

Infrastruktura prostorových dat geografické sekce Přírodovědecké fakulty UK v Praze

Jan Kolář, Stanislav Grill, Markéta Potůčková, Tomáš Bayer, Michal Lodin,
Přemysl Štych, Tomáš Hudeček, Jan D. Bláha, Eva Štefanová, Miroslav Čábelka

Katedra aplikované geoinformatiky a kartografie
Přírodovědecká fakulta, Univerzita Karlova v Praze
Albertov 6
12843, Praha 2, Česká republika
jkolar@natur.cuni.cz

Abstrakt. Stále intenzivnější využívání GIS ve výzkumu a výuce geografie a příbuzných oborů přineslo potřebu vytvoření infrastruktury prostorových dat (SDI), která by umožnila sdílení těchto dat. Nezbytnou podmínkou je vytvoření standardů a jejich dodržování. Článek se zabývá koncepcí SDI navrženou Katedrou aplikované geoinformatiky a kartografie, zejména pak technickým řešením datového skladu a mapového serveru.

Klíčová slova: geoinformatika, SDI, datový sklad, mapový server.

Abstract. Intensive application of GIS in research and education of geography and related fields has brought a need for establishing a spatial data infrastructure (SDI). It enables sharing data and it is based on standards that have to be kept. The article discusses the concept of SDI as it has been designed at the Department of Applied Geoinformatics and Cartography, especially the technical solution of the data warehouse and the map server.

Keywords: geoinformatics, SDI, data warehouse, map server.

1 Úvod

Geografické informační systémy (GIS) již nejsou brány jen jako nástroje pro úzkou a velmi specializovanou skupinu odborníků, ale jsou nedílnou součástí téměř všech geografických, tj. k prostoru vztahených studií, analýz a projektů. GIS mohou být velmi užitečným a efektivním nástrojem, ale pouze za předpokladu existence vhodných dat požadované kvality a možnosti přístupu k těmto datům.

Katedra aplikované geoinformatiky a kartografie (KAGIK) se v rámci výzkumu a výuky zabývá sběrem, zpracováním, vizualizací a distribucí prostorových dat. Aktuální potřeba sdílení datových zdrojů s ostatními katedrami geografické sekce vedla k požadavku na vytvoření infrastruktury prostorových dat (dále SDI – Spatial Data Infrastructure). V následujícím textu budou popsány hlavní stavební kameny předpokládané SDI. Vzhledem k využívání prostorových dat i v oborech mimo geografickou sekci, jako např. geologie či biologie, se nechá předpokládat budoucí rozšíření navržené infrastruktury na úroveň fakulty či celé univerzity.

2 Potřeba SDI v rámci geografické sekce

Z pohledu práce s prostorovými daty na geografické sekci PřF UK v Praze je třeba vyzdvihnout dobré softwarové vybavení, jak v oblasti GIS, tak dálkového průzkumu Země. Softwarová řešení jsou ve valné většině síťová, umožňující hromadné využití v počítačových učebnách či přístup z osobních PC akademických pracovníků. Oproti tomu skladování dat samotných je značně roztržité jak v rámci geografické sekce jako takové, tak jednotlivých kateder. Jelikož data jsou z převážné většiny získávána v rámci výzkumných projektů, jsou následně spravována řešitelem projektu. Získání informace, zda určitá data v rámci sekce existují, v jakém formátu a za jakých podmínek je lze využít pro výuku či další projekt, je časově náročné. Současně to může vést k duplicitě dat a zbytečným výdajům.

Na podzim 2006 byla na KAGIK zřízena přijímací stanice dat EUMETSAT umožňující příjem družicových snímků nízkého prostorového, ale vysokého časového rozlišení (s intervalem až 15 min), což představuje vysoké nároky na správu dat. To je jen další z důvodů, že se KAGIK rozhodla pro vytvoření SDI, která by umožnila efektivní správu a přístup k prostorovým datům různého typu. Za tímto účelem je třeba nalézt odpověď na otázku jaká data budou spravována (typ, formát, objem dat), jak budou data spravována (datový sklad), jak bude vyřešen přístup k datům a jak budou tato data prezentována (mapový server).

3 Struktura SDI

3.1 Spravovaná data

Prostorová data používaná pro výzkum a výuku jednotlivých kateder geografické sekce jsou velmi různorodá. Jako příklad může posloužit Tabulka 1.

Jako další příklad lze uvést data statistická sloužící jako podklad pro analýzy a vizualizaci nejrůznějších jevů v GIS. Dále je třeba mít přístup k textovým i mapovým výstupům výzkumných projektů, diplomových a bakalářských prací studentů.

Při vytváření koncepce SDI je proto potřeba uvážit jaká data se budou dále uchovávat a spravovat, musí být stanovena pravidla, v jakém formátu se do databáze budou předávat data původní i výsledky nejrůznějších zpracování ve formě 2D či 3D kartografických produktů, statistických přehledů a tabulek.

Spravovaná prostorová data budou jednak data základní neboli rámcová a pak data se speciálním tématickým obsahem, který vznikla jako produkt určité analýzy základních prostorových dat. Složení rámcových dat je zásadní pro celou funkci infrastruktury, protože vedle definování standardů pro tato data ovlivňuje možnosti různých prezentací vybraných rysů reálného světa. Model rámcových dat pro SDI by měl klást co nejmenší nároky na pořizování zvolených dat a přitom zahrnout co nejvíc

Tabulka 1. Příklad dat využívaných KAGIK pro výuku.

Data	Typ/Formát	Pokrytí
ZABAGED	vektorová data (dgn, shp)	místní
ArcČR	vektorová data (shp)	národní
Mapy stabilního katastru (skenované)	rastrová data (tif)	národní
LUCC databáze PřF UK v Praze	vektorová data (shp)	národní
Meteosat	rastrová data (tif)	globální
Ortofoto ČR	rastrová data (tif)	národní
Landsat TM, ETM+	rastrová data (tif)	místní
QuickBird	rastrová data (tif)	místní
Barevné letecké snímky (analogová kamera)	rastrová data (tif)	místní
Barevné letecké snímky (digitální kamera)	rastrová data (tif)	místní
Letecké laserové skenování	ASCII (souřadnice X, Y, Z)	místní

zdrojů. Měl by být ale dostatečně flexibilní, aby mohl reagovat na změny způsobené dalším vývojem technických možností i požadavků a cílů různých uživatelů. Při volbě rámce bude klíčové stanovení optimální volby mezi dvěma extrémy. Na jedné straně je přílišná jednoduchost, kdy data nemají žádnou nebo jen malou informační hodnotu, takže nestojí za to, aby se do databáze ukládala. Na druhé straně příliš komplexní stanovení rámce je obtížně pochopitelné, zvláštní přínosy využije jen malý počet partnerů a celkové naplnění je nákladné.

Možná sestava rámcových dat může obsahovat vrstvy digitálního modelu terénu (DEM), ortogonální obrazová data, hydrografickou informaci, administrativní hranice, dopravní infrastruktura (silnice, železnice), geodetické body a sítě.

Kromě rámcových dat jsou uživatelsky důležitá různá další specifická tématická data. Patří mezi ně vrstvy krajinného krytu a nebo využívání krajiny, které se zkoumají v mnoha různých měřítkových pohledech, od studií založených na regionální úrovni až po sledování kontinentální (globální) úrovně. Dalšími příklady jsou půdní mapy

mapy, mapy biodiversity, lesní a geologické mapy, meteorologické mapy, zaznamenání katastrofických událostí a také sociální a ekonomické údaje.

Neodmyslitelnou součástí každé datové sady jsou metadata, Metadata poskytnou inventurní přehled o současných vnitřních zdrojích prostorových dat na fakultě. Naplní funkci katalogu, v němž je možné provádět strukturované hledání a srovnání s daty v jiných databázích. Hlavně ale poskytnou dokumentaci jak data získat, informace o kvalitě a původu dat, které napomohou uživateli data používat ve správné souvislosti.

3.2 Návrh datového skladu a mapového portálu

Datový sklad bude založen na výkonném serveru s několika rychlými disky propojenými v RAID s možnostmi zrcadlení a překonání výpadků jednotek v nepřetržitém provozu. Data budou uložena v objektově-relační databázi a to včetně prostorových vymezení objektů (atributová a prostorová složka dat). Dále spolu s daty bude vytvořen metadatový katalog pro snadné vyhledávání a orientaci v datech. Rastrová data budou v databázi v první fázi projektu uložena jen jako symbolické linky na fyzické soubory (rastry) uložené souborově na disku. Data budou pravidelně aktualizována a to nejenom v rámci zrcadelní disku, ale i na páskovou mechaniku. Podle zvolené databázové platformy bude přístup k prostorovým datům umožněn přímo (např. PostGIS, Oracle Spatial) nebo pomocí aplikační vrstvy mezi desktop aplikací a klientem (např. ArcGIS server řešení nebo ArcSDE).

Pro využívání SDI budou mít důležitou roli služby, které budou podporovat práci s uloženými daty i ukládání nových dat. Zásadní je vyhledávací služba, která umožňuje prostřednictvím metadat zjistit jaké datové zdroje jsou k dispozici. Je to hlavní funkcí datového skladu, s jejíž pomocí uživatel může procházet jednotlivé položky i textové zprávy ve struktuře metadat a zjišťovat informace o vlastnostech dat a možnostech jejich dostupnosti. Druhou skupinou jsou přístupové služby k prostorovým datům. Převážně on-line nebo staticky zpřístupňují původní „syrová“ data, nikoliv mapy. Řešení této služby zahrnuje několik úkolů, jako je definice datového modelu pro data, která se poskytují, navržení přenosového formátu a skladba protokolů. Obecným požadavkem přitom je, aby úkony spojené s přenosem dat byly co nejjednodušší.

Předpokládané funkce a možnosti použití:

- prohlížení dat (celkové procházení dat, výběr dat na základě dotazu)
- distribuce dat výše uvedenými způsoby
- možnost editace dat
- možnost vizualizace uložených dat (pomocí mapového souboru)
- možnost přidání nových dat.

Vedle webového mapového portálu (mapová služba), záznamových datových zařízení (CD-ROM, DVD, BD-R) může být použit i jiný způsob, například velkoplošná

obrazovka, na níž budou zobrazována obrazová data (statická i animace) či mapové výstupy.

Realizace datového skladu bude probíhat v několika fázích:

- zprovoznění na úrovni operačního systému
- přímý přístup k souborově uloženým datům
- webová aplikace pro přímý přístup
- přístup přes mapovou aplikaci „tenkého“ klienta
- přístup přes „těžkého“ klienta (databázová aplikace, GIS aplikace)

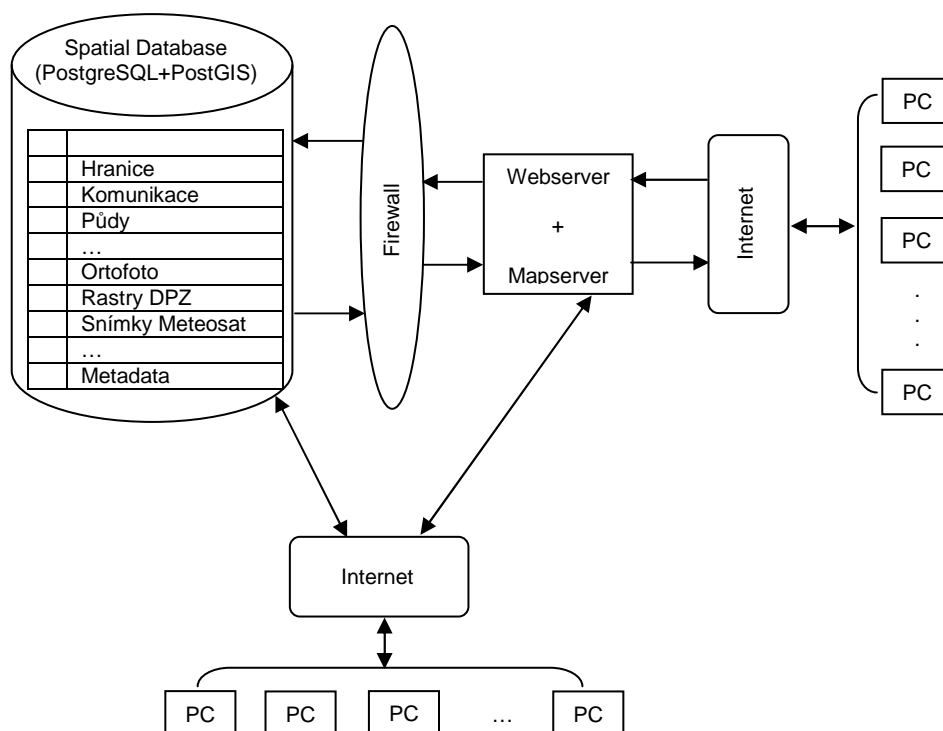
Mapový server bude provádět vizualizaci výsledků jednotlivých projektů v rámci fakulty a případně může být využíván jako brána k volně nebo licencovaným datům spravovaných fakultou (katedrou) pro studijní a vědecké účely. Mapový server bude samostatnou jednotkou vně fakultního intranetu a jediný možný přístup bude přes webové rozhraní. Základní komunikace s daty bude na základě definovaných standardů OGC (WMS, WFS), ve výhledu pak bude spravovat ArcGIS server, resp. bude přístupovou bránou k serveru za fakultním firewallem, který bude obsahovat licencovanou verzi softwaru ArcGis server a kde poběží „geoprocessingové“ služby (ESRI definované a dále dle WFS standardu OGC, resp. dle W3C pravděpodobně na základě WSDL, UDDI a SOAP).

Služby poskytované na úrovni mapového serveru tvoří třetí skupinu služeb. Ty budou tvořit funkce dávající uživateli možnost dodatečného zpracování dat a tím rozšíření a zlepšení informační obsah syrových dat. Mezi takové bude typicky patřit možnost převodů mezi souřadnými systémy, topologické překrytí datových vrstev a další funkce pro analýzu a úpravu rastrových i vektorových dat. Nejvyšší úroveň služeb budou poskytovat modely určitých přírodních i umělých jevů a procesů s možností simulace a předpovědi. Do této skupiny funkcí resp. služeb budou připraveny i vhodné kartografické nástrojů (např. vytváření izolinií, bodově a plošně lokalizované kartodiagramy, sestavení legendy a mapového popisu) a navrženy symboly pro digitální tématické mapové výstupy (mapy krajinného krytu, geologické mapy aj.)

Podle typu poskytovaných služeb a uživatelských kategorií lze mapový server strukturovat do několika dílčích částí:

- „metadatový“ mapserver obsahující metadata o dostupných datech
- „tématický“ mapserver ukazující hlavní mapové a analytické výstupy katedry (fakulty)
- „výzkumný“ mapserver se speciálními službami pro podporu vědeckých a výzkumných prací na fakultě a ukládání jejich výsledků
- „výukový“ mapserver a oddělenou část pro výuku administrace vlastního mapového serveru studenty, nutná restrikce některých administrátorských práv na rovní celého systému)

Obr. 1. Schéma popisovaného řešení správy a přístupu k datům.



U všech úkolů je třeba mít na zřeteli stále to, že jde o části jedné infrastruktury a tedy, že jde současně i o stanovení pravidel, podle kterých bude výsledná SDI fungovat. Tedy standardy, které by měly být optimální pro další rozvoj a usnadnění kompatibility s vnějším světem. Standardy se týkají každé části SDI a každé její funkce nebo akce. Půjde o technické specifikace, formální standardy a nebo pravidla pro vypracování dokumentace. Vodítkem u všech typů služeb a funkcí budou existující dokumenty hlavních institucí vytvářející standardy na různé úrovni jako jsou národní standardy (datový obsah), ISO (základy pro vytváření standardů) a OpenGIS (specifikace interface).

4 Využití navrhované SDI

Navrhovaná SDI bude v první fázi sloužit jako spojovací článek dílčích pracovních skupin v rámci KAGIK. Tyto skupiny se specializují na

- databázové systémy a distribuci prostorových dat
- sběr prostorových dat metodami dálkového průzkumu Země (včetně fotogrammetrie a laserového skenování) a jejich zpracování (obrazové analýzy, prostorové modelování)
- vizualizaci prostorových dat (mapové výstupy, animace)

Zatímco úkolem první skupiny bude návrh datového skladu, přístupů a komunikačních nástrojů pro uživatele, druhé dvě skupiny musí definovat, jakými daty jsou schopny tento sklad naplnit a v jaké formě je mohou poskytnout dalším uživatelům. Úlohou mapového serveru bude nejen dát přehled o poskytovaných datech, ale také co největší míře je uživateli zobrazit při dodržení základních kartografických pravidel, což opět bude vyžadovat spolupráci dílčích skupin.

Navrhovaná SDI bude ve svých jednotlivých fázích vývoje představována katedrám v rámci geografické sekce a později i dalším pracovištím fakulty. Předpokládá se jejich postupné zapojení zejména v otázce specifikace datových struktur a testování funkčnosti navrženého řešení. S využitím standardů a domluvených procedur bude zapotřebí vytvořit partnerské vztahy mezi účastníky zapojenými do infrastruktury.

Uživatelé budou členěni do několika skupin s rozdílnými právy a podle účelu v rámci přístupu k datům. Základní skupiny uživatelů mohou být následující:

- *systemový administrátor* zodpovídající za funkčnost a modifikaci systému
- *databázový administrátor* řídící správu a pravidla databáze prostorových dat
- *editor prostorových dat* s právem data na serveru editovat a aktualizovat
- *vnitřní uživatel* – přístup z vnitřního prostředí fakulty
- *vnější uživatel* - neautorizovaný uživatel mimo fakultu

Dosažením skutečné spolupráce uživatelů prostřednictvím infrastruktury bude dosaženo cíle SDI iniciativy a tím i kvalitativní pokrok v práci s prostorovými daty. Jejím charakteristickým rysem bude stav, kdy jednotliví partneři přestanou upřednostňovat soupeření s ostatními o množství vlastněných dat a nastoupí sdílení dat s ostatními účastníky a spolupráce při pořizování nových dat. Zavedení navrhované SDI tak v průběhu několika příštích let zefektivní a zjednoduší práci s prostorovými daty jak pracovníkům, tak studentům nejen PřF UK v Praze.