

## MOŽNOSTI KONVERZE DAT MEZI FORMÁTY MICROSTATION A ARCGIS A TVORBY DATOVÉHO SKLADU S VYUŽITÍM PROSTOROVÝCH DATABÁZÍ POSTGRESQL A MS SQL

Jiří Ardielli<sup>1</sup>, Jiří Horák<sup>2</sup>, Jiří Kostelníček<sup>3</sup>, Milan Krmíček<sup>4</sup>

<sup>1</sup>Institut geoinformatiky, Hornicko-geologická fakulta, Vysoká škola báňská – technická univerzita Ostrava, 17. listopadu 15, 70830, Ostrava-Poruba, Česká republika  
ard01@vsb.cz

<sup>2</sup>Institut geoinformatiky, Hornicko-geologická fakulta, Vysoká škola báňská – technická univerzita Ostrava, 17. listopadu 15, 70830, Ostrava-Poruba, Česká republika  
hor10@vsb.cz

<sup>3</sup>Moravské naftové doly, a. s., Úprkova 807/6, 69501, Hodonín, Česká republika  
kostelnicekj@mnd.cz

<sup>4</sup>Moravské naftové doly, a. s., Úprkova 807/6, 69501, Hodonín, Česká republika  
krmicekm@mnd.cz

**Abstrakt.** Problémem, který může nastat při nasazení GIS/CAD aplikací různých výrobců, může být vzájemná nepřenositelnost souborových formátů mezi těmito aplikacemi. Obecným řešením této situace může být použití jednotného datového skladu s využitím prostorové databáze. Do této databáze by pak bylo možno společně s použitím konverzních programů ukládat vícero souborových formátů a ty následně exportovat v různé podobě.

Příspěvek se zabývá možnostmi zpětných konverzí mezi souborovými formáty DGN aplikace MicroStation a formáty Shapefile firmy ESRI. Dále se pak zabývá analýzou možností tvorby datového skladu, s využitím prostorové databáze PostgreSQL/PostGIS nebo MS SQL, který by byl v souvislosti s touto konverzí využíván. V rámci použití prostorové databáze MS SQL jsou pak řešeny možnosti využití nástroje ArcSDE firmy ESRI.

Z charakteristiky výše zmíněných formátů souborů pramení problematika jejich vzájemné konverze. Při samotné konverzi je totiž nutno řešit otázku parametrické a neparametrické konverze. V rámci souborového formátu DGN existuje velké množství typů geometrických tvarů, které jsou vyjádřeny parametricky. Soubor DGN může rovněž obsahovat veliké množství hladin. K požadavkům zadavatele patřila i schopnost pracovat se všemi typy geometrických tvarů, které formát souboru DGN může obsahovat, přitom využívat všechny atributy spojené s jednotlivými typy geoprvků daného souboru. Dále bylo zapotřebí nalézt způsob, jak tyto geoprvky a informace ukládat do prostorové databáze a zpětně z nich rekonstruovat původní soubory formátu DGN. V případě souborového formátu ESRI Shapefile lze využít pouze čtyř typů geoprvků - bodů, linií, polygonů a anotací. Při vzájemných konverzích je nutno zohlednit problematiku hladin a typů geometrických tvarů, jejich atributů a mapování atributů mezi DGN a ESRI Shapefile soubory.

**Klíčová slova:** MicroStation, DGN, ESRI, Shapefile, konverze, databáze.

**Abstract.** Possibilities of data converting between MicroStation formats and ArcGIS formats and creation of data storage based on PostgreSQL and MS SQL spatial databases.

The problem, which can occur during usage various types GIS/CAD applications of various producers, can be mutual non-portability of files format between this applications. The universal solution of this situation can be the usage of uniform data storage with exploitation of spatial database. Together with usage of converse application, there would be possibility to store of multiple file formats into this database, and then to export them into different representation.

This article considers of backward converting possibilities between DGN file formats of MicroStation application and Shapefile file formats of ESRI Company. Next the paper occupy by analysis of possibility of data storage creation, with usage of PostgreSQL/PostGIS spatial database and MS SQL spatial database, which would be used in context of data conversion. Within the frame of MS SQL spatial database usage are solved possibilities of ArcSDE instrument utilization.

From the characteristics of above mentioned file formats arises the question of their mutual converting. By the single converting it is necessary to solve the question of parametric or nonparametric converting. Within the frame of DGN file format there exist many types of shapes, which are notated parametric. The DGN files can also contain a lot of levels. One of the submitter's requirement was to work with all of spatial types, which can DGN file contain, and exploit all of attributes linked with several types of geo items of a given file. Then the objective was to find method, how to store this geo items and information about them to the spatial database and how to export them and retroactively create preferably original DGN file. In case of ESRI Shapefile file format, there exist only four types of geo items - points, lines, polygons and annotations. By the mutual converting it is necessity to take into account problems of levels and types of shapes, their attributes and mapping of this attributes between ESRI Shapefile and DGN file formats.

**Keywords:** MicroStation, DGN, ESRI, Shapefile, converting, database.

## 1 Úvod

Studium problematiky převodu mezi souborovým formátem ESRI Shapefile a MicroStation DGN, za využití prostorové databáze MS SQL nebo PostgreSQL/PostGIS, je rozsáhlá co do množství způsobů, jak tato data konvertovat, pomocí kterých nástrojů, a jakým způsobem data předávat SŘBD. Otázku konverze těchto souborových formátů a jejich uchování v prostorové databázi lze vnímat v několika rovinách. Je nutno navrhnout způsob a nástroje, kterými budou jednotlivé formáty souborů konvertovány, eventuálně do jaké intermediální podoby budou převedeny, jakým způsobem se bude jejich obsah ukládat do prostorové databáze a dále pak za pomoci jakých prostředků bude probíhat analýza nad prostorovými daty. Z pohledu možností užití výše zmíněných SŘBD se jeví jako ideální použití souborového formátu ESRI Shapefile jako vstupně/výstupního formátu na straně databází. Použití obou uvedených databází totiž mají jedno společné - bezproblémový import/export z/do formátu ESRI Shapefile. Proto byl tento formát zvolen za cílový při ověřování DGN

konverzí (co lze uložit do formátu ESRI Shapefile, lze uložit i do MS SQL/ArcSDE nebo PostgreSQL/PostGIS).

## **2 Charakteristika souborových formátů a komponent**

### **2.1 Souborový formát ESRI Shapefile**

Formát ESRI Shapefile je otevřený, plně dokumentovaný digitální vektorový formát geografických dat. Tvoří jej kolekce minimálně tří souborů s příponami .shp, .shx a .dbf, které obsahují samotné geometrické prvky, indexaci těchto prvků pro rychlé vyhledávání a tabulku formátu dBASE obsahující atributy a hodnoty pro jednotlivé prvky. Shapefile je popisován primitivními prostorovými prvky, jako jsou body, linie a polygony. Každý z těchto prvků může být charakterizován množinou atributů. V jednom Shapefile se může vyskytovat množina pouze jednoho typu prvku. V GIS aplikacích je potom možno tyto Shapefile obsahující různé typy prvků na sebe vrstvit. [4]

### **2.2 Souborový formát MicroStation DGN**

Základním formátem výkresů v produktech firmy Bentley Systems, jako je např. MicroStation, je CAD formát označovaný jako DGN. Původní formát, který vycházel z Integraph Standard File formátu, dostal přívlastek V7 v době, kdy se objevila nová generace produktu MicroStation V8 a s ní i nový formát DGN V8 s odlišnou interní strukturou dat. DGN V7 je datový formát určený pro uložení technické dokumentace - výkresů, map, schémat, 3D modelů apod. Jedná se primárně o vektorový formát, ačkoli v něm mohou být uložena i rastrová nebo popisná data. Některé definice nutné pro správné zobrazení dat jsou uloženy externě (mimo soubor), konkrétně písma a ve většině případů i uživatelské styly (druhy) čar, měrné jednotky, 2D nebo 3D dimenze, verze souboru a přesnost souřadnic. Prostorové prvky, které nabízí tento formát, lze rozdělit na dvě skupiny, na geoprvky definované neparametricky a geoprvky definované pomocí parametrického vyjádření. Každý typ geoprvku je pak charakterizován jinou množinou popisných dat. Formátem je podporováno maximálně 32 vrstev. [5]

### **2.3 PostgreSQL/PostGIS**

PostgreSQL/PostGIS je Open source řešení prostorové relační databáze organizace Refraction. Programy obsažené v distribuci zajišťují konverzi ESRI Shapefile souborů do SQL (resp. PostGIS) a inverzní transformaci PostGIS dat do Shapefile souborů (pgsql2shp a shp2pgsql).

PostgreSQL je plnohodnotným relačním databázovým systémem s otevřeným zdrojovým kódem. Předností systému PostgreSQL je rozšiřitelnost. Systém může být bezproblémově rozšiřován o nové datové typy, funkce, operátory, agregační funkce, procedurální jazyky. Díky tomu mohly vzniknout rozšíření, jako je např. PostGIS - podpora pro geografické informační systémy.

PostGIS je implementace podpory prostorových dat pro RDBMS PostgreSQL dle standardu OpenGIS. Vychází z SQL92 rozšířeného o geometrické typy (SQL92 with Geometry Types), definuje metadata popisující funkcionality systému co se týká geometrických dat, a definuje datové schéma.

## 2.4 ArcSDE

ArcSDE (Spatial Database Engine) je klient/server technologie firmy ESRI pro ukládání a správu prostorových dat v prostředí zvoleného relačního databázového systému (MS SQL Server, Oracle, Informix, IBM DB2). Takto uložená data pak dodává do jiných aplikací. Funguje tedy jako tzv. middleware, neboť zprostředkovává požadavky mezi klientem a serverem. Mezi přednosti této technologie patří možnost ukládání prostorových dat, víceuživatelská editace dat a možnosti analýzy prostorových dat. Pro nativní souborový formát, se kterým umí ArcSDE manipulovat, tedy ESRI Shapefile, existuje množina příkazů, tzv. Data management commands. S použitím těchto příkazů je možno provádět zpětné konverze mezi formátem ESRI Shapefile a ArcSDE. [1]

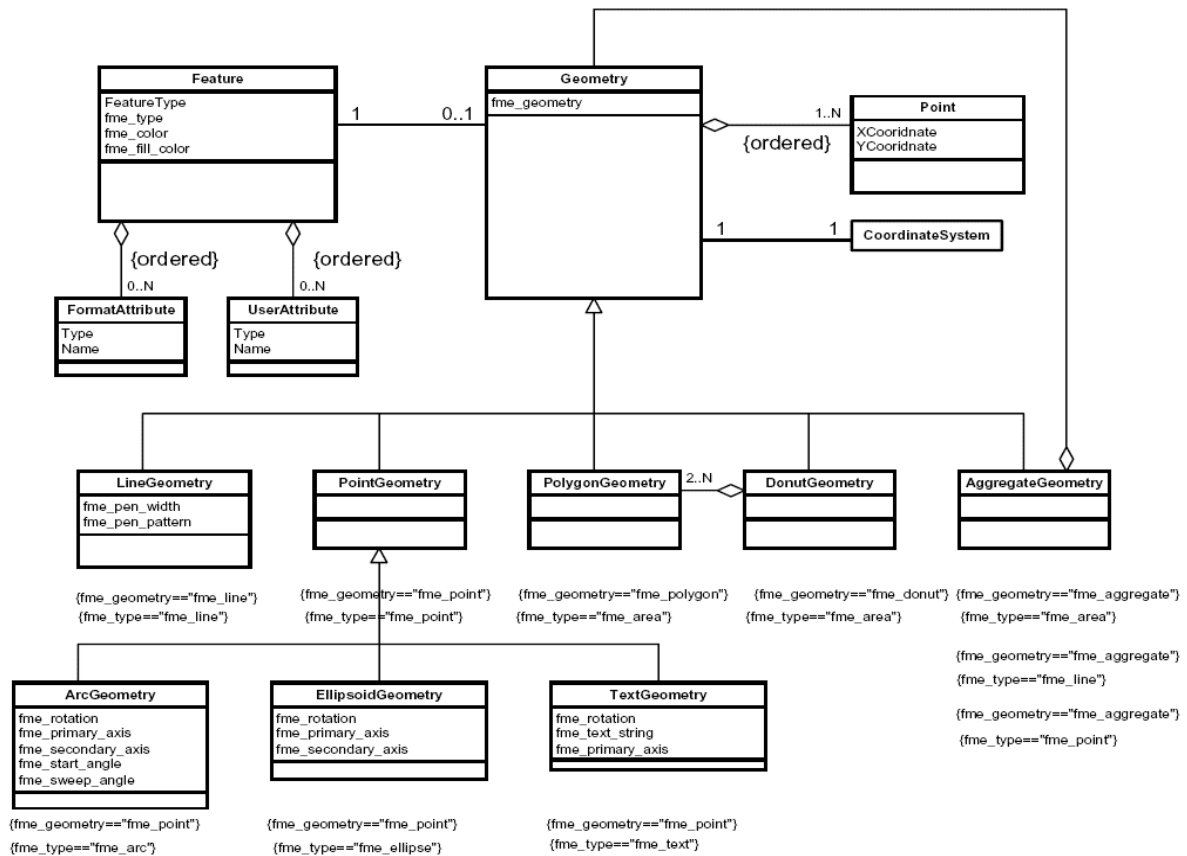
## 2.5 ArcGIS Data Interoperability

Jedná se o další produkt firmy ESRI, který je nadstavbou pro ArcGIS Desktop produkty. Rozšiřuje funkcionalitu o další možnosti přímého čtení a využití více než 60 vektorových formátů. Produkt byl vyvinut ve spolupráci se společností Safe Software na základě jejího populárního produktu FME. Mezi podporované datové formáty se řadí formát DGN, ESRI Shapefile a ESRI Geodatabase (SDE). Jako desktopová komponenta sebou nese i nevýhodu v nemožnosti jejího nasazení do architektury klient/server na straně serveru a vazbu pouze na produkty firmy ESRI. [2]

## 2.6 FME Suite

Prostředí „Feature manipulation Engine“ (FME), vyvinuté firmou Safe Software, umožňuje převod a transformaci formátů z oblastí GIS, CAD jako např. ArcSDE, ArcInfo či Integrgraph a databázových formátů. Pro snadné použití je možné jej přímo integrovat do prostředí GIS aplikací, jako např. ArcGIS, ArcView či Integrgraph GeoMedia. Existuje několik verzí FME, které se od sebe liší množstvím podporovaných vstupních a výstupních formátů. Ve verzi Professional Edition je možno provádět konverzi mezi souborovými formáty DGN a Shapefile. Z charakteristiky souborového formátu Shapefile se jedná o neparametrickou konverzi, tedy převod parametricky vyjádřených geoprvků souborového formátu DGN na jejich neparametrickou podobu uchovanou ve formátu souboru Shapefile. Výhodou této konverze je možnost bezproblémového využití formátu Shapefile jako vstupně/výstupního pro prostorové databáze. Nejnižší verze produktu, která podporuje práci s ArcSDE, je verze ESRI Edition. Ta rovněž podporuje práci s formáty DGN a Shapefile. Pomocí nástroje FME ESRI Edition lze formáty DGN i Shapefile prostřednictvím rozhraní ArcSDE ukládat do prostorové databáze MS SQL v parametrické podobě. FME poskytuje mimo desktopových konverzních programů také API (Application Programming Interface) pro programovací jazyky jako C++, Java, .NET atd., díky čemuž lze vytvářet serverové aplikace umožňující provádět požadované konverze. [3] [6]

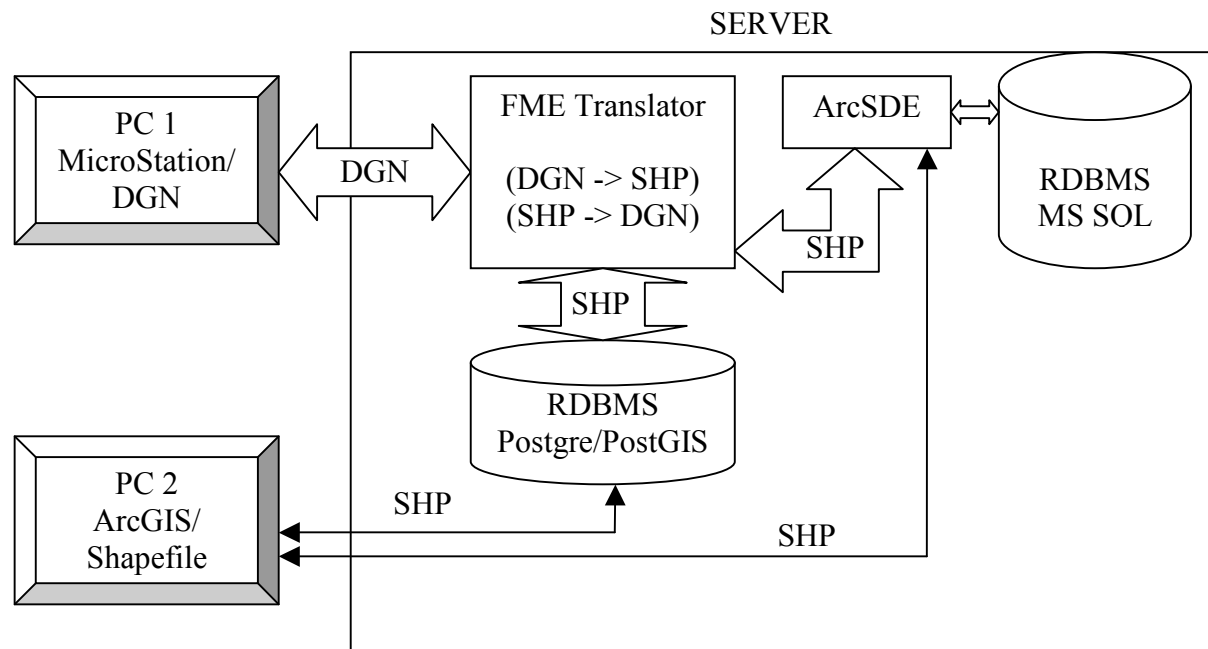
FME Suit používá vlastní obecný objektový model (obr. 1), do kterého je chopen transformovat data z formátů, které podporuje. Takto načtená data je možné modifikovat a uložit do formátů do kterých je FME Suite schopen zapisovat. V rámci projektu je potřeba načíst data z formátu DGN, tato data uložit do vybraného formátu a následně potom provést obrácený postup.



Obr. 1. Objektový model FME Suit.

### 3 Architektura systému

S existencí počítačových sítí je vhodné nasazení architektury typu klient/server. Toto řešení umožňuje centralizovaně uchovávat data, následně je spravovat a dále pak nad těmito daty provádět prostorové analýzy a potřebné konverze. Klientům se pak dostává výsledků serverově zpracovaných dat. Vhodným nasazením takové architektury odpadají problémy, jako je redundance dat nebo nekonzistence. Jádrem na straně serveru je tedy databázový stroj, který uchovává prostorová data. V případě použití PostgreSQL/PostGIS prostorové databáze je možno provádět analýzy nad daty s použitím komponent, které jsou součástí tohoto řešení. Za předpokladu využití MS SQL jako úložiště dat, by bylo využito rozhraní ArcSDE jednak jako konverzního nástroje pro převod ESRI Shapefile formátu do úložiště a zpět, a jednak jako nástroj umožňující prostorovou analýzu dat (obr. 2). V případě zpracování dat na straně klientů a nasazení desktopových aplikací různých výrobců by bylo zapotřebí řešit jednotlivé kroky zpracování dat a konverzí dat pro každou aplikaci zvlášť.



Obr. 2. Architektura systému.

#### 4 Možnosti konverzí

Cílem této části bylo ověřit možnosti zpětných konverzí souborů ve formátu DGN V7 do prostorové databáze PostgreSQL/PostGIS a MS SQL/ArcSDE. Důraz byl kladen na získání pokud možno identického DGN souboru po provedení zpětné konverze z prostorové databáze, v porovnání s DGN souborem, který byl do této prostorové databáze importován. Jedná se zejména o zachování stylů kresby čáry, jejich barvy, tloušťky a zachování textové anotace.

Klíčovým problémem při převodu z DGN souboru bylo nalézt vhodný nástroj podporující práci se všemi typy geometrií, které formát DGN může obsahovat, a získat všechny atributy spojené s jednotlivými geoprvky daného souboru. Dále bylo nutné nalézt způsob, jak tyto geoprvky a informace uchovávat tak, aby byly jako vrstvy zobrazitelné v ESRI produktech a zároveň aby bylo možné zpětně sestavit původní DGN soubor.

Konverze DGN -> ESRI Shapefile za použití serverové aplikace naprogramované nad API FME spočívá v těchto hlavních krocích:

- zjištění atributů, hladin a typů geometrie, které obsahuje daný DGN soubor
- vytvoření ESRI Shapefile souborů podle jednotlivých hladin a typů geometrie v DGN souboru (každá hladina a každý typ geometrie je ukládána zvlášť do ESRI Shapefile souboru)
- vytvoření atributů v ESRI Shapefile souborech nesoucí nezbytné informace ke zpětné konverzi do DGN
- mapování atributů mezi DGN a ESRI Shapefile soubory
- provedení konverze

Geometrické elementy evidované ve formátu DGN mají řadu předem definovaných atributů (např. u liniového prvku je možné specifikovat více než 70 atributů), nicméně prakticky se při

záznamu využívá jen několik atributů jako je tloušťka čáry, barva, hladina. Formát ESRI Shapefile umožňuje u jednotlivých geoprvků definovat libovolné atributy. Vzhledem k tomu, že není možné před samotným převodem zjistit, které atributy byly u jednotlivých DGN elementů použity, bylo by nutné je transformovat všechny. Při převodu DGN -> ESRI Shapefile byly tedy do SHP atributů mapovány jen nejčastěji se vyskytující DGN atributy.

Zpětná konverze ESRI Shapefile -> DGN obnáší postup podobný, přičemž ještě musí být zajištěno správné kódování znakové sady všech atributů, neboť v těchto attributech je v názvech stylů obsažena i diakritika, a hodnoty by v případě nesprávného kódování nebyly identické s hodnotami v Seed File původního DGN souboru.

Formát DGN V7 umožňuje ukládat i geoprvky, jejichž geometrie je vyjádřena parametricky - např. kružnice. Formát ESRI Shapefile však umožňuje ukládat pouze geoprvky definované lomovými body - proto bylo přistoupeno v průběhu konverze k transformaci parametricky definovaných prvků na liniové segmenty o zadané přesnosti.

Výsledkem konverzí mezi DGN souborem s neparametricky vyjádřenými prvky a ESRI Shapefile souborem a zpět, bylo dosažení kýžené shody mezi vstupním a výstupním DGN souborem. U DGN souborů obsahujících parametricky definované geoprvky byly tyto převedeny na liniovou nebo polygonovou reprezentaci s konečným počtem úseků.

Specifika pro uložení dat v jednotlivých prostorových databázích:

#### PostgreSQL/PostGIS

Tento způsob uložení nenabízí v souvislosti s konverzí DGN formátu žádné výhody, které by oproti ESRI Shapefile formátu vedly k odstranění nežádoucí transformace parametricky vyjádřených geoprvků na jejich reprezentaci s konečným počtem úseků. Do prostorové databáze PostgreSQL/PostGIS lze tedy ukládat pouze prvky, které jsou reprezentované neparametricky. Zpětná konverze do formátu DGN tedy nevede k vytvoření souboru s původním formátem dat.

#### MS SQL/ArcSDE

Tento způsob uložení umožňuje ukládat i všechny parametricky definované prvky definované formátem DGN V7. V případě konverze DGN -> MS SQL za použití ArcSDE by už tak nebylo zapotřebí provádět generalizaci parametricky vyjádřených prvků, ale tyto by se ukládaly přímo v parametrickém vyjádření. Výsledek zpětné konverze do DGN formátu by tak byl identický i v tomto ohledu. Pro dosažení tohoto výsledku by bylo nutno využít FME Esri Edition a jeho API.

## 5 Závěr

Na základě provedených analýz a testů a studia problematiky lze konstatovat, že shody při zpětné konverzi DGN -> prostorová databáze -> DGN lze dosáhnout, a to za využití nástroje FME ESRI Edition a jeho API. Pomocí této konverze a využití ArcSDE je zajištěno ukládání parametricky definovaných geoprvků do prostorové databáze MS SQL. Takto uložená data lze rovněž zobrazovat v produktech firmy ESRI. Naopak využití prostorové databáze PostgreSQL/PostGIS pro datový sklad se jeví jako méně výhodnější varianta z důvodu nemožnosti ukládání parametricky definovaných geoprvků obsažených v souborech formátu DGN a tedy nemožností zpětného sestavení původního obsahu souboru DGN. Původní záměr konverze souborů formátu MicroStation DGN na formát ESRI Shapefile a zpět za použití aplikace pracující nad objektovým modelem FME Professional Edition a jejich import do

prostorových databází PostgreSQL/PostGIS nebo MS SQL/ArcSDE byl prakticky ověřen. Výsledky takto konvertovaných DGN souborů s neparametricky vyjádřenými prvky byly uspokojivé. Naopak zpětná konverze DGN souborů, které obsahují parametricky definované prvky, nezajišťovala sestavení původních DGN souborů z důvodu charakteru formátu souboru ESRI Shapefile použitého jako intermediálního formátu.

## Reference

1. Arcdata Praha, s.r.o. <http://www.arcdata.cz>.
2. Arcdata Praha, s.r.o. <http://www.arcdata.cz/download/doc/2006/Data-Interoperability.pdf>
3. CSMAP, s.r.o. <http://www.csmap.cz/>.
4. ESRI. <http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>.
5. GISOFT, v.o.s. <http://www.gisoft.cz/Bentley/DGN>
6. Safe Software. <http://www.safe.com/>.

## Annotation:

*Possibilities Of Data Converting Between MicroStation Formats And ArcGIS Formats And Creation Of Data Storage Based On PostgreSQL And MS SQL Spatial Databases*

On the base of performed analyses and tests and studies of problems it is possible to claim, that it's possible to get agreements on reverse converting of DGN -> spatial databases -> DGN by usage FME ESRI Edition tool and its API. Through the use of this converting and utilization of ArcSDE is guaranteed storage of parametrical defined geo items into spatial database MS SQL. Data deposited by this way can be used by products of ESRI company. On the contrary usage of spatial database PostgreSQL/PostGIS as a data storage appear as less advantageous alternate by reasons of impossibility of storage parametric defined geo items contained in DGN file format and inability of backward assemblage of original DGN file content. The primary intention of MicroStation DGN file format conversion into ESRI Shapefile format and rearwards by disposition of application worked over FME Professional Edition object model and its import into PostgreSQL/PostGIS or MS SQL/ArcSDE spatial databases was practically attested. Converting results of DGN files containing non-parametric phrased geo items was satisfactory. On the contrary backward converting of DGN files containing parametric defined geo items, do not ensure assemblage of original DGN files by reasons of ESRI Shapefile format character used as intermediate format.