

TimeMap - georeferencované historické dokumenty a mapy

Bc. Petr Přidal

Anotace: Zdarma dostupný projekt TimeMap umožňuje hledání dokumentů prostřednictvím webového prohlížeče na základě zvolené zeměpisné oblasti, časového období a klíčových slov. Nabízí i nástroje usnadňující indexaci a zpřístupnění nových dokumentů z vlastních sbírek. Bude popsána struktura tohoto projektu, dostupné pomocné nástroje, ukázka použití a nastíněny možnosti dalšího rozšíření či napojení na knihovní informační systémy.



Úvodní slovo

Se stále narůstajícím množstvím digitalizovaných dokumentů nejrůznějších typů (mapy, rukopisy, staré tisky, fotografie, obrazy, zvukové či video záznamy, trojrozměrné předměty) ve sbírkách mnoha institucí po celém světě vzniká potřeba tyto dokumenty zpřístupnit co nejjednodušší formou široké veřejnosti, čtenářům i badatelům.

Knihovny hromadně převádí celé sbírky do elektronické podoby. Často se digitalizují i dokumenty, které se velmi obtížně vyhledávají pomocí klíčových slov. Těžko hledat například mapu daného území z určité doby pomocí bibliografického záznamu v knihovním systému. Bez znalosti jména autora, přesného roku vydání či vhodných klíčových slov je to se stávajícím systémem indexace obtížné.

Zkombinovat systém plnotextového hledání s určením zeměpisné oblasti a časového rozmezí je právě pro takové dokumenty velmi vhodné. S technickou realizací může pomoci projekt TimeMap - jeden z pilířů mé diplomové práce na téma "Zpracování a zpřístupnění historických dokumentů", kterou dokončuji na Fakultě informatiky Masarykovy univerzity v Brně. Diplomová práce je součástí řešení výzkumného záměru MK00009494301 "Historické fondy Moravské zemské knihovny v Brně" a je vedena Petrem Žabičkou.

Projekt TimeMap

Jádrum projektu TimeMap je vlastní prohlídací systém s názvem TMJava, který pomocí internetového prohlížeče nabízí dokumenty na základě zvolené zeměpisné oblasti na mapě, časového období a klíčových slov. Součástí je ale i sada nástrojů (zejména TMWin), která usnadňuje použití tohoto vyhledávacího systému a především vytváření nezbytných metadat pro nově indexované dokumenty. Kompletní softwarové vybavení je možné získat z webových stránek projektu zdarma. Navíc byla ohlášena dostupnost zdrojových kódů pod svobodnou licencí GPL, TimeMap tedy bude možné si jakkoliv přizpůsobit pro konkrétní aplikaci i na programátorské úrovni.

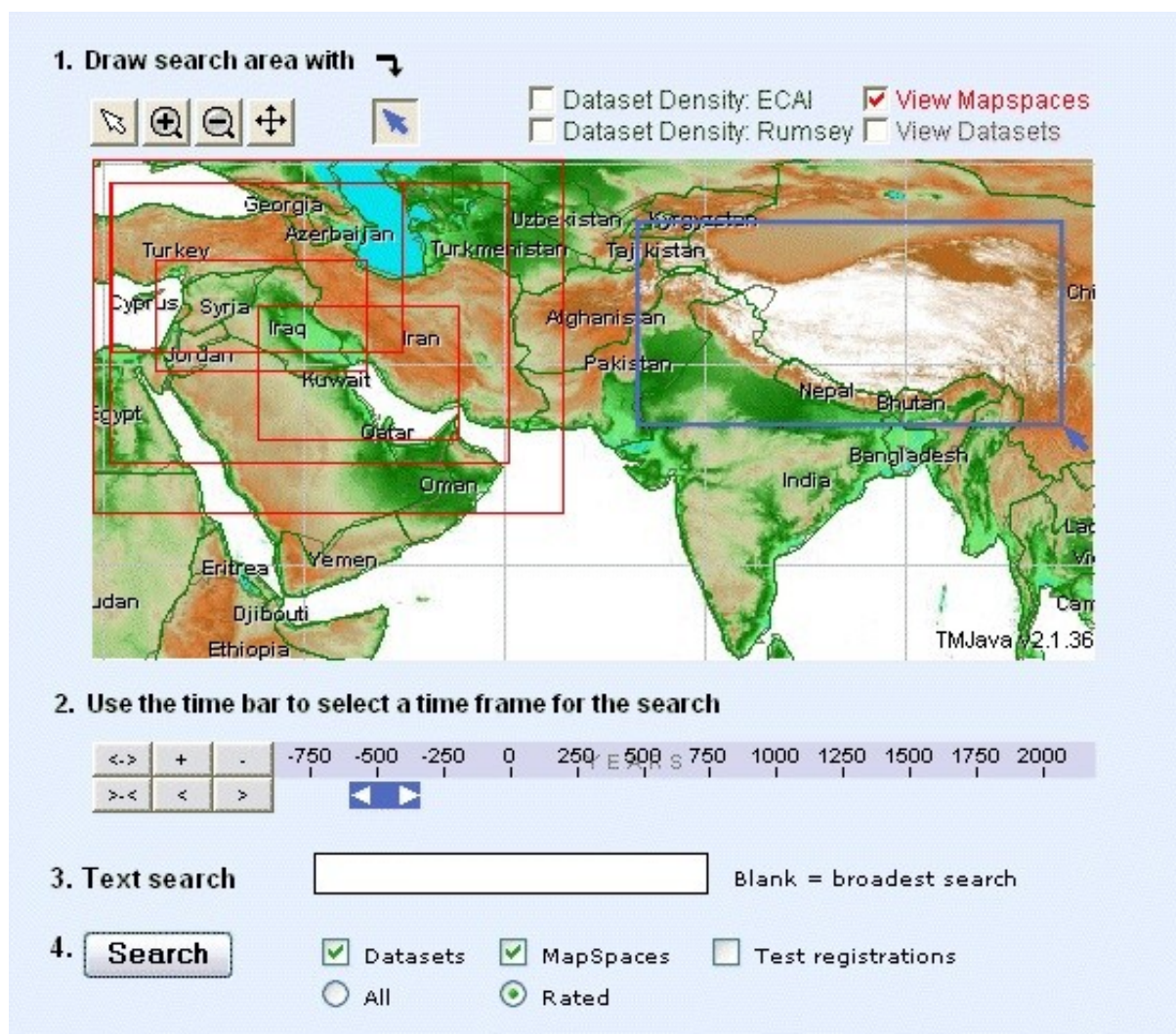
Celý projekt byl původně vytvořen skupinou vývojářů z University of Sydney. Vedením projektu se zabývá Ian Johnson z této univerzity. Na projektu ale participují i další instituce a univerzity z celého světa a TimeMap je používán např. v National Library of Australia, nebo pro zpřístupnění americké David Rumsey Map Collection, San Francisco a dalších.

Technicky je prohledávání řešeno pomocí Java appletu v internetovém prohlížeči, proto je při navštívení webových stránek nejprve nutné vyčkat než se načte speciální rozšíření prohlížeče, které umožňuje vlastní hledání pomocí TimeMapu. Podmínkou funkčnosti je nainstalovaná Java Virtual Machine na počítači uživatele, která by měla být standardně přítomna na všech

nových instalacích Microsoft Windows XP. Do budoucna se počítá i s možností zrušit tuto závislost, pak by se mapa načítala okamžitě a na internetový prohlížeč uživatele by nebyly kladeny žádné speciální požadavky. Použije se podobného rozhraní jako má nyní např. služba Maps od společnosti Google, tzn. technologie AJAX.

Prohledávání pomocí TimeMap je možné využít i bez dostupného připojení k Internetu - to je vhodné např. pro distribuci sbírek na disku CD-ROM - v tzv. *standalone módu*. Pak se načítají do TimeMapu všechna data při prvním spuštění prohlédavače.

Pokud je hledání realizováno on-line, nabízí se naopak tzv. *servlet mód*, při kterém jsou přenášeny jen informace nezbytně nutné k zobrazení konkrétního zvoleného výseku zeměpisné oblasti, pro rozsáhlejší datové zdroje je tento mód nutný.



Obrázek č.1 - Hledání v ECAI Clearinghouse na základě označení území a časového období.

Jak vypadá práce s TimeMap?

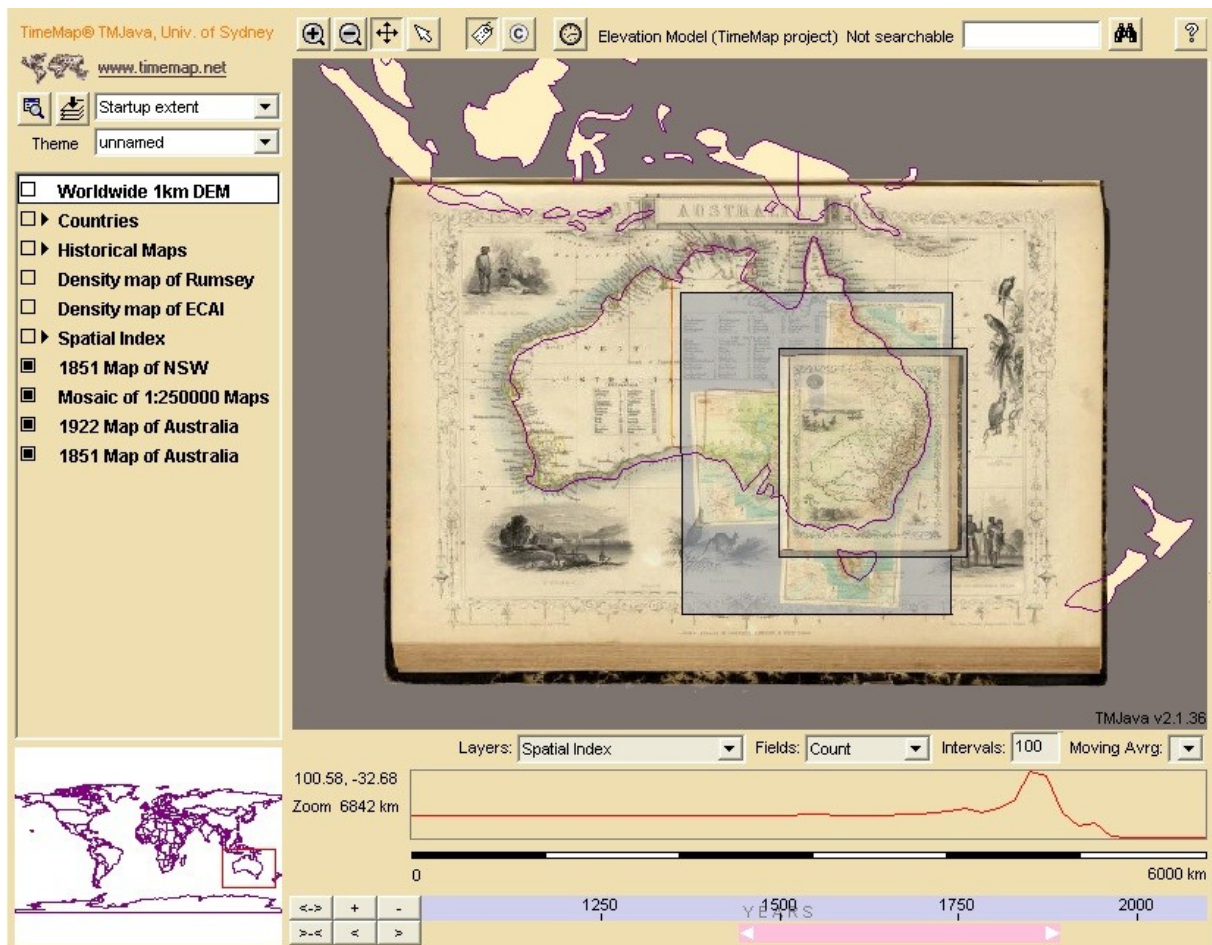
Hlavním prvkem prostředí TimeMap je mapa, na které se uživatel může pohybovat, zvětšovat a zmenšovat výseky lupou a označovat objekty zájmu. Při měnícím se přiblížení mapy mohou být dostupné různé informace, takže při celkovém pohledu se konkrétní město jeví jako tečka, při přiblížení se mohou objevit ulice či městské čtvrti, pokud je mapa pro to připravena.

Druhým důležitým prvkem je lišta pro nastavení časového rozmezí, kdy je možné měnit podrobnost od let, přes měsíce, dny, hodiny až po sekundy.

Posledním prvkem je pole, do kterého je možné zadat frázi pro plnotextové hledání. Prohledávat se budou jen dokumenty splňující předchozí kritéria výběru.

Standardně se data zobrazují jako body, případně jiné symboly na mapě a oblasti (např. pokrytí území konkrétní digitalizovanou mapou) se zobrazují jako obdélníky. Pro hranice oblastí se používá lomené čáry. Každému záznamu nějakého dokumentu na mapě může být přiřazen jak konkrétní časový okamžik (např. datum tisku knihy), tak i rozsah let.

Při vyhledávání se tedy na mapě zobrazí zvolená oblast, provede se výběr časového rozsahu dle zájmu a eventuálně se může použít hledání zadané fráze.



Obrázek č. 2 - Mapy s transparentností, Rumsey Map Collection

TimeMap umožňuje také přímé zobrazení vlastních map v prohlídacím prostředí, nastavitelná je i průhlednost jednotlivých vrstev, zobrazené mapy je tak možné komfortně srovnávat. Podporované jsou mimo jiné mapy ve formátu zdarma dostupného Zoomify, spolupráce s komerčním image serverem MrSID i jinými image servery je také možná.

Vzhled i uspořádání jednotlivých prvků v pracovní ploše programu není fixní, barevné ladění projektů založených na TimeMap i pozice pomocných ovládacích prvků se může lišit dle návrhu designerů prostředí, kde je TimeMap použit.

ECAI (Elektronic Cultural Atlas Initiative)

Elektronický kulturní atlas se snaží shromažďovat historická a kulturní data a umožnit jejich jednoduché prohledávání a zpřístupnění na jednom místě s pomocí webového prohlížeče. Projekt TimeMap je jádrem celého atlasu. Mělo by tak být možné si zobrazovat na mapě informace o významných historických událostech (taženích vojevůdců, obchodních stezkách, válkách), o statistických informacích (jako růstu populace, šíření nemoci) společně s dokumenty dostupnými v knihovnách, muzeích a soukromých sbírkách po celém světě.

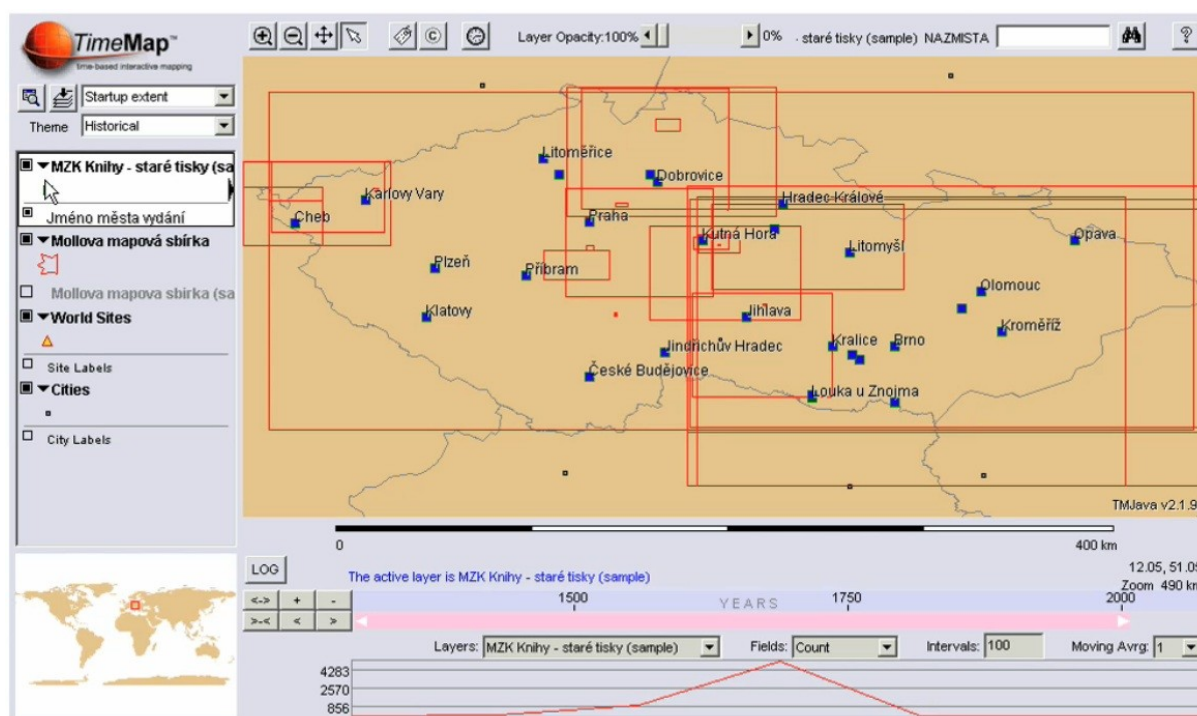
Jde tedy také o jakési centrum sběru metadat pro sbírky dokumentů indexovaných pomocí TimeMap.

Pokud totiž použijete pro svoje dokumenty TimeMap, je velmi jednoduché registrovat metadata týkající se sbírek právě do souborného katalogu (Metadata Clearinghouse) provozovaného ECAI. Můžete tak dát světu vědět o dokumentech, které spravujete.

TimeMap a fondy MZK

Pomocí TimeMap byla zpřístupněna část fondů Moravské zemské knihovny v Brně. Konkrétně se jedná o výběr z Mollovy mapové sbírky (digitalizované mapy jsou dostupné online) a výběr starých tisků publikovaných na území ČR (s vazbou do knihovního systému).

Vše si můžete sami vyzkoušet na adrese uvedené pod článkem. Zatím se jedná o testovací verzi, záznamy budou postupně přibývat i nastavení TimeMapu ještě čeká několik změn k lepšímu.



Obrázek č. 3 – Vzhled prostředí TimeMap s daty historických fondů MZK

Data a metadata

Pro zpřístupnění vlastních sbírek pomocí projektu TimeMap je nejvhodnější použít dodaný nástroj TMWin, který nabízí formou grafického rozhraní snadnější tvorbu metadat nutných pro TimeMap.

Metadata projektu jsou ale čitelné textové XML soubory, není tedy nezbytně nutné grafické prostředí TMWin používat a otevírají se tak možnosti automatického generování těchto souborů. Většinou je ale TMWin nejsnadnější cesta.

TimeMap třídí metadata na tři základní logické celky:

MapSpace popisuje pracovní prostředí prohlédavacího rozhraní, pozici a vzhled jednotlivých prvků (mapy, časové osy, atd.). Odkazuje se také na ostatní dostupná metadata a tím určuje, které informace bude možné prohlédávat.

DataStore je zdrojem dat, která mají být zobrazena, nebo prohlédávána. Daty jsou pro TimeMap převážně záznamy knihovního systému opatřené georeferencí (zeměpisnou šířkou a délkou) a časem, tedy vlastně metadata z knihovnického pohledu. Podporované jsou formáty DBF, CSV nebo ShapeFiles. Také je možné využít SQL serveru a data získávat např. ze serveru Oracle či jiných.

Přímá komunikace s knihovním systémem na straně TimeMap zatím není implementována. Je tedy nutné naprogramovat skript pro automatizovanou výměnu dat. Tímto způsobem je např. jednou denně proveden export dat z knihovního systému do podporovaného formátu.

Layer představuje vrstvu na prohlédavací mapě. Je vytvořena na základě DataStore. V prostředí TimeMap je v přednastavené verzi nalevo od mapy viditelný seznam vrstev, takže má uživatel možnost zvolit si aktuální vrstvu a ostatní vrstvy libovolně zapínat a vypínat.

Zprovoznění TimeMapu krok za krokem

Součástí prezentace byla i série sejmutých obrazovek, kdy byl předveden krok za krokem návod jak zpřístupnit pomocí zdarma dostupných nástrojů vlastní sbírku v TimeMapu.

Vlastní obrázky najdete na stránkách této prezentace (internetová adresa je pod článkem), sem nebyly z prostorových důvodů vloženy.

Nyní tedy stručný popis:

1) Vytvoření dat pro TimeMap v tabulkovém editoru (OpenOffice Calc, Microsoft Excel nebo jiné). Do prvního řádku je nutné zadat názvy jednotlivých sloupců, nejlépe bez diakritiky. Nezbytné je uvést nějakou jednoznačnou identifikaci každého záznamu (ideálně ID z knihovního systému), nějaký popis, časové určení, a zeměpisnou délku a šířku.

2) Získání souřadnic zeměpisné šířky a délky pro území ČR je možné například na stránkách služby Atlas.cz Mapy (adresa <http://mapy.atlas.cz>), kde nad pravým horním rohem zobrazené mapy je tlačítko s funkcí souřadnice. Po kliknutí na mapu je k dispozici údaj o zeměpisné šířce i délce, který je pro TimeMap nutné zkonvertovat do stupňů s desetinným místem. Konverzi podle jednoduchého vzorce můžeme provést sami v tabulkovém editoru.

3) Nebo využijeme některý z on-line nástrojů, které se dají snadno najít pomocí Google (např. zadáním „convert wgs84 degrees online“).

4) Soubor v tabulkovém editoru uložíme jako formát dBASE (DBF). Tím máme data pro TimeMap připravena k použití.

5) Nyní spustíme program TMWin (volně k dispozici na stránkách projektu TimeMap) abychom vytvořili Metadata naší sbírky. V TMWin se při startu automaticky otevře ukázková MapSpace. Stačí tedy stisknout tlačítko “Add Layer/Dataset” vlevo dole.

6) Dále budeme pokračovat přes “Local file Wizard”, tzn. průvodce, který nám tvorbu metadat pro naši sbírku usnadní.

7) Otevřeme vytvořený soubor DBF z disku

8) Zadáme pojmenování celé sbírky a případně i další informace. Ty ale můžeme doplnit i později. Po stisku tlačítka “Next” se dostáváme dál.

9) V tomto kroku určíme význam jednotlivých názvů sloupců, které jsme zadali v tabulkovém editoru při tvorbě dat. Poté dáváme stále tlačítka “Next”.

10) Otevře se Metadata Editor, ve kterém by již většina nutných záznamů měla být vyplněna, pomocí tlačítka “Verify Metadata” si ověříme správnost a úplnost zadaných hodnot. Pokud je nějaká hodnota vyplněna chybně, dojde k podsvícení daného pole červeně a zároveň je k dispozici text popisující, co se programu nelíbí. Uložit se podaří jen metadata, která jsou korektní. Mimochodem pro metadata je použit v TimeMapu standard Dublin Core.

11) Tím jsme se zadáváním dat hotovi. Pomocí tlačítka “Map”, vpravo dole, je k dispozici test hledání nad novými daty. Pomocí “Publish” je celý vyhledávací systém exportován včetně příslušných dat do formy vhodné pro vystavení na Internetu, eventuálně pro vypálení na CD-ROM.

12) Ukázka hledání v nově vytvořené sbírce.

Závěr

TimeMap nabízí zajímavé možnosti pro zpřístupnění historických fondů. Celý projekt je dostatečně otevřený a je dostupný zdarma a proto věřím, že se bude nadále rozvíjet a zlepšovat a že najde své uplatnění. Pořizujte tedy do bibliografického záznamu georeferenci! Je podporována jak v UNIMARC, tak MARC21. Zprovoznit prohledávání pomocí TimeMap pak není obtížné.

Odkazy

Projekt TimeMap (<http://www.timemap.net>)

Fondy MZK v TimeMapu (<http://timemap.mzk.cz>)

Stránky prezentace (<http://www.klokan.cz/presentations/2005timemap/>)

ECAI Clearinghouse (<http://www.ecai.org>)

David Rumsey Map Collection (<http://www.davidrumsey.com>)

Zoomify (<http://www.zoomify.com>)

Atlas mapy.cz (<http://mapy.atlas.cz>)