important to consider these limitations depending on the type of use which will be made of SIGEOM data. For data queries and background display, Shapefiles are appropriate, as well as our [WMS Services](#). However, if the purpose is to process and manipulate data for analysis, then FGDB is the most appropriate format. For more details, please refer to « How may I use SIGÉOM data? » under the [FAQ](#) section.