**6.2 Structure**

**6.2.1 Data format**

Today, all geospatial data are stored in commonly accepted specialized data formats. Those formats have a specific structure and include metadata. Some of the formats are standardized by ISO and/or IEC, others are de-facto standards.

A geospatial dataset shall be archived together with a full documentation of its data format.

One of the key components of sustainable long term preservation is detailed knowledge of the file format that houses data. Some file types are well understood by the technical user communities (e.g. TIFF v6) while other less common file format types suffer from significantly less technical comprehension.

The documentation of the data format shall include the format structure, its properties, the metadata, and eventually a means of interpreting the data to access and interpret the data.

In all cases it is desirable to ensure that the primary file format is recorded in a consistent and comprehensive way, ideally linking via unique identifiers to well established file format registries.

In consideration of experience, cost, and resources, an implementation of this standard shall link to existing and acknowledged format-registries of which an example is PRONOM [13].

The use of a specified profile for any given file format is highly desirable knowledge to long term preservation, and as such it should be meticulously recorded, along with pointers towards any conformance and/or validation tools and methods used to assert the quality of the profiled file.

An analogue graphic representation of file content shall be archived where feasible.

**6.2.2 Data base**

Most of the geospatial data are object-structured and stored in data bases. In order to preserve this structure a simple storage of the data is not sufficient. In any many of the cases, the archive requires the software to access and interpret the data. This is named representation information. Consequently, the full data base content shall be transferred to the archive, which demands an archiving strategy that allows a persistent understanding of the technology for accessing this dataset.

Unless this is possible, the content of objects is modified and thus a decision about the future use of the archived data required before archiving.

**6.2.3 Properties of geospatial data**