PERIODICA POLYTECHNICA SER. CIV. ENG. VOL. 43, NO. 1, PP. 107–116 (1999)

GIS DATA EXCHANGE PROBLEMS, SOLUTIONS

Zoltán SIKI

Department of Surveying
Technical University of Budapest
H–1521 Budapest, Hungary

Received: June 1, 1998

Abstract

The present situation in Hungary in the field of GIS data exchange is introduced in this paper. Most of the data exchanges concentrate on map data (graphical data) without attributes.

The new Hungarian standard DAT1 (MSz 7772-1) has changed the situation. It involves map, topology and attribute data as well and is the official data exchange format for Land Offices. The new standard probably effects most of the GIS community as the existing digital maps should be converted to DAT1 data exchange format. A solution is given in this paper. Finally the demand of a more general data exchange format is discussed.

Keywords: GIS, digital maps, data conversion.

1. Preface

Why is GIS data exchange so important nowadays? Today's GIS users' biggest investment lies in three areas:

- data conversion,
- development of custom application-specific extensions to general- purpose GIS products,
- learning of use of the software and the data.

The main advantage of using GIS standards implies optimization of the use of the hardware/software, people and data resources. Standards not only solve the data exchange between systems and organizations but also reduce the training period of employees.

Although standards are changed over time, they afford users a more stable production environment and help reduce dependence on the capabilities of individual vendors. Even when standards are obsolete, new standards must usually take older ones into account by allowing for upwards compatibility. This protects the user's investment in his data and makes the integration of GIS more productive in long-term run.

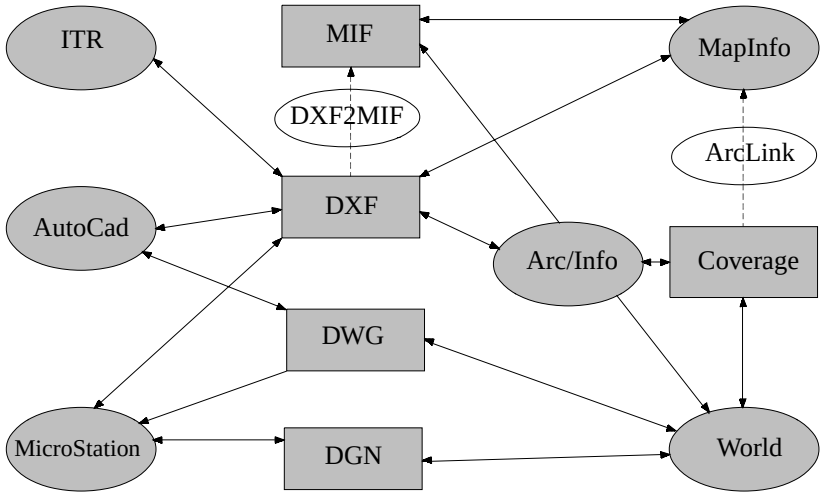

Fig. 1. Possible import/export connections among mapping, GIS software

2. Overview of Mapping and GIS Software Used in Hungary

During the last decade I came across several map data exchange and conversion problems. I have written some programs to solve system-to-system map data conversion. Besides summing up this experience I overview the Hungarian Map/GIS data exchange demands and possibilities.

Several different software are used in Hungary for digital mapping purpose. The main user groups are the land offices and surveyors, the local municipalities, the public utility providers.

In *Table 1* there are three CAD-like software (ITR, AutoCAD, MicroStation), which have sophisticated construction functionality. They are mainly used to digitize, construct digital maps.

The other group involves GIS software which have good analysis tools but they are weak in construction functions.

The digital maps can be easier created with the software in the first group. But once the digital maps is created, the software in the second group gives better functionality for the user to display and analyze the map and attributes.

The most popular mapping/GIS software are the following:

Table 1. Some features of mapping, GIS software used in Hungary

	ITR[1]	AutoCAD	Micro Station	ARC/INFO, ArcView	MapInfo	Autodesk World
Operating system	MS DOS	Win NT	Win NT, Unix	Win 3.x, Unix, Win NT	Win 3.x, Win NT, Mac OS	Win NT
Topology	no	yes	no	yes	no	yes
Cleaning	no	yes	no	yes	no	yes
Programming environment	no	AutoLisp, ARX	MDL, macro	AML, SML, Avenue	MapBasic	VBA, C++
Address matching	no	no	no	yes	yes	no
Export-import formats	**DXF**, ascii	**DXF**, IGES	**DXF**, DWG	**DXF**, IGES, TIGER, E00	**DXF**, MIF	DWG, DGN, coverage, SHP, MIF
Database	no	DBF, ODBC	DBF, Oracle, ODBC	INFO, DBF, ODBC	DBF, ODBC, MapInfo	MS Access, DBF, ODBC
Automatic polygon creation	no	no	no	yes	no	yes
Thematic map	no	yes	no	yes	yes	yes
Shortest route	no	yes	no	yes	no	no
Raster support	yes	yes	yes	yes	yes	yes
Best feature	simplicity	construction functions	construction functions	analysis	simplicity	data integration

[1] ITR is Interactive Mapping System made by DigiCart Ltd Hungary specially for surveyors and Land Offices

DXF is the only commonly supported data exchange format. Autodesk World does not support DXF format directly, you need AutoCAD or AutoCAD Map to convert DXF files to DWG (AutoCAD drawing).

The first group of this software is used not only for mapping purpose, AutoCAD and MicroStation are general CAD software. ITR is a relatively cheap and simple Hungarian software with no attribute capabilities at all. AutoCAD/AutoCAD Map and Microstation are used with and without attributes, too. One of the largest AutoCAD/AutoCAD Map users is the Hungarian Telecom (MATÁV). MicroStation is used in the integrated cadastral system of the Land Offices financed by PHARE. This software – ITR, AutoCAD and MicroStation – has similar data model based on layers and entities. The DXF (Drawing eXchange Format) is supported by all of them. Only few but probably very annoying incompatibilities can arise exchanging 2D map data among these systems using DXF format.

DXF is the native data exchange format of AutoCAD, so if a DXF file cannot be imported, it is the fault of the software which generated that file. The AutoCAD releases have their own DXF version. But all AutoCAD releases can load the DXF file of previous versions.

The main problem to import DXF data into MicroStation is the limited number of levels (64 layers). The 2D/3D entities can be loaded into MicroStation without any problem by an integrated converter.

In the case of ITR there are more differences. ITR does not support polyline entities so they are broken into line segments importing the DXF file. The coordinates are stored as fixed point numbers (cm precision). It can cause problems when we have arcs in the DXF file. In the DXF file the center point, the radius and the start and end angle are stored for an arc. ITR uses start and end points instead of angles. Due to rounding errors we can get unpredictable result after the DXF import, e.g. connected arc and line become disconnected. It will be an error in topology.

The second group of programs in the table is used and offered only for mapping purpose. The Arc/Info, the ArcView and MapInfo have DXF import/export modules which support zero and one-dimensional entities of DXF files (e.g. points, inserts, lines, polylines, arcs, circles, texts). The arcs (circular arcs) in Arc/Info are converted into line segments during the DXF import. Double lines, intersections can be eliminated automatically by Arc/Info commands. Another difference in Arc/Info is that we cannot have points and polygons in the same coverage. ArcView can display DXF/DWG files directly without conversion.

In GIS/LIS systems the areas are more important than in CAD/mapping software. It is not accepted to store only the border of parcels in such a system. No thematic maps can be generated for parcels, analysis tools (polygon overlay) cannot be used if the border is stored only. Arc/Info can generate polygons from the border lines but no overlapping polygons are allowed in the same coverage. Arc/Info has topological data model. Topology is also built for polygons and lines by Arc/Info.

The spaghetti data model is used in MapInfo (no topology) and no tools are provided to generate polygons from borders. MapInfo will generate polygon from

closed polylines in the DXF file. It is not enough to have the same start and end point for the polyline but the 'closed' flag must be set in the DXF data. External programs can solve this problem. I developed such a program which converts the DXF file into MIF (MapInfo Interchange Format) file and creates areas from border lines. It can be used if the polygons to be generated have no overlapping parts. So it is useful, for example, for parcels, buildings, country borders, etc.

The third and the latest one, the Autodesk World has no DXF support. It uses another approach to handle foreign data formats. This software provides drivers to load/save the data files of the most popular GIS/CAD software. AutoCAD Map has similar capabilities. Autodesk World can load and save AutoCAD drawing (DWG), MicroStation design (DGN), MapInfo interchange format (MIF/MID), Atlas GIS (BNA), Arc/Info coverage and ArcView shape (SHP) files. Attributes data are also maintained in the case of Arc/Info, ArcView, and MapInfo files. For DWG and DGN files the entity identifiers are reserved, so data bases can be connected through these identifiers.

Until we have widely accepted international standards (like DXF for CAD data) the solution offered by Autodesk World is the most comfortable for the end users. Of course the conversion never can be perfect because of the incompatibilities of graphical elements and data models of the different systems. World is also able to build polygons from borders.

The DXF2MIF program was specially written to build polygon from polygon borders stored in DXF file. This program can add a text inside the polygon as an attribute in the MIF file. It bridges the gap between line oriented (DXF) and area oriented (MIF) representations. ArcLink is an extra tool offered to MapInfo to import Arc/Info coverages into MapInfo.

3. New Demands in Map and Attribute Data Conversion

Since the last year an Act dictates to the local municipalities and public utility providers to use the official map – created and updated by the Land Offices – as a basis for their GIS/LIS systems. The Land Offices must use the Hungarian standard (DAT) data exchange format. It must supply data to customers and accept data from customers in this format. The DAT standard and its data exchange file contains the following groups of relational data tables:

- Geometric tables (7),
- Topological tables (11),
- Object tables (36),
- Data quality parameter tables (12),
- Object code tables (4),
- Collector tables (15),
- Code tables (64).

These nearly 150 tables are written sequentially into an ASCII file for data exchange. This is not a general GIS data exchange format. It is specific to Land Offices, with fixed table structure and it is not extendible. No commercial software exists yet to create, manipulate data sets conforming to the standard. Some converters are under development based on the mentioned software. Besides these I developed my own software to convert existing DXF files into DAT format. I choose the AutoCAD DXF file format as a start point because it is the most widespread in Hungary and most of the existing maps are available in this format.

Unfortunately there are many differences in logic and content between DAT data exchange format and DXF format. In the DAT format graphic is not separated into layers and no graphic element (points, line strings, borders, surfaces) can be repeated, e.g. there is only one table for each different graphic element. For example, if a parcel has a common border with a building, that line is stored only once and two surfaces refer to that line. In the DXF file usually it is represented by two separate lines. The digital maps created with CAD software usually have no areas, only the border of areas are drawn. So from the lines, polylines of DXF file surfaces should be built, surfaces can have islands, too. DXF format stores no topologic information while topology is part of the DAT format.

I broke the creation of DAT data exchange file into two steps. First the geometric, topologic and object tables are created from the DXF file by a non interactive program. It fills fully the geometric and topologic tables but it cannot fill the object tables fully because a DXF file cannot store all the necessary information. After the first step the connection between object and graphic elements can be established and one text by object can be put into the object tables. For example, if the parcel numbers are put on the map the parcel number column in the object table of parcels can be filled by that text automatically. The transformation of DXF files to DAT data exchange format is based on several parameters. The user can set up parameters using a parameter file. At the end of this step a pre DAT file is created, not all the attributes are filled.

Features of the first step are:

- elimination of points using snap tolerance,
- elimination of double lines,
- check of topology,
- automatic creation of surfaces (islands are allowed),
- creation of line strings between nodes,
- creation and filling of geometric and topologic tables,
- creation and partial filling of object tables.

Spatial index is generated and used during this step to speed up the process. The program is available for Unix, DOS and Windows NT/95 operating systems.

During the second step an interactive program (GeoEasy DAT) is used to display DAT files and to fill, to update object tables. Graphic cannot be changed here, only the attributes. A separate program is used at this step, so any DAT data exchange file can be displayed and manipulated not only those created at the first

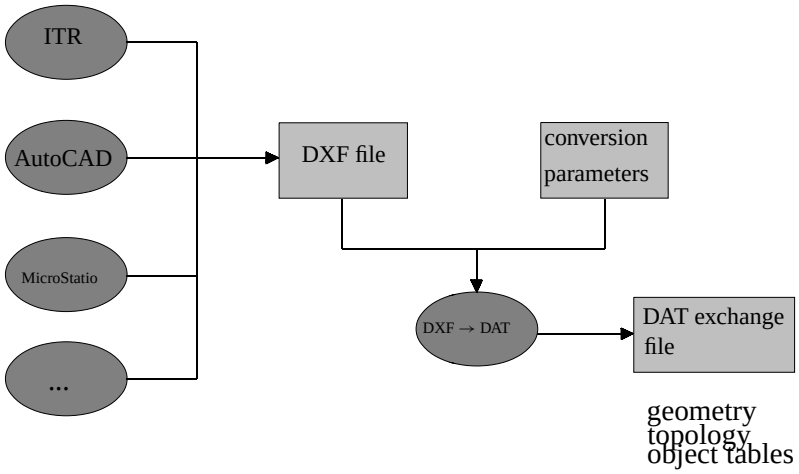

Fig. 2. Data flow of DXF to DAT conversion

step. The ASCII DAT data exchange file is converted to an internal format GeoEasy graph file and database tables for more effective process. After finishing editing attributes, the internal format data are exported to DAT data exchange format file again.

Features of GeoEasy DAT program are:

- graphic display of DAT files,
- zoom in, zoom out,
- maintaining attributes graphic connections,
- interactive editing of object tables,
- form oriented database editing.

The GeoEasy DAT program is available for X Windows, Windows 3.x, Windows NT and Windows 95. This program is used to demonstrate the new standard to the students and to help to create and check their mid-term assignments.

4. Further Demands in GIS Data Exchange

Hungary made an important step in standardization of GIS data by announcing the DAT standard. But the standard that will satisfy a wider user expectation and create such a spatial data infrastructure will cover a wider field than just data exchange. Some efforts were made to involve attribute data of other organizations than Land

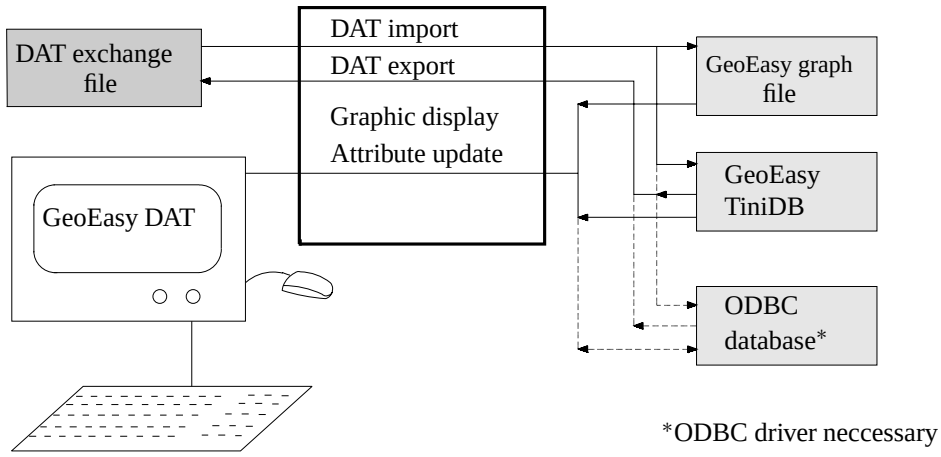

Fig. 3. Data flow diagram of GeoEasy DAT

Offices in DAT but it gives weak result. There are a lot other national standards similar to our standard with rigid structure. Nowadays a GIS data exchange standard implies much more than defining tables and columns. There are big gaps in logic among the commercial GIS software and data sets used by different user areas. No fixed structure can bridge these gaps. Efforts are made to define descriptive language for data exchange purpose. EXPRESS is such a language. A committee was set up in Hungary to develop a more flexible data exchange model. This committee plans to use EXPRESS definitions, and their work is based on European Norms for Geographic Information (CEN TC287).

The European Norms for Geographic Information (CEN TC287) represents the effort of EC countries on standardization in the field of digital geographic information. It is divided into several chapters such as

- CEN ENV 12009 Geographic information – Reference Model,
- CEN prENV 12160 Geographic information – Spatial Schema,
- CEN prENV 12656 Geographic information – Quality,
- CEN prENV 12657 Geographic information – Metadata,
- CEN prENV 12658 Geographic information – Transfer,
- CEN prENV 12762 Geographic information – Position,
- CEN CR 12660 Geographic information – Query and Update,
- CEN prENV 12661 Geographic information – Geographic Identifiers.

This group of inter-related standards provides techniques for developing schemas for geometric primitives and constructs, quality information, directory information and dictionary information. Techniques for mapping from conceptual schemas to

transfer schemas are also defined. The Query and Update part will define a query language that incorporates appropriate spatial operators and identification systems for querying and updating geographic data, including metadata. The conceptual schemas for the CEN standards are defined using the EXPRESS language defined for STEP in Part 11 of ISO 10303.

The Spatial Data Transfer Standard reflects similar efforts in the United States. SDTS is a language for communicating spatial information. SDTS was developed to allow U.S. Federal agencies to share spatial data among applications, which use different hardware, software, and operating systems.

SDTS is designed to support all types of spatial data. However, developing a single translator that could support all of the SDTS options and all the varieties of spatial data models is not practical. What is practical is implementing SDTS through the use of profiles. The first profile developed was the Topological Vector Profile or TVP. The TVP is for use with geographic vector data with planar graph topology. The second profile developed was the Raster Profile, which can accommodate image data, digital terrain models, gridded geographic information system layers, and other gridded data. The Raster Profile applies to data that is georeferenced (data that has been geographically registered to the Earth's surface) and (or) geocoded (georeferenced spatial data representing digital spatial entities).

Environmental Systems Research Institute (ESRI) has included in Version 7 of the ARC/INFO product an SDTS translator. Besides the commercial products there are some public domain packages. The SDTS++ is a C++ class library developed by the USGS to be used in SDTS application development. It was released in February, 1998 supporting both read and write SDTS files. It currently runs under Windows 95/NT, and was built using Microsoft Visual C++. This library is located at:
http://mcmcweb.er.usgs.gov/sdts/sdtsxx/index.html. There are some translators for SDTS data at ftp://ftp.blm.gov/pub/gis/sdts.

SDTS to mif (MapInfo) (sdts2mif.exe - 04/15/97),
SDTS to dxf (AutoCad) (sdts2dxf.exe - 04/15/97),
SDTS to DLG (sdts2dlg.exe - 07/15/97),
SDTS2ARC (ARC/INFO) (sdts2arc.exe - 03/25/98).

5. Conclusion

Drawing eXchange Format (DXF) is an industry standard exchanging 2D map data among systems. Most of the software used in Hungary support this format. The possible connections among the software used in Hungary are presented.

The new DAT GIS data exchange standard is obligatory in the data exchange with the Land Offices. The DAT standard involves attributes and topology besides geometric data. New converters have to be developed to import and export GIS data sets into this format. A solution is introduced in this paper to convert existing maps from DXF file into the new Hungarian standard.

There are several national and international standards. The latest ones try to create intelligent data sets using descriptive language. In this case not only the data is involved in the data exchange file but the information about the interpretation of the data, too.

References

[1] LOFFREDO, D.: Efficient Database Implementation of Express Information Models. Ph.D Thesis, http://www.uspro.org.
[2] STEP standard (ISO-10303).
[3] Graphics file formats, http://www.wotsit.org.
[4] GIS related file formats, http://www.geocities.com.
[5] Spatial Data Transfer Standard SDTS, http://nsdi.usgs.gov.
[6] US FIPS 173 standard, http://mcmcweb.er.usgs.gov/sdts/standard.html.
[7] DAT1 szabályzat, Hungarian Standard 24.459-1996.
[8] STUBKJAER, E.: GIS Exchange Standards – From Danish DSFL Towards European Standards, http://www.civil.auc.dk.
[9] COLEMAN, D. J. –. MCLAUGHLIN, J. D.: Standards for Spatial Information Interchange: A Management Perspective. Department of Geodesy and Geomatics Engineering, University of New Brunswick, http://fas.rugers.edu.