

**VYSOKÁ ŠKOLA BÁŇSKÁ –
TECHNICKÁ UNIVERZITA OSTRAVA**

Hornicko-geologická fakulta

Institut geoinformatiky

**Srovnání datových zdrojů o posledních
zemětřeseních**

Bakalářská práce

Autor:

Jan Lojek

Vedoucí bakalářské práce:

Ing. Jan Růžička, Ph.D.

Ostrava 2010

Prohlášení

- *Celou bakalářskou práci včetně příloh, jsem vypracoval samostatně a uvedl jsem všechny použité podklady a literaturu.*
- *Byl jsem seznámen s tím, že na moji bakalářskou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo.*
- *Beru na vědomí, že Vysoká škola báňská – Technická univerzita Ostrava (dále jen VŠB-TUO) má právo nevýdělečně, ke své vnitřní potřebě, diplomovou práci užít (§ 35 odst. 3).*
- *Souhlasím s tím, že jeden výtisk bakalářské práce bude uložen v Ústřední knihovně VŠB-TUO k prezenčnímu nahlédnutí a jeden výtisk bude uložen u vedoucího diplomové práce. Souhlasím s tím, že údaje o bakalářské práci, obsažené v anotaci, budou zveřejněny v informačním systému VŠB-TUO.*
- *Bylo sjednáno, že s VŠB-TUO, v případě zájmu z její strany, uzavřu licenční smlouvu s oprávněním užít dílo v rozsahu § 12 odst. 4 autorského zákona.*
- *Bylo sjednáno, že užít své dílo – bakalářskou práci nebo poskytnout licenci k jejímu využití mohu jen se souhlasem VŠB-TUO, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly VŠB-TUO na vytvoření díla vynaloženy (až do jejich skutečné výše).*

V Ostravě dne 14.2.2010

Jan Lojek

Anotace

Tato bakalářská práce se zabývá srovnáním dvou datových zdrojů o aktuálních zemětřeseních. Jedním ze srovnávaných zdrojů je nezisková organizace ORFEUS a druhým je americká vědecká agentura USGS. Hlavním cílem práce je zhodnocení zdrojů z hlediska rozsahu, aktuálnosti a stability.

Úvodní část práce se zabývá popisem zemětřesení. Jedním z důležitých bodů práce je popis vytvořených aplikací určených k automatickému ukládání dat z RSS kanálu projektu ORFEUS a Atom kanálu agentury USGS. Obě aplikace jsou napsány v jazyce C#. Získaná data jsou podrobena analýze. Kromě analýzy dat práce popisuje i další poskytované služby.

Klíčová slova: zemětřesení, ORFEUS, USGS, RSS, Atom, C#

Summary

This bachelor thesis focuses on the comparison of two data sources about current earthquakes. One of the compared resources is a nonprofit organization ORFEUS and the second is the U.S. scientific agency USGS. The main aim of the thesis is to evaluate the sources in terms of scope, timeliness and stability.

The introductory part of the thesis deals with the description of earthquakes. One of the important points of the thesis is the description of developed computer programs, which are used to store data from the RSS channel of project ORFEUS and from the Atom channel of USGS. Both programs are written in C# programming language. The stored data are then analyzed. In addition to the data analysis the thesis also describes other offered services.

Keywords: earthquake, ORFEUS, USGS, RSS, Atom, C#

Obsah

Úvod	1
1 Zemětřesení	2
1.1 Dělení zemětřesení	3
1.2 Měření zemětřesení	3
1.2.1 Intenzita zemětřesení.....	3
1.2.2 Velikost zemětřesení	4
1.3 ORFEUS.....	5
1.4 USGS.....	6
2 Technologie	7
2.1 XML.....	7
2.2 RSS.....	7
2.3 GeoRSS	8
2.4 Atom.....	8
2.5 KML.....	8
2.5.1 Ukázka KML souboru	9
2.6 C#	10
2.6.1 Ukázka zdrojového kódu aplikace	11
3 Sběr dat	12
3.1 Databáze	12
3.2 Struktura RSS ORFEUS	13
3.3 Aplikace ORFEUS	14
3.3.1 Zdrojový kód	15
3.4 Struktura Atom kanálu USGS.....	18
3.5 Aplikace USGS	19
3.5.1 Zdrojový kód	19
3.6 Omezení aplikací.....	21
3.7 Úprava dat v databázi.....	21

4	Analýza rozdílů v datech.....	24
4.1	Rozdíly v časech.....	25
4.2	Rozdíly v poloze epicenter.....	25
4.3	Rozdíly v hloubkách hypocenter.....	26
4.4	Rozdíly v hodnotách magnituda.....	27
5	Data a služby poskytované agenturou USGS.....	28
5.1	Google Earth KML.....	28
5.2	RSS a CSV.....	30
5.3	Vizualizace dat.....	30
5.4	System ENS.....	31
5.5	System PAGER.....	32
6	Data a služby poskytované projektem ORFEUS.....	33
6.1	Vizualizace dat.....	33
6.2	System Wilber II.....	35
7	Závěrečné zhodnocení.....	36
7.1	Rozsah dat.....	36
7.2	Aktuálnost a stabilita.....	36
7.3	Služby a informace.....	37
	Závěr.....	38
	Literatura a internetové zdroje.....	39
	Seznam obrázků.....	41
	Seznam tabulek.....	41
	Přílohy na CD.....	42

Seznam použitých zkratk

České zkratky

VŠB – TUO Vysoká škola báňská – Technická univerzita Ostrava

Cizojazyčné zkratky

EMSC	European-Mediterranean Seismological Centre
ENS	Earthquake Notification Service
FTP	File Transfer Protocol
IETF	Internet Engineering Task Force
KML	Keyhole Markup Language
NEIC	National Earthquake Information Centre
NERIES	Network of Research Infrastructures for European Seismology
ORFEUS	Observatories and Research Facilities for European Seismology
RSS	Really Simple Syndication
SGML	Standard Generalized Markup Language
USGS	United States Geological Survey
XML	Extensible Markup Language

Úvod

Pravděpodobnost, že se dnes někde na Zemi odehraje zemětřesení, je stoprocentní. Zemětřesení není na této planetě nijak neobvyklou událostí, a přesto jej většina obyvatel Země nikdy nepocítí, protože otřesy půdy bývají tak malé, že je dokážou zaznamenat pouze citlivé vědecké přístroje. Ročně se odehraje několik miliónů zemětřesení. Podle odhadů má destruktivní potenciál přibližně 700 z nich a to pouze za předpokladu, že se odehrají v obydlených oblastech, což se naštěstí stává velmi zřídka. Pokud však zemětřesení zasáhne obydlenou oblast, rázem se z něj stává jedna z nejničivějších přírodních katastrof.

K výzkumu zemětřesení jsou zapotřebí data, která jsou získávána seismologickými stanicemi po celém světě. Data ze seismologických stanic bývají publikována pomocí různých technologií v rámci internetu.

Tato práce se zabývá srovnáním dvou volně dostupných datových zdrojů o aktuálních zemětřeseních. Cílem práce není určit, který zdroj poskytuje přesnější data, protože to není možné spolehlivě určit. Důraz je kladen na jiné důležité faktory, jako jsou rozsah a aktuálnost dostupných dat. Kromě samotných dat se práce zabývá také popisem a srovnáním doprovodných služeb, které jsou v rámci obou zdrojů uživatelům k dispozici.

Důležitým bodem práce je automatizace sběru dat prostřednictvím vytvořených aplikací a následné ukládání dat do připravené databáze, za účelem analýzy rozdílů v poskytovaných datech.

Výsledky práce mohou být využity několika způsoby. Jednak mohou posloužit při rozhodování, který ze srovnávaných datových zdrojů bude vhodnější pro další využití, ale rovněž mohou být využity jako návod, jak porovnat dva nebo i více zdrojů dat, prezentovaných v podobné formě.

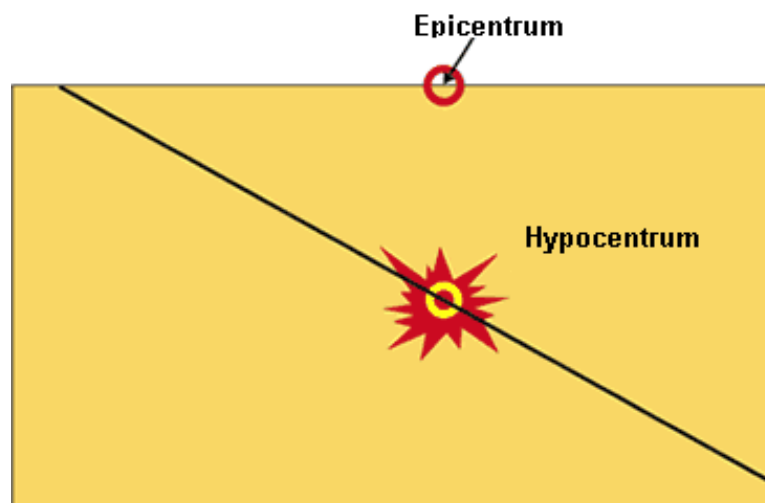
1 Zemětřesení

Jako zemětřesení označujeme náhlý otřes zemské kůry, jehož příčinou je uvolnění velkého množství energie v nitru Země. Zemětřesení se často objevují ve skupinách označovaných jako zemětřesné posloupnosti, které se skládají z několika slabších předtřesů, následovaných hlavním otřesem a slabšími dotřesy. Doba dotřesů se může pohybovat v rozmezí několika měsíců až let. [13]

Zemětřesení se nejčastěji vyskytují v oblastech, kterými prochází zlomy litosférických desek. Jmenovitě se jedná například o oblasti Japonska, západního pobřeží Severní i Jižní Ameriky, jihovýchodní Asie, Íránu nebo Turecka.

Zemětřesení patří k vůbec nejhorším přírodním katastrofám, ať už z hlediska počtu obětí nebo rozsahu postiženého území. Za nejkatastrofálnější zemětřesení se považuje zemětřesení v čínské provincii Shaanxi, které se odehrálo v roce 1556. Na následky otřesů půdy tehdy zahynulo více než 800 tisíc lidí. Což bylo zhruba 60% obyvatel celé provincie. [19]

Při určování polohy zemětřesení se používají dva základní pojmy – hypocentrum a epicentrum. Hypocentrum představuje těžiště ohniska vzniku otřesů pod zemským povrchem. Epicentrum je pak kolmý průmět hypocentra na zemský povrch. [13]



Obr. 1: Znázornění hypocentra a epicentra

Během zemětřesení se z hypocentra šíří seismické vlny, které se dělí na vlny podélné a příčné. Na základě zpoždění mezi oběma druhy vln se určuje poloha hypocentra. [2]

1.1 Dělení zemětřesení

Rozlišujeme tři typy zemětřesení podle jejich původu. Nejčastější (přibližně 95%) a také nejnebezpečnější jsou zemětřesení tektonická, která vznikají vlivem pohybu tektonických desek podél jejich zlomů. Druhým nejčastějším typem jsou zemětřesení sopečná, která většinou předcházejí erupcím sopek a výlevům lávy a jsou způsobena pohybem desek vlivem tlaku vylévající se lávy nebo unikajících plynů a par. Posledním typem jsou zemětřesení řítivá. Ty mají sice lokální charakter, ale následky mohou mít rovněž katastrofální. Vznikají zřícením podzemních dutin krasového nebo důlního původu. [9]

Dále je možné zemětřesení dělit podle hloubky ohniska. Mělká zemětřesení vznikají v zemské kůře a svrchní části zemského pláště v hloubce maximálně 60 km. Středně hluboká zemětřesení vznikají v zemském plášti v hloubkách od 60 do 300 km a hluboká zemětřesení vznikají v hloubkách od 300 km. [9]

Podle oblasti vzniku se zemětřesení mohou rozdělit také na kontinentální a na podmořská. Kontinentální otřesy způsobují škody zpravidla okamžitě, vlivem podmořských otřesů vznikají vlny tsunami. Tyto vlny dosahují u pobřeží několikametrových výšek a mají devastující účinek. Jako příklad z poslední doby může posloužit podmořské zemětřesení v Indickém oceánu z prosince roku 2004. Vlna tsunami si vyžádala přes 200 tisíc obětí v Asii a v Africe. [9]

1.2 Měření zemětřesení

1.2.1 Intenzita zemětřesení

Sílu zemětřesení můžeme popsat ze dvou úhlů pohledu. Na subjektivním pozorování je založena veličina označována jako intenzita zemětřesení. Jedná se o popis projevů otřesů v krajině, což můžeme chápat například jako popis úrovně poškození budov. Intenzita není ve všech zasažených oblastech stejná. Ve většině případů klesá směrem od epicentra. K vyjádření intenzity se nejčastěji využívá modifikované Mercalliho stupnice, kterou vytvořil italský seismolog Giuseppe Mercalli. V Evropské unii se v současnosti používá stupnice modifikovaná stupnice s označením EMS-98.

Intenzita	Definice	Zkrácený popis typických účinků
1	Nepocítěno	Nepocítěno.
2	Zřídka pocítěno	Pocítěné jen jednotlivci na některých místech v domech.
3	Slabé	Zemětřesení uvnitř budov cítí jen někteří lidé (0-20%). Cítí jej nanejvýš jako houpání nebo lehké chvění.
4	Značně pozorované	Zemětřesení uvnitř budov cítí mnozí (10-60%), venku jen výjimečně. Někteří jsou probuzeni. Okna a dveře rachotí.
5	Silné	Zemětřesení uvnitř budov cítí většina (50-100%), venku někteří. Mnozí spící se probudí. Někteří jsou vystrašení. Budovy vibrují. Visící objekty se značně houpají. Malé předměty se posouvají. Dveře a okna se otvírají a zavírají.
6	Mírně ničivé	Mnozí jsou vystrašení a vybíhají ven. Některé předměty padají. Mnohé budovy utrpí malé nestrukturální škody jako např. vlásečnicové trhliny nebo odpadnuté malé kousky omítky.
7	Ničivé	Většina lidí je vystrašená a vybíhá ven. Nábytek je posunutý. Předměty padají z polic ve velkém množství. Mnohé dobře postavené běžné budovy utrpí střední škody: opadáva omítka, padají části komínů; ve stěnách starších budov jsou velké trhliny a příčky jsou zřícené.
8	Těžce ničivé	Mnozí mají problémy udržet rovnováhu. Mnohé domy mají velké trhliny ve stěnách. Několik dobře postavených běžných budov má vážně poškozené stěny. Slabé starší budovy se mohou zřítit.
9	Destruktivní	Všeobecná panika. Mnoho chatrných budov se řítí. I dobře postavené běžné budovy utrpí velmi těžké škody: těžké poškození stěn a částečně i strukturální škody.
10	Velmi Destruktivní	Mnohé dobře postavené běžné budovy se řítí.
11	Devastující	Většina dobře postavených běžných budov se řítí. I některé dobře anti-seismicky postavené budovy jsou zničené.
12	Úplně devastující	Téměř všechny budovy jsou zničené.

Tab. 1: Stupnice EMS-98

1.2.2 Velikost zemětřesení

Druhou veličinou je velikost zemětřesení. Jedná se o objektivní veličinu, jejíž hodnota se stanovuje na základě měření. Velikost zemětřesení se měří pomocí seismografů, což jsou

přístroje, které zaznamenávají pohyb půdy. Za vynálezce seismografu je považován čínský astronom a matematik Chang Heng. Seismografy mají pro seismologii zcela zásadní význam. Bez nich by se zemětřesení zkoumaly velmi složitě. Seismografy se skládají ze seismometru a analogového nebo digitálního záznamového zařízení. Analogové zařízení vykresluje údaje v kontinuální podobě na role papíru. Digitální zařízení pak ukládá data v podobě číselných hodnot. Výstupem seismografu je v obou případech seismogram. [1]

K vyjádření velikosti zemětřesení se používá řada stupnic, z nichž je zcela jistě nejznámější Richterova stupnice, kterou vytvořil v roce 1935 americký seismolog Charles F. Richter ve spolupráci s Beno Gutenbergem. Magnitudo zemětřesení se počítá pomocí logaritmu vlny zaznamenané seismografem. Problémem této stupnice je fakt, že byla sestavena pro srovnávání středně velkých zemětřesení v Kalifornii měřených v poměrně malé vzdálenosti od epicentra. S rostoucí vzdáleností od epicentra a s rostoucí velikostí zemětřesení se přesnost stupnice zmenšuje, nehledě na to, že podmínky ve světě jsou jiné než podmínky v Kalifornii, na jejichž základě byla stupnice sestavena. Stupnice je vhodná pro zemětřesení s magnitudem v rozmezí 3-7. [15][18]

V současné době vědci nejčastěji používají stupnici momentového magnituda, která přesněji vystihuje velká zemětřesení. Stupnici sestavili v roce 1979 seismologové Thomas C. Hanks a Hiroo Kanamori. Číselné hodnoty u menších a středních zemětřesení jsou z pravidla téměř totožné jako u Richterovy škály. Rozdíly se nejvíce projevují u velkých zemětřesení s magnitudem větším než 7. Kromě momentového magnituda se používají ještě další stupnice, které jsou také sestaveny tak, aby odpovídaly Richterově stupnici. [7]

1.3 ORFEUS

ORFEUS je zkratka pro Observatoře a výzkumná zařízení pro evropskou seismologii. Jedná se o neziskovou organizaci, založenou v roce 1987. Organizace si klade za cíl koordinaci a propagaci digitální širokopásmové seismologie v oblasti Středozeemí.

Projekt je financován a řízen zakladateli ze 13 evropských států. Činnost organizace je rozdělena mezi Datové centrum (ODC), které zajišťuje sběr, archivaci a poskytování dat o průběhu vlny, a mezi čtyři pracovní skupiny, které se starají o dostupnost dat a vývoj. Každodenní činnost zaměstnanců je řízena výkonným výborem, jmenovaným správní radou.

ODC slouží jako regionální datové centrum v Mezinárodní federaci digitálních seismografických sítí a hostí jej Královský meteorologický institut v Holandsku.

ORFEUS působí pod záštitou Evropské seismologické komise (ESC) a úzce spolupracuje se svou sesterskou organizací v Evropě, Evropsko-středozezemním seismologickým centrem (EMSC).

ORFEUS v současné době koordinuje archivaci a přístup k datům o průběhu vlny zemětřesení ze seismických stanic ve středomořské oblasti. ORFEUS rovněž koordinuje projekt NERIES.

Na adrese <http://www.orfeus-eu.org/orfeus-rss.xml> poskytuje projekt ORFEUS data o posledních zemětřeseních v podobě RSS kanálu. [10]

1.4 USGS

United States Geological Survey, česky Geologická služba Spojených států, je nezávislá vědecká agentura spadající pod Ministerstvo vnitra Spojených států. USGS byla založena 3. března 1879 několik hodin před koncem 45. Kongresu Spojených států. V současnosti se zabývá čtyřmi hlavními disciplínami – biologií, geologií, geografii a hydrologií. Agentura zaměstnává přibližně 10 000 lidí a její hlavní sídlo se nachází ve městě Reston ve Virginii. Další větší sídla se nacházejí v Denveru v Coloradu a v Menlo Park v Kalifornii. Rozpočet pro rok 2010 byl stanoven na 1,1 miliardy dolarů.

Jedním z programů, kterými se USGS zabývá, je Earthquakes Hazards Program (Rizika zemětřesení). Cílem programu je poskytování vědeckých informací a znalostí o zemětřeseních, které by vedly ke snížení počtu úmrtí, zranění a škod na majetku prostřednictvím pochopení vlastností a účinků zemětřesení. Program je oblastně zaměřen především na Spojené státy, ale monitoruje i zemětřesení po celém světě. [16]

2 Technologie

2.1 XML

Extensible markup language (XML) je jednoduchý a velmi flexibilní značkovací jazyk odvozený od staršího jazyka SGML. Určen je především pro strukturování dat. Pod pojmem strukturovaná data si můžeme představit například adresář, databázi, záznam finančních transakcí atd. V XML se podobně jako v jazyce HTML používají tagy. Rozdíl spočívá v tom, že v HTML jsou tagy pevně definovány a v XML si musí jednotlivé tagy definovat uživatel podle své potřeby. [20]

Díky jednoduchosti a rozšiřitelnosti XML vzniklo velké množství implementací. Patří mezi ně i RSS, KML a Atom. XML dokumenty mají sice textovou podobu, ale k přímému čtení uživatelem nejsou určeny, k tomu slouží aplikace. Nicméně právě díky textovému formátu dávají data v případě potřeby smysl i při přímém čtení, což je zásadní rozdíl oproti datům uloženým v binární podobě. [20]

2.2 RSS

RSS je formát určený primárně k syndikaci obsahu mezi webovými stránkami. Pro informační portály se staly samozřejmostí RSS kanály, které návštěvníkům umožňují odebrat novinky pomocí RSS čtečky, aniž by museli stránky sami navštívit, čímž se udržuje trvalý kontakt mezi webovou stránkou a návštěvníkem. Moderní internetové prohlížeče jako jsou Firefox, Internet Explorer nebo Safari mají již takovouto čtečku přímo v sobě zabudovanou. U prohlížeče Google Chrome tomu tak není, nicméně čtečku lze přidat pomocí pluginu. Kromě zabudovaných čteček existují také samostatné aplikace jako třeba SharpReader. Čtečky v určitém časovém intervalu stahují data z RSS kanálu a upozorňují uživatele na změny. RSS dokumenty mohou obsahovat například nadpis článku, datum a čas, kdy byl článek zveřejněn, a odkaz na celý článek. [8]

První verzi RSS, označovanou jako RSS 0.9, vytvořil v roce 1999 indický programátor Ramanathan V. Guha pro společnost Netscape. V srpnu téhož roku Dan Libby formát RSS zjednodušil vytvořením nové verze označované jako RSS 0.91. Historie RSS je složitá, protože se o formát přetahovalo několik organizací. Důsledkem je fakt, že jednotlivé verze

RSS nejsou mezi sebou kompatibilní. Přehled verzí a jejich tvůrců je uveden v tabulce. V současné době se využívá převážně verze RSS 2.0 z roku 2002. [6]

Verze	Tvůrce	Komentář
RSS 0.90	Netscape	Zastaralý po vzniku 1.0
RSS 0.91	Netscape, přebráno f. UserLand	Velmi jednoduchý, oficiálně zastaralý po vzniku verze 2.0
RSS 0.92, 0.93 a 0.94	UserLand	Obsahují bohatší metadata než 0.91. Zastaralé po vzniku verze 2.0.
RSS 1.0	RSS-DEV Working Group	Založený na RDF, rozšiřitelný pomocí modulů.
RSS 2.0	UserLand	Současná verze

Tab. 2: Verze RSS

2.3 GeoRSS

GeoRSS umožňuje rozšířit stávající klasické RSS kanály o geografické informace. Mezi nejvyužívanější možnosti patří bodová lokalizace publikované informace pomocí souřadnic. U GeoRSS rozlišujeme dva druhy kódování – GeoRSS Simple a GeoRSS GML. GeoRSS Simple je velmi odlehčená verze, pomocí níž lze definovat bod, linii, obdélník, polygon a kruh. Pomocí elementu elevation lze definovat i výšku nad elipsoidem WGS-84. GeoRSS GML je pokročilejší verze, nabízející více možností, z nichž vyčnívá především možnost používat odlišné souřadnicové systémy. [4]

2.4 Atom

Atom Syndication Format je stejně jako RSS formát určený k syndikaci obsahu a odebírání novinek z webových stránek. Hlavním důvodem vzniku vývoje Atomu byla snaha vytvořit jednotný formát, jelikož jednotlivé verze RSS nebyly mezi sebou kompatibilní. V roce 2005 je komisi IETF předložen návrh standardu. Označen je jako RFC 4287 a dosud se nachází ve fázi standardizace. Hlavní rozdíly mezi formáty Atom a RSS přiblíží podkapitola, věnující se popisu struktury Atom kanálu agentury USGS. [8]

2.5 KML

Keyhole Markup Language je jazyk určený pro vizualizaci geografických dat v prostředí aplikací Google Earth, Google Maps a Google Maps for mobile (Google Maps pro

mobilní telefony). KML vyvinula firma Keyhole pro svou aplikaci EarthViewer. V roce 2004 firmu odkoupila společnost Google a aplikace EarthViewer byla přejmenována na Google Earth. Jazyk KML je podobně jako RSS založen na standardu XML, z čehož vyplývá relativní jednoduchost tohoto formátu. V současnosti se používá KML ve verzi 2.2. [5]

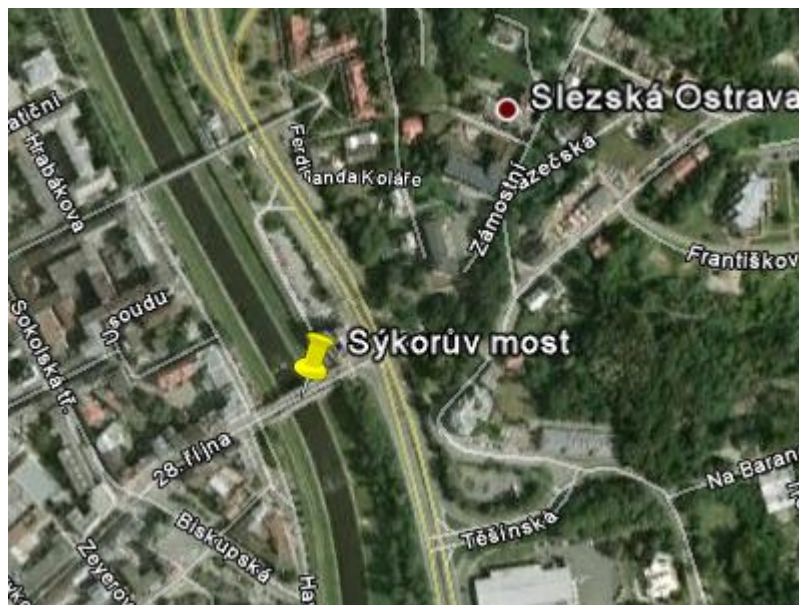
Pomocí jazyka KML lze vizualizovat řadu geografických prvků, jako jsou body, linie nebo polygony, rovněž lze do aplikace Google Earth importovat trojrozměrné modely budov. KML však nezůstává pouze u klasické vizualizace prvků, ale podporuje například i změnu směru pohledu kamery nebo umožňuje definovat pohyb kamery podél určené trasy, což mohou uživatelé využít například pro znázornění letu letadla. K jednotlivým bodům lze přidávat fotografie, komentáře, hypertextové odkazy atd.

KML soubory je možné distribuovat v nekomprimované podobě s koncovkou KML a nebo v podobě komprimované s koncovkou KMZ. KMZ archívy je nutné komprimovat pomocí formátu ZIP. Google doporučuje komprimaci v případě, že KML soubor přesáhne velikost 10 kB, případně když je soubor doplněn o další data v podobě obrázků atd. Dvě nejzajímavější místa, odkud mohou uživatelé již vytvořené KML soubory stahovat, jsou Galerie Google Earth (<http://earth.google.com/gallery>) a komunitní webové stránky <http://bbs.keyhole.com>. Právě v Galerii Google Earth se pod názvem Real-time Earthquakes nachází i KML zobrazující poslední zemětřesení. [11]

2.5.1 Ukázka KML souboru

V následující jednoduché ukázce kódu je zobrazena struktura KML souboru. Na začátku souboru se vždy objevuje XML hlavička s informací o verzi XML. Součástí hlavičky je zde i informace o použitém kódování. Následuje element `kml` s deklarací jmenného prostoru KML. Pokud uživatel tento soubor otevře v prostředí Google Earth, označí se mu poloha Sýkorového mostu v Ostravě pomocí ikony žlutého špendlíku. Vzhled této ikony není pevně dán a je možné jej změnit. Přesná poloha je definována souřadnicemi v elementu `coordinates`. Element `name` slouží k definování názvu označeného bodu a element `description` k jeho popisu.


```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://earth.google.com/kml/2.2">
  <Placemark>
    <name>Sýkorův most</name>
    <description>Sýkorův most v Ostravě</description>
    <Point>
      <coordinates>18.29636886293216,49.83750623947687,0</coordinates>
    </Point>
  </Placemark>
</kml>
```



Obr. 2: Sýkorův most v aplikaci Google Earth

2.6 C#

C# (C Sharp) je objektově orientovaný programovací jazyk vyvinutý společností Microsoft spolu s platformou .NET. Jeho vývoj byl započat v roce 1999 pod názvem Cool (C-like object oriented language), ale ve finální fázi vývoje bylo rozhodnuto o přejmenování jazyka na C#. V současnosti se používá ve verzi 3.0 z roku 2007. Na rok 2010 je plánována verze 4.0. Jazyk C# je jednoduchý, ale zároveň velmi výkonný. Po svých předchůdcích C++ a Microsoft Visual Basic zdědil celou řadu dobrých vlastností a těch špatných naopak velmi málo, což má za následek čistší a logičtější jazyk. Jazyk je také velmi podobný programovacímu jazyku Java. Programátor James Gosling, tvůrce jazyku Java, jej dokonce

označil za pouhou imitaci jazyku Java. V jazyce C# je možné vytvářet téměř libovolné aplikace. [2][12]

2.6.1 Ukázka zdrojového kódu aplikace

Jako názorná ukázka poslouží známá aplikace Hello World. Po spuštění aplikace, jejíž zdrojový kód je uveden níže, se otevře konzole a na obrazovku se vypíše text „Hello, World!“ Okno konzole se zavře po zmáčknutí libovolné klávesy.

```
using System;
class HelloWorld
{
    public static void Main()
    {
        Console.WriteLine("Hello, World!");
        Console.ReadKey();
    }
}
```

3 Sběr dat

Jedním z hlavních bodů této práce je srovnání dat z obou zdrojů. Aby bylo možné toto srovnání uskutečnit, je nutné data nějakým způsobem získávat. Samozřejmě se nabízí možnost ručního ukládání dat, ale to by bylo časově velmi náročné a neefektivní. Pro tento účel byly tedy vytvořeny aplikace, které tuto úlohu téměř zautomatizují. Podle původního předpokladu měly být vytvořeny dvě aplikace, z nichž jedna měla ukládat data z RSS kanálu projektu ORFEUS a druhá měla ukládat data z KML Real-time Earthquakes. Vzhledem k tomu, že tvůrcem Real-time Earthquakes je americká agentura USGS, která na svých webových stránkách nabízí stejná data v několika různých formátech, je v této práci namísto KML využit kanál Atom. Toto rozhodnutí však nic nezměnilo na faktu, že byly vytvořeny dvě aplikace, protože struktura kanálů RSS a Atom je natolik odlišná, že nelze pro oba zdroje použít totožnou aplikaci.

3.1 Databáze

Data získaná pomocí dvou zmíněných aplikací jsou ukládána do databáze, která byla vytvořena v prostředí aplikace na správu relačních databází Microsoft Office Access. Data z obou zdrojů jsou ukládána do jedné tabulky. V případě potřeby je možné data rozdělit do dvou tabulek pomocí jednoduchého SQL dotazu.

Tabulka obsahuje celkem 8 sloupců. Do prvního sloupce označeného jako location, se ukládají názvy oblastí, kde se zemětřesení odehrála. Datum a čas se do databáze ukládají odděleně. Následují 4 sloupce, které obsahují základní popis vlastností zemětřesení – zeměpisná délka a šířka, hloubka hypocentra a samozřejmě magnitudo. Do posledního sloupce je ukládán zdroj daného zemětřesení.

Přehled datových typů:

- **location** – Text; Velikost pole – 255 znaků
- **datum** – Datum a čas; Formát – datum (krátké)
- **cas** - Datum a čas; Formát – hh:nn:ss (hodiny:minuty:sekundy)
- **latitude** – Číslo; Velikost pole – desetinné číslo
- **longitude** – Číslo; Velikost pole – desetinné číslo
- **depth** – Číslo; Velikost pole – desetinné číslo

- **mag** – Číslo; Velikost pole – desetinné číslo
- **source** – Text; Velikost pole – 255 znaků.

3.2 Struktura RSS ORFEUS

```
<?xml version="1.0"?>
<rss version="2.0">
<channel>

<title>Recent earthquakes - ORFEUS</title>
<link>http://www.orfeus-eu.org</link>
<description>Recent earthquakes determined by the VEBSN at ORFEUS</description>
<image>
<url>http://www.orfeus-eu.org/gif/ORFEUS_logo.gif</url>
</image>

<item>
<title>2010-03-15, Off coast of central Chile, M = 6.0</title>
<link>http://www.orfeus-eu.org/cgi-
bin/wilberII/wilberII_page3.pl?evid=27112</link>
<description>2010-03-15; 11:08:28; lat=-35.9; lon=-73.3; depth=10; mag=6.0; Off
coast of central Chile; author=NEIC-BQ; product=7112</description>
</item>
<item>
<title>2010-03-14, South Indian Ocean, M = 6.0</title>
<link>http://www.orfeus-eu.org/cgi-
bin/wilberII/wilberII_page3.pl?evid=27111</link>
<description>2010-03-14; 20:33:10; lat=-2.8; lon=83.7; depth=10; mag=6.0; South
Indian Ocean; author=NEIC-BQ; product=7111</description>
</item>

</channel>
</rss>
```

Výše zobrazena ukázka pochází ze dne 15. 3. 2010. Z důvodu úspory místa je ukázka omezena na pouhá dvě zaznamenaná zemětřesení. Na začátku souboru je v hlavičce definováno, že se jedná o XML soubor a o RSS verze 2.0.

Element channel obsahuje základní informace o kanálu RSS. Podle specifikace W3C musí tento element obsahovat tři další elementy – title, link a description. Element title definuje název celého kanálu. V tomto případě se kanál jmenuje Recent Earthquakes – ORFEUS. Element link obsahuje odkaz na domovskou stránku kanálu. A poslední element description slouží ke slovnímu popisu obsahu kanálu. Channel může dále obsahovat řadu dalších volitelných elementů. V případě kanálu ORFEUS je využit pouze element image, který slouží k zobrazení loga organizace.

Informace o zemětřeseních jsou obsaženy v elementu item, přičemž každý element item představuje právě jedno zemětřesení. Obdobně jako u popisu kanálu i v elementu item nalezneme elementy title, link a description. Title v tomto případě obsahuje datum, kdy se zemětřesení odehrálo, jméno oblasti a také magnitudo. Element link obsahuje odkaz na podrobnější informace o daném zemětřesení v prostředí systému Wilber II, kterému se podrobněji věnuje jedna z následujících kapitol. Veškeré důležité informace nalezneme v elementu description – jsou to datum a čas, zeměpisné souřadnice, hloubka, magnitudo a název oblasti. Pomocí aplikace je tedy nutné celý element rozložit na jednotlivé položky a ty následně uložit do databáze.

Na předposledním místě v elementu description se nachází položka author, uvádějící původní zdroj těchto dat. V ukázce je u obou zemětřesení jako zdroj uvedeno NEIC spadající pod USGS. Na první pohled se jedná o velmi cennou informaci, která by velmi usnadnila práci při zjišťování příčin rozdílů v datech. Bohužel však Atom kanál USGS podobnou informaci neposkytuje.

3.3 Aplikace ORFEUS

V této části se konečně podíváme na již několikrát zmíněnou aplikaci. Podkapitola je zaměřena především na základní principy práce aplikace a na popis potencionálních problémů. Po přečtení této podkapitoly by měl být čtenář schopen vytvořit podobnou aplikaci. Předpokladem je alespoň základní znalost programování v jazyce C#.

Aplikace funguje na jednoduchém principu. Po každém spuštění se aplikace připojí k nadefinovanému XML souboru a načte jeho obsah. Základním prvkem fungování aplikací je cyklus, který postupně prochází celý XML soubor, jednotlivé záznamy o zemětřeseních upravuje do potřebné podoby a následně je ukládá do připravené databáze.

3.3.1 Zdrojový kód

V prvním kroku je nutné definovat nový XML dokument. To provedeme pomocí třídy XmlDocument. Instance je zde pojmenována jako xDoc. Následně pomocí metody Load() načteme XML soubor z RSS kanálu ORFEUS.

```
XmlDocument xDoc = new XmlDocument();  
xDoc.Load("http://www.orfeus-eu.org/orfeus-rss.xml");
```

Jakmile máme soubor načtený, musíme nějakým způsobem načíst i jeho strukturu a obsah. K tomuto účelu využijeme třídu XmlNodeList, která pomocí metody GetElementsByTagName() načte jednotlivé elementy definované v závorkách do seznamů elementů. Je nutné podotknout, že tímto se načtou všechny elementy, které mají stejný název, jaký je uvedený v závorkách. To znamená, že se načtou i elementy z počátku dokumentu, které slouží pouze k popisu kanálu a jsou tudíž bezvýznamné.

```
XmlNodeList item = xDoc.GetElementsByTagName("item");  
XmlNodeList title = xDoc.GetElementsByTagName("title");  
XmlNodeList desc = xDoc.GetElementsByTagName("description");
```

Pomocí textového řetězce database nadefinujeme zprostředkovatele databáze (Microsoft.Jet.OleDB.4.0) a také cestu k databázi. V případě, že je cesta k databázi definována jako v ukázce níže, musí se databáze nacházet ve stejné složce jako aplikace. Třída OleDbConnection představuje připojení k databázi.

```
string database = "provider=Microsoft.Jet.OleDB.4.0; " +  
"data source=earthquakes.mdb";  
OleDbConnection connection = new OleDbConnection(database);  
connection.Open();
```

V dalším kroku je třeba nadefinovat číselnou proměnnou i. Tato proměnná bude sloužit při procházení jednotlivých záznamů. Nastavením její počáteční hodnoty na i = 1 předejdeme potencionálnímu problému, zmíněnému o několik řádků výše, protože první záznam, který uložit nepotřebujeme, je v XML dokumentu vedený jako nultý. Textový řetězec „orfeus“ slouží k rozlišení zdrojů dat po uložení do tabulky.

Následuje základní kámen celé aplikace. Tím je cyklus foreach, který projde všechny elementy načtené do seznamů v jednom z předcházejících kroků. Ve složených závorkách budou uvedeny všechny operace, potřebné k úpravě načtených údajů a jejich následnému

uložení do databáze. Po provedení všech operací dojde k navýšení hodnoty proměnné `i` o 1 a cyklus projde další záznam v XML souboru. Celý proces se bude opakovat, dokud nebudou všechny záznamy upraveny a uloženy.

```
int i = 1;
string source = "orfeus";
foreach (XmlNode node in item)
{
...
i++;
}
```

Nyní je třeba si přiblížit výše zmíněné operace, probíhající v cyklu `foreach`. Všechny údaje, které potřebujeme získat, se v XML souboru RSS ORFEUS nachází v elementu `description`. Na první pohled se může jako problém jevit fakt, že údaje jsou naskládány za sebe. Ale není tomu tak. Každý údaj je oddělen středníkem, čehož lze využít a pomocí metody `Split()` se středníkem jako parametrem celý záznam rozdělit na jednotlivé části. `InnerText` je vlastností třídy `XmlNode`, která vrací text nacházející se mezi počátečním a koncovým tagem elementu. Výsledkem je pole textových řetězců, na které se můžeme odkazovat čísly. Například číslo 0 odpovídá datum, protože se nachází na první pozici.

```
string[] descSplit = desc[i].InnerText.Split(';');
```

Tím s metodou `Split()` ještě nekončíme, protože u údajů popisujících zeměpisnou délku a šířku, magnitudo a hloubku, se musíme zbavit názvů a rovníčka. Parametrem metody `Split()` bude v tomto případě samotné rovníčko. Zeměpisná šířka se v elementu `description` nachází na třetí pozici, proto je v hranatých závorkách číslo 2. Po této operaci zůstanou u zeměpisných souřadnic pouze číselné hodnoty.

```
string[] lat = descSplit[2].Split('=');
string[] lon = descSplit[3].Split('=');
```

Mohlo by se zdát, že získané hodnoty již lze ukládat do databáze. Bohužel to však ještě není možné, protože dané hodnoty mají podobu textových řetězců (`string`) a do databáze se musí ukládat v podobě desetinných čísel. Problém se snadno vyřeší parsováním textu na desetinná čísla (`float`) metodou `Parse()`. Nyní lze tyto hodnoty bez problémů uložit do databáze. Stejná úprava se týká i magnituda a hloubky.

```
float latitude = float.Parse(lat[1]);  
float longitude = float.Parse(lon[1]);
```

Při zpracování se vyskytl problém, týkající se ukládání času a oblasti. Pokud totiž použijeme metodu Split() na rozdělení řetězců na základě definovaného oddělovače, budou jednotlivé řetězce obsahovat i mezery, které se vyskytují za oddělovačem. V případě údajů, které se dále dělí, to samozřejmě problém není, protože se ukládá až druhá část a ta mezeru již neobsahuje. Čas a oblast se ale dále nedělí a proto se musí mezera odstranit, jinak by nebylo možné data uložit. K odstranění poslouží metoda Trim() bez dalších parametrů.

```
string time = descSplit[1].Trim();  
string location = descSplit[6].Trim();
```

Poslední důležitou operací v cyklu je samotné uložení upravených dat. U tohoto kroku je třeba nadefinovat SQL příkaz (String SQL), který uložení dat provede. Z ukázky je patrné, že se data budou ukládat do tabulky Zemetreseni. V závorce se definují sloupce tabulky. Následují již konkrétní ukládané hodnoty (VALUES). Třída OleDbCommand definuje příkaz potřebný k vykonání SQL příkazu. Třída má dva parametry. Prvním parametrem je SQL příkaz a druhým připojení k databázi (connection). Připojení jsme definovali ještě před samotným cyklem. Veškeré potřebné operace s databází by bylo možné zapsat přímo do cyklu foreach, ale to by znamenalo, že by se s každým záznamem opětovně vytvářelo připojení k databázi, což by ukládání zpomalovalo. V této podobě se tedy připojení vytvoří pouze jednou, všechny záznamy se uloží a připojení se ukončí.

Po spuštění aplikace by mohly všechny zmíněné operace proběhnout, aniž by o tom uživatel věděl. A jelikož jsou uživatelé často netrpěliví a nervózní, když nevědí, co program zrovna provádí, je cyklus doplněn o výpis aktuálně ukládaného zemětřesení.

```
String SQL = "INSERT INTO Zemetreseni (location, datum, cas, latitude, longitude,  
depth, mag, source) " + "VALUES('" + location + "', '" + czdate + "', '" + time +  
"', '" + latitude + "', '" + longitude + "', '" + depth + "', '" + magnitude + "',  
'" + source + "')";  
OleDbCommand insertValues = new OleDbCommand(SQL, connection);  
insertValues.ExecuteNonQuery();  
  
Console.WriteLine("Zemětřesení v oblasti " + location + " přidáno do databáze");
```


Po vykonání cyklu se metodou Close() ukončí připojení k databázi. Okno konzole s výpisem zůstane otevřeno, dokud uživatel nestiskne libovolné tlačítko.

```
connection.Close();
Console.WriteLine("Všechny záznamy byly uloženy do databáze.");
Console.WriteLine("Aplikaci ukončíte stisknutím libovolné klávesy.");
Console.ReadKey();
```

3.4 Struktura Atom kanálu USGS

```
<?xml version="1.0"?>
<feed xmlns=http://www.w3.org/2005/Atom
xmlns:georss="http://www.georss.org/georss">
  <updated>2010-03-16T17:16:08Z</updated>
  <title>USGS M5+ Earthquakes</title>
  <subtitle>Real-time, worldwide earthquake list for the past 7 days</subtitle>
  <link rel="self" href="http://earthquake.usgs.gov/earthquakes/catalogs/7day-
M5.xml"/>
  <link href="http://earthquake.usgs.gov/earthquakes/" />
  <author><name>U.S. Geological Survey</name></author>
  <id>http://earthquake.usgs.gov/</id>
  <icon>/favicon.ico</icon>
  <entry><id>urn:earthquake-usgs-gov:us:2010twam</id><title>M 6.0, offshore Bio-
Bio, Chile</title><updated>2010-03-15T11:08:28Z</updated><link rel="alternate"
type="text/html"
href="http://earthquake.usgs.gov/earthquakes/recenteqsww/Quakes/us2010twam.php"/><
summary type="html"><![CDATA[<p>Monday, March 15, 2010 11:08:28 UTC<br>Monday, March 15, 2010 08:08:28 AM at
epicenter</p><p><strong>Depth</strong>: 10.00 km (6.21
mi)</p>]]></summary><georss:point>-35.8805 -73.2827</georss:point><georss:elev>-
10000</georss:elev><category label="Age" term="Past week"/></entry>
  <entry><id>urn:earthquake-usgs-gov:us:2010tvb3</id><title>M 6.0, South Indian
Ocean</title><updated>2010-03-14T20:33:10Z</updated><link rel="alternate"
type="text/html"
href="http://earthquake.usgs.gov/earthquakes/recenteqsww/Quakes/us2010tvb3.php"/><
summary type="html"><![CDATA[<p>Sunday,
March 14, 2010 20:33:10 UTC<br>Monday, March 15, 2010 02:33:10 AM at
epicenter</p><p><strong>Depth</strong>: 10.00 km (6.21
mi)</p>]]></summary><georss:point>-2.7634 83.6777</georss:point><georss:elev>-
10000</georss:elev><category label="Age" term="Past week"/></entry>
</feed>
```

Struktura Atom kanálu USGS je oproti RSS kanálu ORFEUS komplikovanější. Na začátku je opět v hlavičce definováno, že se jedná XML verze 1.0. Prvním velkým rozdílem ve struktuře je absence elementů rss a channel. Místo nich je použit element feed, doplněný o deklarace jmenných prostorů Atom a GeoRSS. Podobně jako u RSS následují informace o samotném kanálu. V případě RSS byly povinné tři údaje – název, odkaz na domovskou stránku a slovní popis obsahu kanálu. První dva údaje jsou u Atomu totožné – název kanálu definuje element title a odkaz element link. Popis obsahu se skrývá v elementu subtitle (podtitul). K povinným údajům přibyly další tři údaje - datum a čas poslední aktualizace kanálu (element updated), autor (element author) a unikátní identifikační znak (element id).

Jednotlivá zemětřesení tentokrát spadají pod elementy entry. Každé zemětřesení má přiřazený svůj jedinečný identifikační znak. Element title opět obsahuje magnitudo a název oblasti. Čas se přesunul do elementu updated. Po kliknutí na titulek je uživatel přesměrován na webové stránky USGS s mnohem podrobnějšími údaji o daném zemětřesení. Nejdůležitější informace se nacházejí v elementu summary, který odpovídá elementu description u RSS. Pro lepší představu o poloze epicentra, je kanál doplněn o obrázek Země s vyznačeným epicentrem. Kanál USGS je navíc obohacen o GeoRSS. Použita je specifikace GeoRSS Simple. Celá struktura působí oproti kanálu ORFEUS velmi nepřehledně, alespoň částečným řešením by bylo zalamování řádků za každým elementem.

3.5 Aplikace USGS

Aplikace funguje na stejném principu jako v případě RSS kanálu ORFEUS. Jediné rozdíly spočívají v elementech, z nichž jsou čerpány údaje a v odlišných úpravách některých údajů před jejich uložením do databáze.

3.5.1 Zdrojový kód

V této podkapitole již nebudou zmiňovány postupy, které byly detailně rozebrány v popisu zdrojového kódu první aplikace, ale pouze odlišnosti a nové problémy. Prvním rozdílem oproti RSS je fakt, že data tentokrát nelze čerpat pouze z jednoho elementu. U RSS stačilo načíst všechny informace z elementu description, tomu odpovídá element summary, ale ten má komplikovanější podobu a celý jeho obsah je tvořen sekci CDATA, která obsahuje HTML kód. Pro další zpracování je sekce CDATA nevhodná. Údaje musíme získat z jiných

částí dokumentu. Využijeme celkem čtyři elementy – title, updated, georss:point a georss:elev.

```
XmlNodeList entry = xDoc.GetElementsByTagName("entry");
XmlNodeList titulek = xDoc.GetElementsByTagName("title");
XmlNodeList updated = xDoc.GetElementsByTagName("updated");
XmlNodeList latlon = xDoc.GetElementsByTagName("georss:point");
XmlNodeList elev = xDoc.GetElementsByTagName("georss:elev");
```

Jméno oblasti získáme jednoduše pomocí metod Split() a Trim() z elementu title. U Atom kanálu USGS se objevil problém v podobě častého výskytu apostrofu ve jménech oblastí. Apostrof způsoboval chybu při ukládání dat, protože se špatně vyhodnocoval SQL příkaz. Problém řeší metoda Replace(). Pokud se ve jménu oblasti vyskytne apostrof, dojde k jeho nahrazení dvěma apostrofy. SQL příkaz je pak vyhodnocen správně a do databáze se uloží záznam s pouze jedním apostrofem. U kanálu ORFEUS se jména s apostrofem za celou dobu sběru dat nevyskytla, i když se jednalo o data převzatá právě od USGS.

```
string[] titleSplit = title[i].InnerText.Split(',');
string loc = titulekSplit[1];
string locationTrimmed = loc.Trim();
string location = locationTrimmed.Replace("'", "'");
```

Další problém se objevil u zeměpisných souřadnic. Jako zdroj souřadnic slouží element georss:point. Georss:point se na rozdíl od ostatních elementů neobjevuje na počátku souboru u popisu kanálu. Z toho vyplývá, že první výskyt tohoto elementu je na nulté pozici. Pokud by tedy při úpravě bylo v závorce ponecháno pouze i, nastavené na hodnotu 1, ukládaly by se do databáze zeměpisné souřadnice, které by patřily až k následujícímu zemětřesení. Proto se při úpravě musí používat hodnota i-1. To samé platí i u hloubky hypocentra.

```
string[] latlonSplit = latlon[i-1].InnerText.Split(' ');
string latitudeStr = latlonSplit[0];
string longitudeStr = latlonSplit[1];
```

Hloubka hypocentra v elementu georss:elev je uváděná v metrech. Metry musíme převést na kilometry, aby si hodnoty z obou aplikací odpovídaly.

```
string[] elevation = elev[i-1].InnerText.Split('-');
float depth = float.Parse(elevation[1]) / 1000f;
```

Žádné další zásadní rozdíly se v druhé aplikaci neobjevují. Data jsou ukládána totožným SQL příkazem. Jediným rozdílem, je že jako proměnná source je použit řetězec usgs.

3.6 Omezení aplikací

Obě aplikace mají jedno velké omezení. V případě, že dojde ke změnám ve struktuře kanálů, aplikace s největší pravděpodobností přestanou správně fungovat. Jediným řešením je editace zdrojového kódu.

Před spuštěním aplikací musí být jako desetinný oddělovač nastavena tečka. Desetinný oddělovač se v prostředí operačního systému Windows nastavuje v Místních a jazykových nastaveních.

3.7 Úprava dat v databázi

Agentura USGS poskytuje ve svém kanálu data o zemětřeseních, které se odehrály během posledního týdne. Aplikace proto musí být spouštěna minimálně jednou za sedm dnů, jinak by docházelo ke ztrátám dat. U projektu ORFEUS je situace odlišná. Časový úsek, který data pokrývají, není přesně dán. Limitem je zobrazení dat o posledních 30 zemětřeseních, což zhruba odpovídá době jednoho týdne až dvaceti dnů. Jako ideální se opět jevilo spouštění aplikace každý týden. Toto rozhodnutí však s sebou nese i problém, kterým je vznik duplicitních záznamů v databázi. Ukládání duplicitních záznamů by se dalo předejít přímo v aplikaci, tak že by před uložením záznamu u každého zemětřesení došlo ke kontrole, zda už databáze totožný záznam neobsahuje. Tím by však došlo ke zbytečnému zkomplikování celého programu. Odstraňování duplicit je proto řešeno až v prostředí Microsoft Office Access. [7]

K vyhledání a odstranění duplicitních hodnot lze použít buď vestavěné nástroje v MS Access, nebo vlastní SQL dotaz. Vhodnější volbou je vlastní SQL dotaz, protože Průvodce vyhledávacím dotazem na duplicitní položky zobrazí duplicitní dvojice, které je třeba ručně odstraňovat. Níže uvedený dotaz uloží všechny unikátní záznamy do nové tabulky.

```
SELECT DISTINCT Zemetreseni.* INTO Zemetreseni_upr
FROM Zemetreseni;
```

Aby bylo možné data z obou zdrojů mezi sebou porovnat, musí být odstraněny také záznamy, které netvoří s žádným záznamem z druhého zdroje dvojici. Za dvojici jsou považovány dva záznamy, představující totožné zemětřesení. V tomto případě se nejedná o duplikáty, protože u obou záznamů bude uvedený jiný zdroj dat a s největší pravděpodobností se budou mírně lišit i zaznamenané hodnoty.

K rozlišení, zda se jedná o totožné zemětřesení, je ideální využít rozdílů časů, kdy se zemětřesení odehrála. Pokud rozdíl nepřesáhne stanovenou hranici, budou zemětřesení vyhodnocena jako totožná. Tato podmínka není sama o sobě dostačující a musí být doplněna o další zpřesňující podmínky. Ke zpřesnění je možné použít rozdílů v zeměpisných délkách a šířkách. Opět se stanoví mez, kterou hodnota rozdílu nesmí překročit. Ve výsledku budou jako totožná zemětřesení označeny pouze ty záznamy, které splní všechny tři podmínky.

V prvním kroku se data rozdělí do dvou tabulek pomocí jednoduchých SQL dotazů. SQL dotaz vybere všechny záznamy se zdrojem s hodnotou usgs respektive orfeus. Uveden je pouze dotaz týkající se záznamů z Atom kanálu USGS. Druhý dotaz pro data projektu ORFEUS bude vypadat obdobně.

```
SELECT * INTO USGS
FROM Zemetreseni_upr
WHERE (((Zemetreseni_upr.source)='usgs'));
```

Jakmile jsou záznamy rozděleny do dvou tabulek, je možné implementovat dříve zmíněné podmínky. Opět vzniknou dva SQL dotazy. Oba dotazy budou sice fungovat úplně stejně, ale jeden z nich bude ukládat do nové tabulky s filtrovanými daty projektu Orfeus a druhý bude data ukládat do nové tabulky s daty agentury USGS. Samozřejmě by bylo možné ukládat data do jedné tabulky, ale dvě nezávislé tabulky usnadní následný export dat před jejich analýzou.

V níže uvedené ukázce dotazu je mez u rozdílu časů nastavena na pět sekund a meze u rozdílů souřadnic jsou nastaveny na jeden stupeň. U všech tří podmínek je využívána absolutní hodnota rozdílu. Uvedená implementace má jedno nepříjemné omezení. Jestliže se zemětřesení odehrálo v blízkosti hranice zemských polokoulí a oba zdroje dat jej umístily na jinou polokouli, dojde ke špatnému vyhodnocení podmínky. Pro lepší představu poslouží následující příklad. Agentura USGS uvedla u zemětřesení zeměpisnou šířku -179,93 a ORFEUS uvedl u stejného zemětřesení hodnotu 179,84. Podle stanovené podmínky od sebe

hodnoty odečteme a absolutní hodnota rozdílu vyjde 359,77, což samozřejmě podmínku nespĺňuje. V realitě je, ale situace jiná a uvedené dvě místa od sebe rozhodně nejsou vzdálená o více než jeden stupeň. Stejná situace nastane i v případě hranice v podobě nultého poledníku nebo rovníku u zeměpisné šířky, s tím rozdílem, že podmínka bude vyhodnocena správně, pokud nebude mez větší než 1 stupeň. Pravděpodobnost, že se tento problém objeví je sice malá, ale i tak je dobré výsledky filtrace překontrolovat.

```
SELECT USGS.location, USGS.datum, USGS.cas, USGS.latitude, USGS.longitude,  
USGS.depth, USGS.mag, USGS.source INTO USGS_F1d  
FROM ORFEUS, USGS  
WHERE Abs(DateDiff("s",[ORFEUS].[cas],[USGS].[cas]))<5 AND  
Abs([ORFEUS].[latitude]-[USGS].[latitude])<1 AND Abs([ORFEUS].[longitude]-  
[USGS].[longitude])<1;
```

Po výše uvedených úpravách jsou již data připravena k samotnému vyhodnocení rozdílů, kterému bude věnována následující kapitola.

4 Analýza rozdílů v datech

Již v úvodu bylo zmíněno, že cílem práce není určit, který ze zdrojů poskytuje přesnější data, protože to není možné. Proto se tato kapitola zabývá pouze analýzou rozdílů v získaných datech a jejich možnými příčinami.

Data byla pomocí vytvořených aplikací ukládána ve dvou obdobích - od února 2009 do dubna 2009 a od července 2009 do října 2009. Během sběru dat došlo ke dvěma výpadkům na straně projektu ORFEUS. RSS kanál nebyl v období od 15.3. – 23.3. 2009 a 11.4 – 14.4. 2009 dostupný. V získaných datech se však výpadek nijak neprojevil. V případě, že by došlo k podobnému výpadku na straně USGS, znamenalo by to ztrátu dat za dané období. Ve zpětném pohledu se však jako lepší řešení jeví automatické spouštění aplikací na serveru.

Za celé období došlo k uložení celkem 604 záznamů o zemětřeseních, přičemž 330 z nich pocházelo z RSS kanálu projektu ORFEUS a 274 z Atom kanálu agentury USGS. Po úpravě dat popsané v předešlé kapitole zůstalo v databázi celkem 122 dvojic záznamů.

Samotnou analýzu by bylo možné provést pomocí SQL dotazů přímo v prostředí Microsoft Office Access, ale v rámci zjednodušení byla upravená data exportována do formátu XLS a analyzovaná v uživatelsky příjemnějším prostředí tabulkového editoru Microsoft Excel. Data z obou zdrojů jsou načtena do dvou samostatných listů a samotné analýze jsou věnovány listy další. Analýza je zaměřena na rozdíly v čase, poloze, hloubce a v magnitudu zemětřesení. Podle původního plánu měla být data vyhodnocena i z hlediska oblastí, ale od tohoto plánu se nakonec upustilo, protože naprostá většina zemětřesení se odehrála v oblasti jihovýchodní Asie a Tichého oceánu. Analýza by tudíž měla malou výpovědní hodnotu.

Následující podkapitoly jsou doplněny o tabulky čítající vždy pět záznamů s extrémními hodnotami. Kompletní tabulky lze nalézt v přílohách v souboru `analiza.xls`. Všechny analyzované zemětřesení je možné zobrazit v aplikaci Google Earth pomocí KML souboru `zemetreseni.kml`, který se rovněž nachází v přílohách. Součástí příloh je také aplikace, s jejíž pomocí byl KML soubor vygenerován. Aplikace je napsána v jazyce C#.

4.1 Rozdíly v časech

Rozdíly v určení přesného času, kdy se zemětřesení odehrálo, se v rámci získaných dat pohybují v rozmezí 1 sekundy a 16 sekund. Celkem u 48% všech zaznamenaných zemětřesení uvedly oba zdroje odlišný čas, přičemž u 27% záznamů přesáhl rozdíl hranici 1 sekundy. Hranici 5 sekund přesáhlo pouze 7 záznamů (6%). U těchto zemětřesení již vyvstává na mysl otázka, zda se skutečně jedná o dvojici totožných zemětřesení nebo zda jde o dva různé otřesy během krátkého časového úseku. Možnost selhání SQL dotazu se minimalizovala pečlivou ruční kontrolou. Jako velmi pravděpodobný zdroj takto velkých rozdílů se jeví různé zdroje původních dat nebo různé verze dat. Bohužel se však jedná pouze o domněnku, protože USGS v rámci Atom kanálu neposkytuje informace o původním zdroji.

ID	Oblast	Datum	Rozdíl časů
97	Owen Fracture Zone region	29.8.2009	0:00:16
7	Hokkaido, Japan, region	28.2.2009	0:00:12
113	New Britain, Papua New Guinea	30.9.2009	0:00:06
48	Yunnan, China	9.7.2009	0:00:06
81	Southeast of Honshu, Japan	12.8.2009	0:00:05

Tab. 3: Extrémní hodnoty v rozdílech časů

4.2 Rozdíly v poloze epicenter

V případě polohy epicentra nastává oproti času poněkud odlišná situace. V získaných datech neexistuje jediný záznam, který by měl u obou zdrojů uvedenou stejnou zeměpisnou šířku nebo délku. Situace je však odlišná i v tom, že na první pohled lze spatřit jednu z příčin rozdílů a tou je odlišné zaokrouhlování hodnot. Zatímco zeměpisné souřadnice projektu ORFEUS jsou zaokrouhlovány na jedno desetinné místo, souřadnice z Atom kanálu USGS jsou zaokrouhlovány na čtyři desetinná místa. Z toho vyplývá, že poloha epicenter u dat USGS je určena přesněji než u dat projektu ORFEUS. To ale neznamená, že jsou to data stoprocentně správná. V důsledku zaokrouhlení na menší počet desetinných míst dojde k nárůstu chyby v určení polohy. Na první pohled se sice zdá, že nárůst chyby je malý, ale pouhé rozdílné zaokrouhlení vede k posunutí polohy epicentra o několik kilometrů.

V analýze byl použit vzorec určený k výpočtu vzdálenosti dvou bodů na zemském povrchu na základě jejich souřadnic. Vzorec vychází ze sférické trigonometrie. U vzdálenosti do přibližně 8 kilometrů můžeme obecně říct, že příčinou je zaokrouhlování hodnot u dat z RSS ORFEUS. Toto tvrzení platí u naprosté většiny záznamů, ale objevuje se i několik

záznamů, u kterých zaokrouhlení neodpovídá přesnějším datům USGS, neboť došlo k zaokrouhlení dolů místo nahoru.

U záznamů s vypočítanou vzdáleností větší než 8 kilometrů se již objevují rozdíly v souřadnicích větší než 1 desetina stupně, u kterých nelze za příčinu považovat zaokrouhlování. I tak ale rozdíly nejsou až na výjimky nijak extrémní. U 75% záznamů je vzdálenost mezi epicentry menší než 10 kilometrů. Průměrná hodnota vzdáleností činí 9,4 km.

Největší vypočítanou vzdálenost mezi epicentry mělo zemětřesení v oblasti ostrovů Tonga v Tichém oceánu. Vzdálenost dosahuje téměř 82 kilometrů, což je pro srovnání vzdálenost mezi Ostravou a Olomoucí. Vzhledem k faktu, že rozdíl v časech je pouhá 1 sekunda, se jako pravděpodobnější příčina opět jeví rozdílný původní zdroj nebo rozdílná verze dat. Další extrémní hodnoty zobrazuje tabulka.

ID	Oblast	Datum	Vzdálenost (v km)
27	Tonga Islands region	19.3.2009	81.79
49	North Pacific Ocean	9.7.2009	63.24
10	Fiji Islands region	1.3.2009	56.51
117	Tonga Islands	1.10.2009	54.08
61	Gulf of Alaska	25.7.2009	38.20

Tab. 4: Extrémní hodnoty vzdáleností epicenter

4.3 Rozdíly v hloubkách hypocenter

U hloubky hypocentra zemětřesení se objevuje stejný problém jako u polohy epicenter. Hloubky jsou v datech agentury USGS uváděny s přesností na jedno desetinné místo. U projektu ORFEUS jsou hodnoty uváděny v podobě celých čísel, z čeho plyne jedna z příčin rozdílů.

Průměrná hodnota rozdílů dosahuje 9,8 km. 75% všech záznamů vykazuje rozdíl menší než 10km. Tři z pěti extrémů se již objevily v tabulce extrémních rozdílů časů, což by mohlo nahrávat myšlence, že se opravdu jedná o dva různé otřesy.

ID	Oblast	Datum	Rozdíl hloubek (v km)
97	Owen Fracture Zone region	29.8.2009	162.9
7	Hokkaido, Japan, region	28.2.2009	93.7
49	North Pacific Ocean	9.7.2009	59.4
117	Tonga Islands	1.10.2009	47
113	New Britain, Papua New Guinea, region	30.9.2009	46.8

Tab. 5: Extrémní hodnoty v rozdílech hloubek hypocenter

4.4 Rozdíly v hodnotách magnituda

U hodnot magnituda je důležité si uvědomit, že používané stupnice jsou logaritmické. Pokud by tedy rozdíl magnitud z obou zdrojů dosáhl hodnoty 1, znamenalo by to přibližně 30násobný rozdíl v reálné síle zemětřesení. U hodnoty 2 by se jednalo již o 1000násobný rozdíl. Až na několik výjimek se však takto extrémní rozdíly v datech neobjevily.

Druhým důležitým poznatkem je, že oba zdroje u některých záznamů nepoužívají totožné stupnice a to i v případech, kdy je v datech uveden stejný původní zdroj. Agentura USGS uvádí, že preferovanou stupnicí v jejich datech je stupnice momentového magnituda (MW), ale i tak se velmi často v rámci jejich dat objevují údaje v podobě magnituda na Richterově stupnici určeného z objemové P vlny (Mb). Data získatelná z RSS a Atom kanálu obou organizací neposkytují informace o použité stupnici. Lze je dohledat až v podrobnějších informacích na webových stránkách.

V rámci získaných dat činil průměrný rozdíl 0,1, přičemž u 86% záznamů nepřesáhl rozdíl hodnotu 0,2. Extrémní hodnoty rozdílu ukazuje tabulka. Největší rozdíl v magnitudu se vyskytl u zemětřesení v oblasti Filipín ze dne 16.7.2009 a dosáhl hodnoty 0,7. V systému Wilber II je k tomuto zemětřesení možné dohledat podrobnější informace, v katalogu USGS bohužel nikoliv. Zajímavostí je ale fakt, že jako původní zdroj dat je uveden NEIC spadající pod USGS a je velmi nepravděpodobné, že by byl pro data v Atom kanálu USGS použit zdroj jiný. Nabízí se otázka, jak je možné, aby rozdíl nabyl tak vysokou hodnotu, když se jedná o data ze stejného původního zdroje. Jedním z možných vysvětlení je, že hodnota magnituda byla ještě dodatečně pozměněna, což se v datech z RSS kanálu ORFEUS neprojevílo. V tabulce se opět objevuje zemětřesení ze severní části Tichého oceánu (ID 49), které vykazuje extrémní rozdíly ve všech údajích s výjimkou času.

ID	Oblast	Datum	Rozdíl magnituda
55	Mindanao, Philippine Islands	16.7.2009	0.7
18	Off coast of central America	10.3.2009	0.6
15	Timor, Indonesia, region	7.3.2009	0.5
49	North Pacific Ocean	9.7.2009	0.5
10	Fiji Islands region	1.3.2009	0.4

Tab. 6: Extrémní hodnoty v rozdílech magnituda

5 Data a služby poskytované agenturou USGS

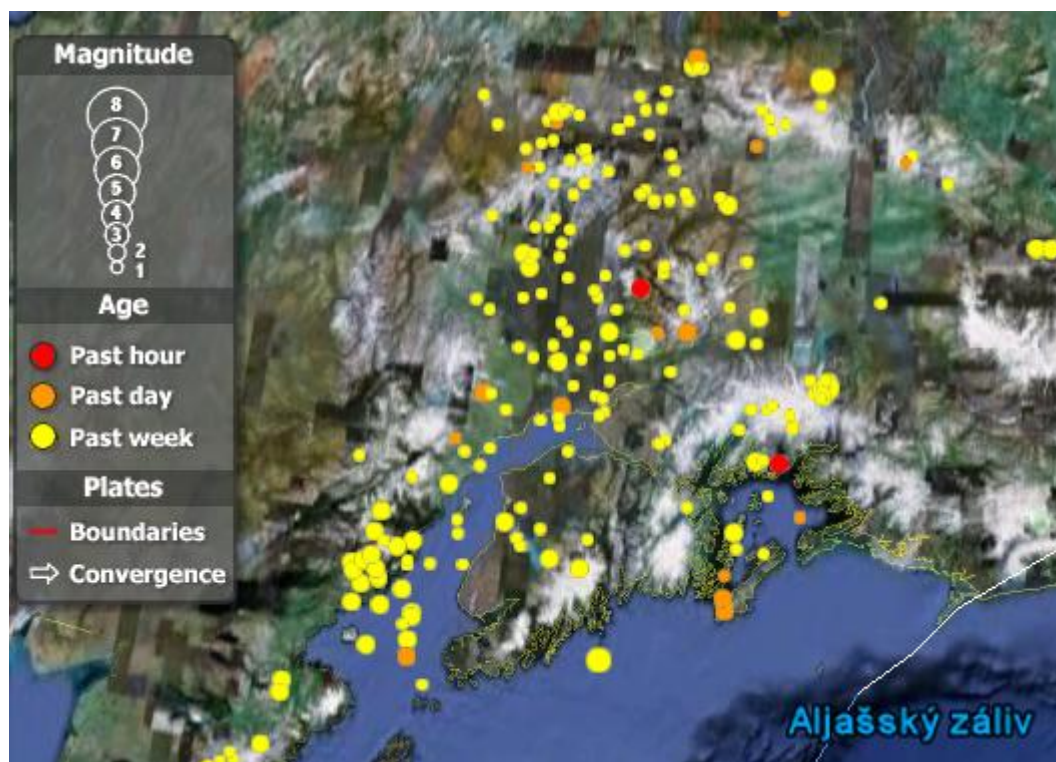
Oba zdroje dat neposkytují data pouze formou RSS nebo Atom kanálů. V této kapitole se blíže podíváme na nabídku dat a služeb agentury USGS. Další kapitola bude věnována nabídce projektu ORFEUS.

5.1 Google Earth KML

Agentura USGS poskytuje data o posledních zemětřeseních i formou souborů KML. V galerii Google je jeden z těchto KML souborů vedený jako Real-time Earthquakes. Markéta Průšová jej zmínila ve své bakalářské práci, která se zabývala vizualizací dat projektu ORFEUS v prostředí Google Earth, což byl jeden z podnětů ke vzniku této práce.

K dispozici jsou dva KML soubory. Po načtení do Google Earth oba zobrazují místa výskytu zemětřesení za poslední týden. Zobrazovány jsou zemětřesení s magnitudem větším než 1. Důležité je dodat, že takto detailní data jsou k dispozici pouze pro oblast Spojených států. Světová zemětřesení s magnitudem menším než 4 zobrazovaná nejsou. Zemětřesení v rámci Spojených států jsou aktualizovány každých 5 minut. Zbytek světa je aktualizován přibližně každých 30 minut. Jediný rozdíl mezi těmito dvěma soubory spočívá v barvě použitých symbolů.

U prvního souboru (<http://earthquake.usgs.gov/earthquakes/catalogs/eqs7day-age.kmz>) jsou symboly barevně odlišeny podle stáří daného zemětřesení. K rozlišení slouží tři barvy. Červená znázorňuje zemětřesení, která se odehrála během poslední hodiny. Oranžová barva představuje poslední den a žlutá poslední týden. K rozlišení magnitudy jsou použity rozdílné velikosti kruhových symbolů. Jak již zmínila Markéta Průšová ve své práci, toto rozhodnutí je nevhodné, neboť rozdíly ve velikostech symbolů jsou velmi malé a rozlišit například zemětřesení s magnitudem 5 od zemětřesení s magnitudem 6 je téměř nemožné. V krajním případě může uživatel k přeměření velikostí symbolů využít poloprůhlednou legendu. Otázkou zůstává, zda toto byl skutečný záměr autorů. Pomoci může také možnost filtrace zobrazených zemětřesení na základě velikosti magnitudy. Nechybí také znázornění hranic litosférických desek. [11]



Obr. 3: Ukázka Real-time Earthquakes

Na uvedeném obrázku lze vidět problém spojený s velikostí symbolů. Po kliknutí na některý ze symbolů se zobrazí základní informace o zemětřesení v podobě data, času, hloubky a hodnotě magnituda. Součástí je také odkaz na webové stránky USGS s podrobnějšími informacemi.

Druhý soubor (<http://earthquake.usgs.gov/earthquakes/catalogs/eqs7day-depth.kmz>) zobrazuje stejná data a používá i stejné symboly, ale tentokrát barevně odlišené podle hloubky hypocentra. K rozlišení je využito šesti barev.

Barvy symbolů podle hloubky hypocentra (v km)					
0 – 33	33 – 70	70 – 150	150 – 300	300 – 500	500 – 800

Tab. 7: Barvy symbolů podle hloubky hypocenter

Během zkoumání KML souborů v prostředí Google Earth byl objeven odkaz na další soubor, který není uvedený přímo na webových stránkách USGS. (http://earthquake.usgs.gov/eqcenter/catalogs/historic/eqs-1970_src.kmz). Po načtení soubor zobrazuje historická zemětřesení zaznamenaná od roku 1970. Všechna zemětřesení jsou zobrazována šedou barvou. Chybí bohužel legenda, ale je zřejmé, že velikosti symbolů opět

odpovídají hodnotě magnituda. Zmíněný soubor načte najednou přibližně 20 000 značek, což nemusí méně výkonné počítače zvládnout.

5.2 RSS a CSV

Na webových stránkách <http://earthquake.usgs.gov/earthquakes/catalogs/> je kromě tří kanálů Atom k dispozici také kanál RSS s názvem Earthquake ShakeMaps. Místo obrázku Země s vyznačeným epicentrem zemětřesení je součástí každého záznamu mapka, znázorňující intenzitu otřesů. Bohužel je však mapka velmi malá a nelze z ní mnoho vyčíst. Kanál ale obsahuje odkaz na stránky USGS s mapami v daleko čitelnější velikosti. Intenzita otřesů a míra potenciálních škod je zobrazena pomocí barevné stupnice, kde červená barva představuje nevyšší intenzitu a modrozelená naopak nejnižší.

Ze stránek USGS je možné stahovat také CSV soubory. Jedná se o textové soubory, v nichž jsou jednotlivé hodnoty oddělovány čárkou. CSV soubory jsou snadno zpracovatelné pomocí aplikací. K dispozici jsou ve třech verzích. Data jsou rozdělena do tří souborů podle stáří zemětřesení – hodina, den a týden. Všechny tři verze obsahují data o zemětřeseních s magnitudem větším než 1.

Data z posledních sedmi dnů lze získat také v podobě archivovaného XML souboru. Oproti datům z Atom kanálů jsou tato data doplněna o další podrobnější informace jako například typ magnituda nebo počet stanic podílejících se na lokalizaci zemětřesení. Totožná data nabízí i formát CUBE.

5.3 Vizualizace dat

Kromě vizualizace dat v prostředí Google Earth nabízí USGS i vizualizace přímo na svých webových stránkách. Vizualizované oblasti jsou rozděleny na dvě části – Spojené státy (<http://earthquake.usgs.gov/earthquakes/recenteqsus/>) a zbytek světa (<http://earthquake.usgs.gov/earthquakes/recenteqsww/>). Mapy mají tři úrovně přiblížení. U mapy světa je u první úrovně přiblížení viditelná celá zeměkoule. Po kliknutí do mapy se zobrazí celý kontinent a po dalším kliknutí se zobrazí oblast o rozměrech 10x10 stupňů. Velký nedostatek spočívá v nemožnosti mapu zpět oddálit, nezbyvá tedy nic jiného než použít funkci Zpět v prohlížeči. U mapy Spojených států je zobrazované území podrobnější. Při

maximálním zvětšení se zobrazí oblast o velikosti 2x2 stupňů. I tak se ale stává, že se na mapě objevují shluky symbolů, které znemožňují přesné kliknutí na vybraný symbol. Tento problém se týká především oblastí s vysokou seismickou aktivitou, jako je Kalifornie a Aljaška.

Zemětřesení jsou na mapách zobrazována formou čtvercových symbolů, barevně odlišených podle stáří. Velikost čtverce odpovídá hodnotě magnituda. Po kliknutí na symbol se otevře stránka s podrobnými informacemi.

5.4 Systém ENS

Earthquake Notification Service je bezplatná služba, která uživatelům zasílá upozorňující zprávy v případě, že se zemětřesení odehraje v jimi definované oblasti. Zprávy jsou zasílány během pěti minut u zemětřesení v rámci Spojených států a během třiceti minut v rámci zbytku světa. Uživatelské rozhraní má anglickou a španělskou verzi.

Oblasti zájmu se definují v prostředí Google Maps. Je možné využít buď předdefinované regiony, nebo si pomocí polygonů označit oblasti vlastní. Na jednom účtu může mít uživatel nadefinováno více profilů, přičemž pro každý profil lze nastavit odlišné hodnoty minimálního magnituda.

Earthquake Notification Service

[Logout](#)

The screenshot displays the Earthquake Notification Service (ENS) web interface. At the top, there are navigation tabs: 'My Profiles', 'Recent Events', 'My Email Addresses', 'Account Preferences', 'Text Only', and 'Help'. The main area features a map of the United States with a red pin and a polygon defining a region. The map is titled 'United States' and is part of a profile named 'United States (Predefined)'. The interface includes navigation controls, a search bar, and a list of regions. The 'United States' profile is selected, and its settings are displayed below the map, including email address, magnitude thresholds, and active status.

Profile Name	United States	(Predefined)
Send email to:	lojekj@seznam.cz <input checked="" type="checkbox"/>	Add another email
Magnitude:	Day 1.0	Night 1.0
Active	Yes <input checked="" type="radio"/>	No <input type="radio"/>
More Options	None	

Obr. 4: Systém ENS

Upozorňující zprávy jsou zasílány v podobě emailů. K dispozici jsou celkem čtyři formáty – HTML, dlouhá textová verze, krátká textová verze a formát CUBE. Za zmínku stojí fakt, že data zasílaná emailem jsou podrobnější než data poskytovaná prostřednictvím kanálu Atom nebo KML. Kromě základních údajů obsahují data také informace o typu magnituda, vzdálenosti epicentra od nejbližších měst nebo nejistotu určení polohy epicentra a hypocentra.

5.5 System PAGER

PAGER je automatizovaný systém, který z naměřených údajů odhaduje, kolik obyvatel bylo vystaveno otřesům půdy. Odhady jsou poskytovány buď v podobě map ve formátech JPEG a PDF nebo ve formě XML souborů. Kromě počtů obyvatel jsou odhadovány také postižená města. Obyvatelé i města jsou dále rozděleni podle síly otřesů na základě modifikované Mercalliho stupnice.

Odhady jsou na webových stránkách publikovány během třiceti minut po zemětřesení, což je mnohdy rychlejší než první výpovědi očitých svědků. Celý systém je určen především pro organizace nouzové odezvy, vládní složky a média. [17]

6 Data a služby poskytované projektem ORFEUS

Oproti USGS je nabídka poněkud chudší, ale na druhou stranu zahrnuje i neupravená data ze seismologických stanic tak jak byla zaznamenána v době zemětřesení. (<http://www.orfeus-eu.org/Data-info/available.html>). Na výběr jsou k dispozici formáty SEED a miniSEED, což jsou formáty vyvinuté speciálně pro výměnu neupravených digitálních seismologických dat. Veškerá data jsou poskytována zdarma a mají podobu buď kontinuálního záznamu, nebo záznamu konkrétní události. [14]

Totožná data jako v RSS kanálu se nachází i přímo na webových stránkách projektu ORFEUS (<http://www.orfeus-eu.org/Earthquakes/recentquakes.html>). Zcela nepochopitelně se na webových stránkách nenachází jediná zmínka o datech poskytovaných prostřednictvím RSS. Uživatel má takto nulovou šanci RSS kanál vůbec najít.

6.1 Vizualizace dat

Podobně jako USG nabízí ORFEUS na svých stránkách vizualizace poskytovaných dat. Zvolit lze mezi vizualizací v prostředí Google Maps (http://www.orfeus-eu.org/Earthquakes/recentquakes_googlemaps.html) nebo v rámci fyzicko-geografické mapy světa respektive Evropy (<http://www.orfeus-eu.org/Earthquakes/quakemonitor.html>).

V prostředí Google Maps jsou epicentra znázorňována pomocí červených kružnic. Legenda sice chybí, ale po bližším prozkoumání je patrné, že zemětřesení s větším magnitudem jsou znázorněna větší kružnicí. Bohužel je však rozdíl ve velikostech minimální a podobně jako map USGS lze velmi špatně rozeznat zemětřesení s podobným magnitudem. Poslední zaznamenané zemětřesení je znázorněné žlutou hvězdou. Po kliknutí na symbol se zobrazí informace o vybraném zemětřesení spolu s odkazem na podrobnější údaje v prostředí systému Wilber II.

K rychlému vyhledání zemětřesení slouží seznam pod mapovým polem. Se seznamem se pojí i nepříjemný nedostatek v podobě nevhodně zvolené úrovně přiblížení po kliknutí na položku v seznamu. Pokud se zvolené zemětřesení odehrálo například někde v oblasti Tichého oceánu, uživatel se po kliknutí v mapě ztratí, protože se mu zobrazí pouze vodní plocha bez záchytných bodů. Nedostatek neřeší ani pomocná mapa v pravém dolním rohu,

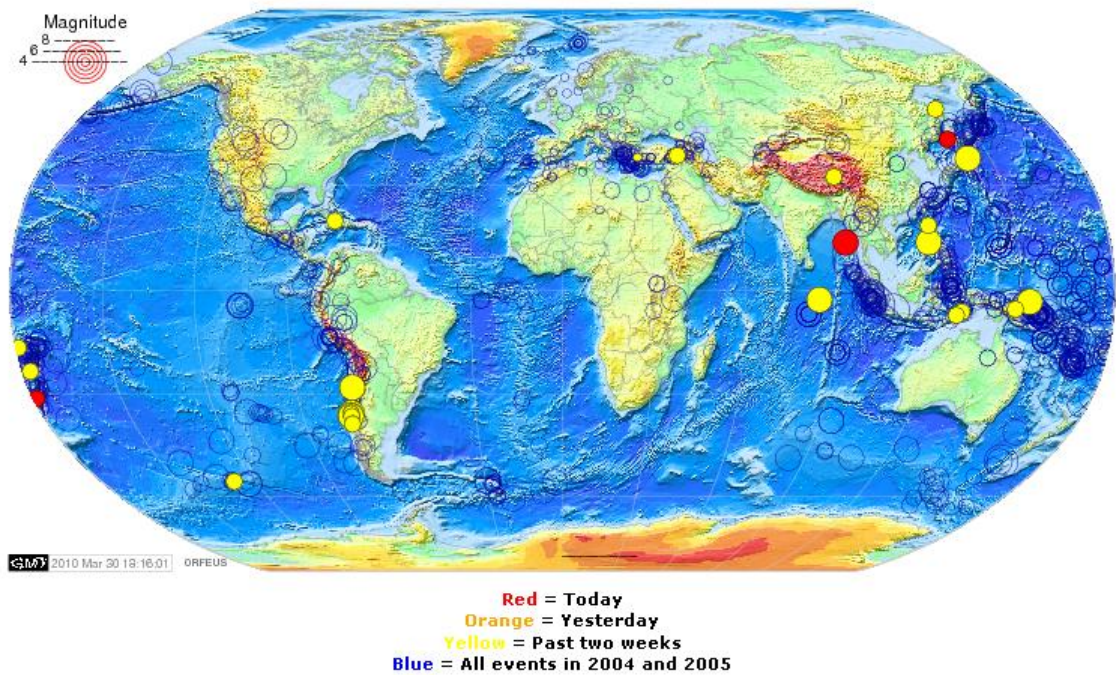
kteřá v daném přiblížení zobrazí také pouze vodní plochu. Řešením může být buď zvětšení mapového pole, nebo změna úrovně přiblížení.



Obr. 5: Vizualizace dat ORFEUS pomocí Google Maps

Velkou výhodou oproti mapám USGS je možnost většího přiblížení, tudíž nehrozí problémy spojené se shlukem symbolů na malé části mapy popsané v minulé kapitole.

To se bohužel nedá říct o vizualizaci v rámci map světa a Evropy. Zde jsou aktuální epicentra znázorněna pomocí kruhových symbolů. Obě mapy se nedají přiblížit a tak v oblastech častého výskytu zemětřesení vznikají nepřehledné shluky kruhů, znemožňující přesné kliknutí na konkrétní symbol. Symboly epicenter jsou rozlišeny velikostně podle magnituda a barevně podle stáří. Velikost symbolů je tentokrát vybrána lépe a i zemětřesení s malými rozdíly magnitud je možné snadno rozeznat. Kromě aktuálních zemětřesení zobrazují mapy i historické záznamy z let 2004 – 2005 v podobě modrých kružnic.



Obr. 6: Vizualizace dat ORFEUS

6.2 Systém Wilber II

Prostřednictvím systému Wilber II lze získat podrobné informace o zemětřeseních a to nejen o těch aktuálních, ale i o historických. Databáze sahá až do druhého čtvrtletí 1988. Po výběru příslušného čtvrtletí ze seznamu jsou odpovídající zemětřesení vyobrazena na mapě světa. K získání podrobných údajů je třeba zemětřesení vybrat kliknutím do mapy. Výběr probíhá na základě zvoleného poloměru výběrové oblasti. Po kliknutí se zobrazí seznam všech odpovídajících zemětřesení, ze kterých lze dále vybrat jedno konkrétní zemětřesení s podrobnými údaji, jejichž součástí je i informace o použité stupnici.

Hlavní síla systému spočívá v nabídce dat přímo ze seismologických stanic v Evropě. U každého zemětřesení je zobrazen seznam stanic, které k němu mají dispozici data. Uživatel si může vybrat požadované stanice a o data zažádat prostřednictvím formuláře. V nabídce je několik datových formátů, včetně zmíněných formátů SEED a miniSEED. Po odeslání formuláře je uživateli na FTP serveru projektu ORFEUS vytvořena osobní složka, do které jsou vyžádána data uložena. Celá operace i s vyřízením žádosti trvá pouze několik minut a je zdarma.

7 Závěrečné zhodnocení

Po důkladném prostudování není možné říci, který zdroj je lepší, protože oba jsou zjevně zaměřené na odlišné cílové skupiny. Zatímco ORFEUS má v první řadě odborný vědecký charakter a jeho hlavním úkolem je poskytování dat, USGS se zaměřuje spíše na širší okruh veřejnosti a na poskytování informací. Následující podkapitoly se věnují zhodnocení z různých úhlů pohledu.

7.1 Rozsah dat

Pokud by se měl uživatel rozhodnout pro jeden ze zdrojů a hlavním faktorem při výběru by byl rozsah dat, jevila by se v celosvětovém měřítku jako vhodnější zdroj agentura USGS. Nicméně ideální by byla kombinace obou zdrojů s eliminací duplicitních zemětřesení, protože ORFEUS nabízí v rámci Evropy větší množství dat než agentura USGS. Pokud by se práce týkala pouze území Spojených států, je agentura USGS bez diskuze vhodnějším zdrojem, neboť poskytuje data o zemětřeseních s hodnotou magnituda větší 1.

Výše zmíněné informace se týkají dat poskytovaných formou RSS respektive Atom kanálů. Pro odborníky z oboru seismologie, kteří potřebují data ze seismologických stanic, je zcela jistě lepší volbou projekt ORFEUS, protože USGS data v této podobě neposkytuje.

Co se týče historických dat, nabízí ORFEUS údaje o zemětřeseních od roku 1988 včetně dat ze seismologických stanic. Agentura USGS poskytuje data od roku 1977.

7.2 Aktuálnost a stabilita

U obou zdrojů dochází k aktualizování dat během několika minut po zemětřesení. USGS uvádí, že data jsou publikována během 5 minut po zemětřesení pro oblast Spojených států a během 30 minut pro zbytek světa. ORFEUS přesné údaje neuvádí.

Z hlediska stability zdrojů byly za celou dobu sběru dat zaznamenány pouze dva několikadenní výpadky u RSS kanálu ORFEUS. K objektivnímu zhodnocení stability by bylo zapotřebí provádět kontroly mnohem častěji, než jen při spouštění aplikací za účelem uložení dat.

7.3 Služby a informace

Nabídka různých služeb, vizualizací a informací je mnohem rozsáhlejší na straně agentury USGS, což je dáno zaměřením obou zdrojů. USGS nabízí kromě služeb a dat zmíněných v kapitole 5, také obrovské množství informací a článků o zemětřeseních, jejichž popis by překročil rámec bakalářské práce. Dvě samostatné sekce webových stránek jsou dokonce věnovány výuce dětí a studentů. Další sekce popisuje například jak se před zemětřesením chránit a jak se během něj chovat. Webové stránky projektu ORFEUS se USGS v tomto směru ani zdaleka nepřibližují, což samozřejmě nelze vzhledem k zaměření považovat za zápor.

Vizualizace dat u obou zdrojů v podobě map nejsou příliš zdařilé a potýkají se s nepříjemnými problémy, které zhoršují jejich čitelnost a ovladatelnost. V tomto ohledu je na vyšší úrovni aplikace Ing. Františka Klímka, která k vizualizaci zemětřesení používá data z RSS kanálu projektu ORFEUS.

Závěr

Cílem bakalářské práce bylo srovnat dva datové zdroje poskytující data o posledních zemětřeseních. Úvodní část práce se zabývá obecnou problematikou zemětřesení a jejich měření. Po prostudování KML souboru s posledními zemětřeseními z galerie Google bylo zjištěno, že autorem je americká vládní vědecká agentura USGS, která poskytuje data v řadě dalších formátů. Namísto KML byl nakonec využit formát Atom, neboť má k formátu RSS formátu blíže.

Po nastudování formátů RSS a Atom byly v jazyku C# vytvořeny dvě aplikace sloužící k ukládání dat z datových zdrojů do databáze. Uložená data posloužila k analýze rozdílů. Zajímavým zjištěním byl fakt, že ačkoli je u obou poskytovatelů dat často uveden stejný původní zdroj, hodnoty se liší.

Závěrečná část práce přibližuje další nabízená data a služby obou zdrojů. Důležitým poznatkem této části je, že oba zdroje mají zjevně odlišné cílové skupiny uživatelů. Agentura USGS se soustředí především na širší okruh veřejnosti. Projekt ORFEUS je určen pro vědecké účely.

V bakalářské práci byly splněny zadané úkoly. Prezentované poznatky mohou posloužit především při rozhodování, který ze dvou zdrojů u dalších prací použít. Kapitola věnující se sběru dat a popisu tvorby aplikací může sloužit jako návod jak získávat data z podobných zdrojů netýkajících se pouze zemětřesení.

V budoucnosti by bylo možné práci rozšířit o zahrnutí dalších zdrojů. Jako vhodný kandidát pro další srovnání se jeví webové stránky Evropsko-středozemního seismologického centra (<http://www.emsc-csem.org>). Další možné rozšíření spočívá také v úpravě aplikací tak, aby dokázaly získávat údaje o použitých stupnicích a původních zdrojích dat, což by mohlo pomoci při analýze rozdílů. S analýzou se pojí také vytvoření SQL dotazů, které by umožňovaly analýzu dat přímo v databázi.

Literatura a internetové zdroje

- [1] Akademie věd ČR, Geofyzikální ústav, Jak fungují seismografy, [online], [2010-02-12], dostupné na WWW: <http://rebel.ig.cas.cz/seismika/Seismograf/sm_text.php>
- [2] CNET News, Rozhovor s Jamesem Goslingem, [online], [2010-02-26], dostupné na WWW: <<http://www.w3.org/XML/>>
- [4] GeoRSS dokumentace, [online], [2010-02-15], dostupné na WWW: <<http://www.georss.org>>
- [5] GOOGLE, Google KML, [online], [2010-02-22], dostupné na WWW: <<http://earth.google.com/kml/>>
- [6] Historie RSS, [online], [2010-02-15], dostupné na WWW: <<http://www.rssboard.org/rss-history>>
- [7] KLÍMEK, F.: Vizualizace seismických dat projektu ORFEUS v prostředí MapServer UMN integrovaná na Live CD, Diplomová práce, VŠB-TUO, Ostrava 2006, 56 stran
- [8] KNAPP, O. Návrh struktury RSS kanálu pro vybrané aplikace IS MU, Bakalářská práce, Masarykova Univerzita, Brno 2007, 46 stran
- [9] Masarykova univerzita, Přírodovědecká fakulta: Přírodní katastrofy a environmentální hazardy, [online], [2010-02-12], dostupné na WWW: <<http://www.sci.muni.cz/~herber/quake.htm>>
- [10] ORFEUS, Webové stránky projektu ORFEUS, [online], [2010-02-12], dostupné na WWW: <<http://www.orfeus-eu.org>>
- [11] PRŮŠOVÁ, M.: Vizualizace epicenter aktuálních zemětřesení v prostředí aplikace Google Earth, 2008
- [12] SHARP, J.: Microsoft Visual C# Krok za krokem, Brno 2006, 528 stran, ISBN 80-251-1156-3
- [13] Stránky převodu jednotek - conVERTER, [online], [2010-02-12], dostupné na WWW: <<http://www.converter.cz/tabulky/richterova-stupnice.htm>>
- [14] University of California, Berkeley, Formát SEED, [online], [2010-03-26], dostupné na WWW: <<http://quake.geo.berkeley.edu/ncedc/SEED-help.html>>
- [15] USGS, Richterova stupnice, [online], [2010-02-15], dostupné na WWW: <<http://earthquake.usgs.gov/learn/topics/richter.php>>
- [16] USGS, O USGS, [online], [2010-02-12], dostupné na WWW: <<http://www.usgs.gov/aboutusgs/>>

- [17] USGS, Systém PAGER, [online], [2010-03-20], dostupné na WWW:
<<http://earthquake.usgs.gov/earthquakes/pager/background.php>>
- [18] USGS, Zásady týkající se magnituda zemětřesení, [online], [2010-02-15], dostupné na
WWW: <http://earthquake.usgs.gov/aboutus/docs/020204mag_policy.php>
- [19] USGS, Zemětřesení s více než 50 000 obětí, [online], [2010-02-12], dostupné na
WWW: <http://earthquakes.usgs.gov/earthquakes/world/most_destructive.php>
- [20] World Wide Web Consortium, XML, [online], [2010-02-22], dostupné na WWW:
<<http://www.w3.org/XML/>>

Seznam obrázků

Obr. 1: Znázornění hypocentra a epicentra	2
Obr. 2: Sýkorův most v aplikaci Google Earth	10
Obr. 3: Ukázka Real-time Earthquakes	29
Obr. 4: Systém ENS	31
Obr. 5: Vizualizace dat ORFEUS pomocí Google Maps	34
Obr. 6: Vizualizace dat ORFEUS	35

Seznam tabulek

Tab. 1: Stupnice EMS-98	4
Tab. 2: Verze RSS	8
Tab. 3: Extrémní hodnoty v rozdílech časů	25
Tab. 4: Extrémní hodnoty vzdáleností epicenter	26
Tab. 5: Extrémní hodnoty v rozdílech hloubek hypocenter	26
Tab. 6: Extrémní hodnoty v rozdílech magnituda	27
Tab. 7: Barvy symbolů podle hloubky hypocenter	29

Přílohy na CD

- bakalarska_prace.doc – text práce ve formátu DOC
- bakalarska_prace.pdf – text práce ve formátu PDF
- zemetreseni.kml – KML soubor vygenerovaný z databáze
- /zdrojove_kody/usgs.cs – zdrojový kód aplikace k ukládání dat USGS
- /zdrojove_kody/orfeus.cs – zdrojový kód aplikace k ukládání dat projektu ORFEUS
- /zdrojove_kody/mdb2kml.cs – zdrojový kód aplikace generující KML z databáze
- /aplikace/usgs.exe – přeložená aplikace k ukládání dat USGS
- /aplikace/orfeus.exe – přeložená aplikace k ukládání dat projektu ORFEUS
- /aplikace/mdb2kml.exe – přeložená aplikace generující KML z databáze
- /databaze/earthquakes.mdb – databáze zemětřesení
- /analyza/analyza.xls – XLS soubor s analýzou rozdílů v datech