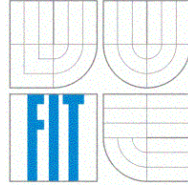


Vysoké učení technické v Brně
Fakulta informačních technologií



Ročníkový projekt

**Informační portál pro prezentaci genealogických
dat**

František Nosek

Brno 2006

Zadání:

1. Seznamte se s existujícími standardy pro ukládání genealogických dat, zaměřte se zejména na standard GEDcom.
2. Seznamte se s možnostmi dynamické prezentace dat na webových stránkách, zejména s technologií PHP.
3. Navrhněte prezentační portál, který bude vycházet z genealogického standardu pro ukládání dat, a vhodným a jednoduchým způsobem je bude prezentovat. Nejlépe pokud vstupem bude přímo datový soubor. Portál navrhněte tak, aby bylo možné používat více dat (stromů, rodů) současně.
4. Implementujte do portálu grafické procházení mezi jednotlivými generacemi a stromové zobrazení informací.
5. Pokuste se rozšířit portál o podporu nekomerčních databází (nejlépe kombinace PHP+mySQL) a možnost importu (resp. exportu) těchto do databází z běžně používaného standardu (GEDcom).

Čestné prohlášení

Prohlašuji, že jsem předkládaný ročníkový projekt vypracoval samostatně pod vedením Ing. Pavla Očenáška a že jsem uvedl veškeré zdroje informací, ze kterých jsem v průběhu své práce čerpal.

.....
podpis

Poděkování

Na tomto místě bych rád vyjádřil své poděkování Ing. Pavlu Očenáškovvi za jeho odborné vedení a vstřícnou pomoc při zpracování tohoto projektu.

Abstrakt

Projekt se zabývá studiem standardů pro ukládání genealogických dat, Nejvíce se věnuje standardu GEDCOM. Dále se zbývá implementací jednoduchého prohlížeče souborů formátu GEDCOM. Prohlížeč je prezentován jako internetová aplikace napsaná v interpretovaném jazyce PHP a zobrazená s pomocí kaskádových stylů CSS

Klíčová slova

genealogie, GEDCOM, GEDCOM XML, PHP, CSS, rodokmen, prohlížeč

Abstract

The purpose of this work is to study standards for storing genealogical data. Specially the GEDCOM standard. The goal of this work is to implement a simple viewer of GEDCOM files. The viewer is presented as internet application written in PHP interpreted language and cascading style sheets CSS.

Keywords

genealogy, GEDCOM, GEDCOM XML, PHP, CSS, genealogical tree, viewer

Obsah

1 Úvod.....	2
1.1 Genealogie a počítačové zpracování.....	2
1.2 Členění dokumentu	2
2 Ukládání genealogických dat.....	3
2.1 PAF.....	3
2.2 GEDCOM.....	3
3 GEDCOM 5.5 a GEDCOM XML 6.0	5
3.1 GEDCOM 5.5.....	5
3.2 GEDCOM XML 6.0	7
4 Dynamické HTML a styly CSS	12
4.1 PHP.....	12
4.2 CSS.....	12
5 Implementace GED prohlížeče	14
5.1 Datové struktury	14
5.2 informací z GED souborů.....	15
5.3 Výpis seznamu osob.....	17
5.4 Vykreslování rodokmenu.....	17
5.5 Ukázka souboru CSS s informacemi pro zobrazení.....	18
6 Závěr	19
6.1 Zhodnocení výsledků.....	19
6.2 Další možné rozšíření projektu	19
Literatura.....	20
Příloha A – Datové modely a diagramy	21
Příloha B - Ukázky prohlížeče.....	23

1 Úvod

1.1 Genealogie a počítačové zpracování

Genealogie neboli rodopis je vědou, která se zabývá rody a rodinami. V minulosti se touto vědou zabývaly hlavně šlechtické rody. Na mnoha zámcích a hradech jsou i dnes k vidění rodokmeny šlechtických rodů zpracované velmi umělecky významnými malíři vystavené jako obrazy. I dnes se touto vědou zabývá velké množství lidí. V poslední době se také začínají digitalizovat záznamy matrik, což přivádí ke genealogii další zájemce z řad amatérů. Ještě dnes je však nutné většinu záznamů hledat po matrikách a zájemce o informace je nucen cestovat do různých míst. Proto se také s počítači objevily i programy pro zpracování rodokmenů a s nimi i různé standardy pro ukládání genealogických dat.

1.2 Členění dokumentu

V druhé kapitole je přehled nejvíce užívaných formátů pro ukládání dat pro genealogii. Kapitola třetí se pak zabývá nejrozšířenějším formátem podrobně. Ve čtvrté kapitole se seznámíme s interpretovaným programovacím jazykem PHP a další kapitoly jsou věnovány vytvořenému programu.

2 Ukládání genealogických dat

Jak již bylo zmíněno v úvodu, existuje dnes mnoho softwaru pro práci s genealogickými daty. Amatérský genealog vystačí i se staršími programy, které mnohdy ukládají informace do databáze Betrievé. Případně využívají interní formáty souborů. Jedním z nich je i formát PAF, jeden z nejpoužívanějších formátů, bude zmíněn dále v této kapitole. Nejvíce se na rozvoji programů pro genealogy podílela mormonská církev. V církvi samotné existuje skupina pro historii rodin a chrámů (Family and Church History Department), která se spravuje jedny z nejobsáhlejších databází s informacemi o rodech. Protože tato skupina potřebovala získávat data z různých zdrojů, vytvořila nový standard GEDCOM¹, jehož import a export je podporován většinou genealogických programů. Pro velké databáze je však formát nevhodný. Mormonská církev jistě pro své rozsáhlé databáze využívá výkonných databázových serverů a data jsou přímo v databázích. Stejně tak jsou uloženy i jiná větší množství genealogických dat v různých informačních systémech firem, které se zabývají digitalizací záznamů z matrik a jiných úřadů.

2.1 PAF

Tento formát je rozšířený, protože jej používá program Personal Ancestral File, který je v základní verzi volně šiřitelný, ale existuje i placená verze poskytující další výhody. Poslední dostupná verze tohoto programu poskytuje i možnost uložit data do přenosných počítačů (Palm/PDA), což je velmi vhodné při cestování při získávání nových informací o předcích. K tomuto programu vzniklo i velké množství různých nástrojů, které umožňují tisk rodokmenu apod. Existují i některé nástroje pro kontrolu správnosti tohoto formátu případně i pro čtení a konverzi dat z tohoto formátu, ale pro ukládání dat se používá téměř jen program Personal Ancestral File.

2.2 GEDCOM

Formát GEDCOM¹ byl vyvinutý skupinou pro historii rodin a kostelů mormonské církve. Tento formát se hojně využívá pro přenos informací mezi různými genealogickými programy. Je i velmi dobře strukturovaný. Ovšem je na druhou stranu i velmi dobře čitelný. Data jsou zde uložena textově, ale zároveň je i zachována hierarchie. Standard i přesně definuje gramatiku souboru [3], což velmi ocení hlavně vývojáři, kterým poskytuje možnosti dobře přečíst či zapsat data. První dvě verze byly pouze návrhy, které byly uveřejněny aby o nich mohlo být diskutováno a bylo možné vznést připomínky. Zatím poslední verze formátu je z roku 1999. Ovšem rozdíl je oproti starším verzím minimální. Jedná se hlavně o upřesnění některých pravidel v gramatice souboru případně o přidání či modifikaci některých tagů, jako jsou informace o absolvování školy apod. Hlavní záznamy o osobách, rodinách a událostech, které jsou využity nejvíce se téměř nezměnily

¹ GEDCOM = GEnealogical Data COMunication – formát pro výměnu genealogických dat

od starších návrhů standardu. Poslední aktuální verze standardu 5.5 je pravděpodobně jednou z posledních modifikací, protože se již v roce 2002 byla vydána betaverze návrhu standardu GEDCOM XML 6.0. I tento standard vzniká v prostředí mormonské církve a jak již napovídá název, snaží se autoři udělat z původního čistě genealogického formátu dokument XML². Změna formátu není jednoduchou záležitostí a vyžaduje dlouhý čas hlavně pro přípravu vývojářů, kteří se musí naučit zapisovat nový formát souboru. V současné době (5/2006) je návrh stále ještě ve stadiu betaverze a vývojáři genealogických programů zasílají připomínky a zároveň se odlaďují vzniklé chyby.

Přechodem pod jazyk XML se však otevírají možnosti využití všech výhod tohoto jazyka. Nové možnosti, které se objevují s novým formátem jsou hlavně z oblasti importu dat ze souboru do databází. Pro import staršího formátu souboru bylo nutné psát programy nebo skripty, které import prováděly a řídily se zpracováním tagů GEDCOM souborů. Naproti tomu u souboru, který je zároveň XML dokumentem lze využít podpory XML ze strany databázových serverů nebo použít některé z volně šiřitelných nástrojů. Přechod na nový formát s sebou nese i změny ve struktuře dokumentu, kde byly tagy nahrazeny elementy jazyka XML. Některé tagy nebylo možné nahradit přesně, také se někde změnila hierarchie. Změny nejdůležitějších tagů a další rozdíly mezi oběma formáty jsou popsány v následující kapitole.

2 XML = Extensible Markup Language – rozšiřitelný značkovací jazyk konsorcia W3C – www.w3c.org

3 GEDCOM 5.5 a GEDCOM XML 6.0

Nově připravovaná verze GEDCOM XML 6.0 již musí samozřejmě dodržovat syntaktická pravidla jazyka XML, proto je zápis informací zcela odlišný, hierarchie již není určena číslem na začátku řádku, ale pomocí elementů jazyka XML. Pro vývojáře, kteří budou jistě postupem času nuceni do svých programů zahrnout i export a import z a do nového formátu, je však přechod na nový formát velkou výhodou.

3.1 GEDCOM 5.5

Ve formátu GEDCOM lze teoreticky ukládat jakékoliv záznamy, v nichž jsou obsaženy vztahy, ale nejvíce je využíván pro přenos záznamů mezi genealogickými programy. Všechny záznamy jsou uspořádány do úrovní, na nejvyšší úrovni jsou záznamy o rodinách, lidech i o autorovi dokumentu. Také jsou na této úrovni řídicí znaky začátku a konce souboru. Další informace související s danou rodinou jsou uvedeny hned pod tagem určujícím číslo rodiny. Všechny tyto záznamy však již mají úroveň jedna nebo vyšší. Například informace o svatbě v sobě ještě zahrnuje datum případně místo svatby. Informace o místě a datu svatby je již označena číslem dvě a je tedy ve třetí úrovni. Kde končí daná informace o rodině lze zjistit z čísla úrovně zanoření. Obdobně jsou uloženy i informace o osobách. Čísla u záznamů o osobách i u záznamů o rodinách samozřejmě musí být v unikátní v rámci jednoho souboru. Tato čísla záznamů zde mají funkci jako primární klíče v databázi a jsou použita k vytváření zpravidla oboustranných vazeb mezi záznamy o rodině a záznamy o osobách. Dále budou popsány nejdůležitější typy záznamů. Přesná gramatika formátu GEDCOM a popis dalších typů záznamů viz [3].

ADDR – informace o adrese, může být použit u záznamů osob, firem škol apod

ADOP – tento typ záznamu vytváří vztah dítě rodič, jedná se o záznam o adopci dítěte.

BAPM, BAPL, BARM, BASM - tento typ slouží pro uložení informací o náboženských obřadech jako je křest aj.

BIRT – záznam o narození dítěte

BURI – záznam o pohřbu

CHIL – slouží pro informaci o narozených a adoptovaných dětech

CONT – pouze pomocný typ který říká že se jedná o pokračování informace z předchozího řádku

DATE – informace o datu vztahující se k nějaké události, např BIRT, DEAT, BURI kdekoliv je třeba přiřadit informaci o datu, je použit tento identifikátor záznamu

DEAT – informace o smrti osoby, nejčastěji spojena s datem případně místem.

DIV, DIVF – pro uložení informací o rozvodu manželů

EMAI – záznam o adrese elektronické pošty (email)

EMIG – informace o emigraci osoby

FAM – záznam o rodině, jedná se nejdůležitější záznam, který obsahuje odkazy definující pozici člena rodiny. Typ záznamu udává pozici kterou člověk v rodině zastává a následuje číslo ID dané osoby

FAMC – součást záznamu o osobě (INDI), v tomto případě se jedná o odkaz na informace o rodině ve které daná osoba vystupuje v pozici dítěte

FAMS – informace o rodině ve které daná osoba vystupuje v roli manžela či manželky

IMMI – záznam o imigraci

INDI – informace o konkrétní osobě, což je základní záznam celého souboru, zde se uchovávají informace o každém člověku uloženém v souboru. Součástí záznamu jsou odkazy na záznamy typu FAM, odkazy podle role ve kterých daná osoba v rodině vystupuje a číslo rodiny

MARR – záznam typu událost - svatba

NAME – hlavní záznam pro identifikaci jedince, ukládá se celé jméno člověka

SEX – pohlaví jedince – záznam určující pohlaví – část záznamu INDI

SOUR – zdroj dat pro GED soubor

SUBM – poskytovatel genealogických dat – firma či osoba které poskytla data v GED souboru

TRLR – ukončení souboru GED pokud je tento tag na úrovni 0

Největší problémy přinášelo zpracování tagu ADOP a odpovídajících vztahů. Ve standardu verze 4.x jsou vazby mezi adoptivními rodiči a dětmi uvedeny v rámci tagu ADOP, který je uveden jako událost, kdežto od verze 5 je možné je vztah modelován na stejné úrovni jako u ostatních záznamů o dětech a jejich rodičích. Obdobně jsou také změněny vztahy mezi dětmi osvojenými v podle pravidel mormonské církve.

V GEDCOM standardu je samozřejmě definováno ještě mnoho dalších typů záznamů. Pro data o osobách je zde možné uvést internetovou adresu domovských stránek, zemi ve které žije, získané vzdělání aj. Tyto záznamy se však mnoho nevyužívají a pokud ano, tak nebývají nikde veřejně vystaveny. Formát GEDCOM se využívá hlavně pro výměnu informací mezi různými programy a proto tato soukromá data nebývají ani exportována. Dále jsou uvedeny ukázky z GED souboru.

Ukázka záznamu o osobě:

0 @I51@ INDI – hlavní záznam, jedná se o člověka a bude odkazován přes číslo I51
1 NAME George /McBride/ - jméno v závorce jméno rodu (příjmení)
1 SEX M – pohlaví - Muž
1 BIRT – záznam o narození
2 DATE 11 AUG 1907 – informace o datu vztahující se k narození
2 PLAC Stenson, Texas – informace o místě, taktéž pro narození v tomto případě
1 DEAT – záznam o úmrtí
2 DATE 3 MAY 1989 – opět informace o datu
2 PLAC San Diego, San Diego Co., California – informace o místě úmrtí
1 BURI – informace o pohřbu
2 DATE 12 MAY 1989 – datum kdy byl pohřben
2 PLAC Rosecrans Natl C, Calif. - místo kde byl pohřben
1 FAMS @F14@ - odkaz na záznam o rodině ve které vystupuje v roli manžela
1 FAMC @F15@ - odkaz na záznam o rodině ve které vystupuje v roli dítěte

Ukázka záznamu o rodině:

0 @F202@ FAM – číslo rodiny
1 HUSB @I624@ - odkaz na manžela
1 WIFE @I154@ - odkaz na manželku
1 CHIL @I625@ - odkaz na dítě
1 CHIL @I626@ - odkaz na druhé dítě
1 DIV Y – informace o rozvodu
1 MARR – informace o svatbě
2 DATE 1948 - datum svatby

3.2 GEDCOM XML 6.0

Specifikace formátu GEDCOM XML ještě není hotova, jedná se o zatím pouze o betaverzi. (Stav v květnu 2006). Nicméně se již nepředpokládají výrazné změny ve struktuře souboru. Pro přizpůsobení formátu GEDCOM tak aby byl aplikací XML bylo potřeba změnit syntaxi a symboly používané ve formátu GEDCOM podle struktury jazyka XML. Z hlediska možností výměny informací jsou oba formáty podobné, oba zachycují vztahy záznamů a umožňují strukturovaný zápis dat, ale jazyk XML je navržen lépe s ohledem na další možnosti použití. Velkým rozdílem který je nutné překonat je určování hloubky, ve formátu GEDCOM je hierarchie dána číslem zanoření před každým řádkem, kdežto v XML je každý tag ukončen. Velkou výhodou formátu GEDCOM XML je možnost zpracování dat všemi dostupnými prostředky vyvinutými pro jazyk XML, existuje velké množství nástrojů pro práci s XML, které je možno z velké části využít pro práci s genealogickými daty obdobně jako jsou dnes využívány v prostředí internetu. Nevýhodou tohoto formátu je že již není zpětně kompatibilní s formátem GEDCOM, který existoval po patnáct

let. Velkou změnou oproti verzi 5.5 je odstranění redundance vztahů. V klasickém formátu GEDCOM může být vztah zaznamenán na dvou místech, v záznamu FAM může být položka CHIL, která ukazuje na záznam o osobě, naproti tomu v záznamu o osobě lze odkázat na rodinu pomocí záznamu FAMC. Při existenci obou záznamů mohou vznikat nesrovnalosti a proto je ve formátu GEDCOM XML umožněna specifikace vztahu pouze na jednom místě a jedním směrem. Tagy formátu 5.5 jsou v návrhu GEDCOM XML 6.0 změněny na elementy jazyka XML. Byla velmi výrazně změněna struktura záznamů, porovnání datových modelů a všech vztahů je nejlépe vidět na datových modelech obou formátů, které jsou uvedeny v příloze A. V následujícím seznamu jsou uvedeny některé důležité elementy a atributy standardu GEDCOM XML 6.0. Přehled všech atributů a elementů lze nalézt v [4].

CHILD – informace o dítěti, v elementu Family, obsahuje odkaz na dítě

DATE – datum, nejčastěji u elementu událost, je zaznamenán ve formátu den, měsíc, rok, ale jsou zde i další možnosti jak upřesnit informaci i když datum není přesně známo, slouží k tomu různé tagy jako APT, které říká že se jedná o přibližný údaj, nebo BEF kde událost se stala před uvedeným datem případně i BET kde je uvedeno rozmezí. U elementu DATE lze také uvést atribut Calendar, který umožňuje hodnoty: Gregorian, Julian, Hebrew, French, Roman, unknown.

EventRec – v tomto elementu se zaznamenávají veškeré události, jsou rozlišeny typem, typů je zde celá řada, protože tento element sdružuje všechny události dříve uvedené pod různými tagy, jako narození, smrt, svatba, pohřeb atd.

FamilyRec – obsahuje záznam o rodině, včetně odkazů na členy rodiny a události týkající se rodiny

Gender – informace o pohlaví, jedná se o atribut v záznamu o osobě, možné hodnoty jsou Male, Female a Unknown

HusbFath – obsahuje záznam o otci v rodině s odkazem na záznamy o dané osobě

IndividualRec – záznam o osobě

IndivName – tento element udává jméno tak jak je například překládáno do jiných jazyků, případně umožňuje zaznamenat například přidávání koncovky -ová v češtině apod. Také lze uvést přezdívkou, rodná jména apod.

Role – identifikuje účastníka nějaké události – například ženich na svatbě, případně otec, matka při narození dítěte

Submitter – odkaz na osobu které dodala informace obsažené v souboru

WifeMoth – odkaz na matku, manželku v rodině

Ukázka záznamu o osobách a událostech:

```
<IndividualRec Id="I0001">
<IndivName>John Smith</IndivName>
</IndividualRec>
<IndividualRec Id="I0002">
<IndivName>Jane Doe</IndivName>
</IndividualRec>
<EventRec Id="E0001" Type="marriage">
<Participant>
<Link Target="IndividualRec" Ref="I0001"/>
<Role>husband</Role>
</Participant>
<Participant>
<Link Target="IndividualRec" Ref="I0002"/>
<Role>wife</Role>
</Participant>
<Date>7 NOV 1891</Date>
<Place>Cove, Cache, Utah, USA</Place>
</EventRec>
```

Ukázka využití možností XML v informacích o osobách:

Duchess Neta Eskelson von Allen

```
<IndivName>
  <NamePart Type="title">Duchess </NamePart>
  <NamePart Type="given name" Level="3"> Neta </NamePart>
  <NamePart Type="maiden name" Level="2"> Eskelson </NamePart>
  von <NamePart Type="surname" Level="1"> Allen</NamePart>
</IndivName>
```

Takto lze zaznamenat všechny součásti jména, při zpracování lze pak vybrat pouze podstatné části, ale lze informace o jméně je kompletní. Tyto změny v zápisu informací se nejvíce projevily u záznamů o osobách. Odkazuje se i podle typu elementu jak je vidět na ukázce. GEDCOM XML uchovává obsahově stejná data, ale jsou uchovány v lépe organizované struktuře a také obsahují méně nejednoznačností. Některé téměř nevyužívané tagy byly vypuštěny. Dalšími vlastnostmi, které je nutné zapsat jiným způsobem jsou uživatelsky definované tagy, které je v novém formátu potřeba přiřadit do jiného jmenného prostoru než je GEDCOM XML a pak je možné je použít. V XML je také preferováno kódování znaků UNICODE oproti kódování ANSEL ve verzi 5.5. V tabulce 3.2.1 na následující stránce jsou uvedeny elementy XML tak jak odpovídají tagům ze standardu verze 5.5.

Tag ve formátu GEDCOM 5.5	Ekvivalent ve formátu GEDCOM XML 6.0
AGE	EventRec → Participant → Age
AUTH	SourceRec → Author
BAPL	LDSOrdRec → Type="B"
BAPM	EventRec → Type="baptism"
BARM	EventRec → Type="bar mitzvah"
BASM	EventRec → Type="bas mitzvah"
BIRT	EventRec → Type="birth" VitalType="birth"
BLES	EventRec → Type="blessing"
BURI	EventRec → Type="burial" VitalType="death"
CHIL	FamilyRec → Child
CHR	EventRec → Type="christening" VitalType="birth"
CHRA	EventRec → Type="adult christening"
CITY	EventRec → Place → PlacePart → Type="city"
	ContactRec → MailAddress → AddrLine → PlacePart → Type="city"
CREM	EventRec → Type="cremation" VitalType = "death"
DATE	EventRec → Date
	LDSOrdRec → Date
	HeaderRec → FileCreation → Date
	LDSOrdRec → OrdStat → Date
	xxxxxRec → Changed → Date
DEAT	EventRec → Type="death" VitalType="death"
DIV	EventRec → Type="divorce"
DIVF	EventRec → Type="filed for divorce"
EDUC	IndividualRec → PersInfo → Type="education"
EMIG	EventRec → Type="emigration"
FAM	FamilyRec
FAMC	FamilyRec → Child
FAMS	FamilyRec → HusbFath
	FamilyRec → WifeMoth
HUSB	FamilyRec → HusbFath
INDI	IndividualRec
LEGA	EventRec → Participant → Role

NAME	IndividualRec → IndivName
	HeaderRec → FileCreation → Product → Name
	ContactRec → Name
	RepositoryRec → Name
	GroupRec → Name
NICK	IndividualRec → IndivName → Type="nickname"
PLAC	EventRec → Place
	LDSOrdRec → Place
RETI	EventRec → Type="retirement"
ROLE	EventRec → Participant → Role
SEX	IndividualRec → Gender
SURN	IndividualRec → IndivName → NamePart → Type="surname"
TRLR	vypuštěno
WIFE	FamilyRec → WifeMoth
WILL	EventRec → Type="will" VitalType="death"

Tabulka 3.2.1: Tagy verze 5.5 a jim odpovídající elementy verze XML 6.0

Jak je vidět z tabulky, jsou tagy z verze 5.5 nahrazeny několika základními elementy, které jsou doplněny o další atributy určující přesný význam. Zápis dat ve formátu GEDCOM XML je také poněkud rozsáhlejší, ale při běžném použití pro přenos informací mezi systémy a programy se nejedná o velký nárůst velikosti souboru.

4 Dynamické HTML a styly CSS

Pro dynamické internetové stránky se velmi často používá jazyk PHP³. PHP³ poskytuje mnoho funkcí pro získávání dat z databází, ale i podporu pro čtení ze souboru a práce s řetězci. Všechny takto zpracované údaje pak lze HTML⁴ stránky jednoduše vypisovat. PHP je proto i velmi vhodným nástrojem pro zpracování prohlížeče souborů formátu GEDCOM. Další důležitou součástí HTML stránek je jejich vzhled. Pro definici barev, písma a dalších vlastností lze velmi vhodně použít kaskádové styly CSS⁵.

4.1 PHP

PHP je interpretovaný jazyk používaný zejména na vytváření dynamických HTML stránek. Jedná se o produkt šířený zdarma (Open Source). Tento jazyk vytvořil v roce 1994 Rasmus Lerdorf pro svoji potřebu. V dnešní době je nejrozšířenější verze PHP 4, ale verze 5 velmi podstatně zlepšuje možnosti objektově orientovaného programování. Velkou výhodou tohoto jazyka je množství knihoven, které umožňují práci s mnoha internetovými protokoly (HTTP, FTP, SMTP, IMAP aj.). Avšak jsou zde i knihovny pro grafiku, zpracování textu a hlavně knihovny pro podporu přístupu k mnoha databázovým serverům (MySQL, Oracle, PostgreSQL aj.). Do stránek HTML se je možno vkládat kód PHP několika způsoby. První možností je vložení kódu mezi značky `<? , ?>`. Při použití PHP pro generování XML je nutné přidat identifikátor preprocesoru pomocí značek `<?php` a `?>`. Poslední možnosti, která zaručuje funkčnost pro některé starší editory HTML je pomocí tagů `<SCRIPT LANGUAGE="php">` a `</SCRIPT>`.

4.2 CSS

Tabulky kaskádových stylů jsou určeny pro popis prezentace dokumentů a jsou nadstavbami značkovacích jazyků HTML, XHTML a XML. V dnešních dynamických stránkách a velkém množství informací v HTML stránkách, se pozice některých prvků řeší pomocí několikanásobně vnořených tabulek a HTML dokument se stává velmi nečitelným. Z velké části obsahuje pouze popis jak zobrazovat obsažené informace. Velkou výhodou, které kaskádové styly přinášejí je oddělení definice vzhledu stránky od obsahu. Každý objekt na stránce může mít nastaveny jiné vlastnosti pro zobrazování, které je možné načítat z externích souborů a lze je také vybrat podle druhu zobrazovacího zařízení. Pro stejné objekty na stránce lze definovat různé styly pro tisk stránky pro zobrazování na monitoru počítače případně i pro přenosné počítače. Všechny objekty, ke kterým chceme přiřadit nějaké vlastnosti pomocí CSS stylů musí být pojmenovány, nebo musí být objektem nějaké třídy. Styl se poté přiřadí objektům dané třídy případně přímo pojmenovanému objektu. Velkou výhodou CSS je možnost dědění, kde lze již existující styl rozšířit o některé

3 PHP – PHP hypertext Processor – interpretovaný jazyk pro vytváření dynamických HTML stránek

4 HTML – Hypertext Markup Language – značkovací jazyk

5 CSS – Cascading Style Sheets – kaskádové styly

vlastnosti. Tímto způsobem lze velmi dobře oddělit různé typy objektů na stránce. Můžeme definovat základní styl například pro všechny tabulky a některé výjimky, u kterých chceme vzhled změnit pouze rozšířit základní vlastnosti. Případně můžeme od nějakého základního stylu odvodit několik dalších odlišných stylů se společným základem. V dnešní době jsou styly CSS široce používány pro formátování a úpravu vzhledu celých informačních serverů. Mnohé internetové servery nabízejí i výběr z několika možností zobrazení, případně umožňují pravidelným návštěvníkům vybrat si z několika možností. Velmi často lze také změnit vzhled celého internetového serveru při změně ročního období případně změnit například vzhled internetového obchodu před vánoci. Kaskádové styly jsou programátorům a webdesignerům v tomto úkolu velkým pomocníkem.

5 Implementace GED prohlížeče

Program pro přehlednou prezentaci dat z formátu GED má několik hlavních částí, které postupně popíšeme v této kapitole. Nejdůležitější částí programu je část načítání informací z ged souborů a jejich uložení do datových struktur tak, aby bylo snadné poté vykreslit vztahy mezi osobami.

5.1 Datové struktury

Pro snadnou práci je velmi vhodné si data získaná ze souboru načíst do paměti. Hlavní informace, které bylo potřeba uložit byly o osobách a o rodinách. Pro každý z těchto dvou druhů záznamů jsem vytvořil dvourozměrné pole. Každý řádek představuje jeden záznam o jedné osobě, případně o jedné rodině. Ve sloupcích jsou poté uloženy detailní informace, jež dané záznamy obsahují. Pozice jednotlivých údajů v polích jsou zobrazeny v tabulce 5.1.1 a 5.1.2. Kde obsah tabulky odpovídá jednomu záznamu o osobě případně o rodině.

0	Jméno
1	Příjmení
2	Pohlaví
3	Narození - datum
4	Narození - místo
5	Smrt - datum
6	Smrt - místo
7	Křest (BAP..) - datum
8	Křest (BAP..) - místo
9	Počet dětí
10	Počet rodičů (uvedených v souboru)
11	Odkazy na děti a rodiče podle počtů
12	Odkazy na děti a rodiče podle počtů

Tabulka 5.1.1. Popis pole se záznamy o osobách

0	Manžel (otec)
1	Manželka (matka)
2	Svatba - datum
3	Svatba - místo
4	Rozvod - datum
5	Rozvod - místo
6	Odkazy na děti
7	Odkazy na děti

Tabulka 5.1.2. Popis pole se záznamy o rodinách

Znázorněny jsou nejdůležitější struktury, při načítání a poté i zobrazování jsou použity ještě další pomocná pole.

5.2 informací z GED souborů

Získávání informací z GED souboru je vcelku snadnou záležitostí, protože PHP obsahuje mnoho funkcí pro práci jak se soubory tak i s řetězci. Soubor se načítá postupně po řádcích. Načítá se v cyklu po řádcích až do konce souboru nebo do ukončovacího tagu TRLR. Z načteného řádku se odstraní mezery na jeho konci a poté se pomocí funkce `explode` řádek rozdělí podle oddělovačů na pole řetězců. Další práce s polem řetězců je pak řízena podle čísel na začátku řádků a podle tagu.

```
while (!feof($file) && !$konec) {
    $line=chop(fgets($file, 100));
    $line=explode(" ", $line);
    $lines++;
#   rozhodování podle prvního slova (číslo úrovně) na řádce o dalším postupu zpracování
    switch($line[0]) {
```

Při nalezení čísla 0 se načítají buď informace o rodině (FAM) , osobě (INDI) , autorovi dokumentu nebo jde o tag TRLR, který značí konec souboru. Pro záznamy o rodinách či osobách se do pole uloží číslo záznamu a vynulují se počty dětí a rodičů, obdobně u záznamů o rodinách.

```
case 0:
    $objekt="";
    $podobjekt="";
    $jiny=false;
    if (isset($line[2])) $temp=$line[2];
    else $temp="";
    switch($temp) { //deli se podle tagu Fam nebo INDI
    case "INDI":
        $objekt="I";
        $cislo=SubStr($line[1], 2, -1); // ulozi cislo zaznamu
        $I[$cislo][0]="";
        $I[$cislo][9]=0;
        $I[$cislo][10]=0;
        $Is[$j]=$cislo;
        $j++;
        break;
    case "FAM":
        $objekt="F";
        $cislo=SubStr($line[1], 2, -1);
        $F[$cislo][6]=0;
        break;
    default: $jiny=true;
    }
}
```

Proměnné objekt a podobjekt určují k jakému záznamu se vztahují informace na další úrovni. Při zjištění tagu FAM se jako objekt označí záznam o rodině. Do podobjektu se poté uloží informace k jaké události se vztahuje následující tag, může jít o svatbu apod. v záznamech o rodině, nebo o narození, smrt aj. v záznamech o osobách.

```

case "TITL": // zaznamy uovne 1
    $I[$scislo][0]='<i>'. $line[2]. '</i> '.$I[$scislo][0];
    break;
case "SEX":
    $I[$scislo][2]=$line[2];
    break;
case "BIRT":
    $podobjekt="BIRT";
    $I[$scislo][3]="";
    $I[$scislo][4]="";
    break;
case "DEAT":
    $podobjekt="DEAT";
    $I[$scislo][5]="";
    $I[$scislo][6]="";
    break;
// zaznamy urovne 2
if ($objekt=="I") {
    if ($podobjekt=="BIRT") {
        switch($line[1]) {
            case "DATE":
                $i=2;
                while (isSet($line[$i])) {
                    $I[$scislo][3].=$line[$i]." ";
                    $i++;
                }
                break;
            case "PLAC":
                $i=2;
                while (isSet($line[$i])) {
                    $I[$scislo][4].=$line[$i]." ";
                    $i++;
                }
                break;
        }
    }
}

```

5.3 Výpis seznamu osob

Výpis osob obsažených v souboru již není problém, projde se celé pole záznamů a postupně se vypíše do tabulky. Tabulka má definovanou třídu pro úpravu zobrazení pomocí CSS.

```
# výpis seřazeného seznamu osob
$i=0;
while (isset($Is[$i])) {
    echo '<tr id="I'.$Is[$i].'" OnMouseOver="je(this);" onMouseOut="ne(this);">
    <td align="center" class="zaz">'.$Is[$i].'</td>
    <td align="center" class="zaz"><a href="?zdroj='.$_GET['zdroj'].'&indi='.$Is[$i].'">;
    if (!mezery($Is[$i][0])) echo $Is[$i][0];
    else echo '<i>&lt;neuvedeno&gt;</i>'.$Is[$i][0];
    echo '</a></td><td align="center" class="zaz">;
    if (isset($Is[$i][1]) && $Is[$i][1]!="") echo '/'.$Is[$i][1].'/';
    else echo '&nbsp;';
    echo '</td><td align="center" class="zaz">&nbsp;';
    if (isset($Is[$i][2])) echo $Is[$i][2];
    echo '</td><td align="center" class="zaz">&nbsp;';
    if (isset($Is[$i][3])) echo $Is[$i][3];
    echo '&nbsp;</td></tr>'. "\n";
    $i++;
}
```

5.4 Vykreslování rodokmenu

Ze seznamu jmen se předá pomocí HTTP číslo osoby pro kterou se má vykreslit strom. Poté se vytvoří hlavní tabulka. Do té se postupně budou přidávat další vnořené tabulky. Prvním řádkem hlavní tabulky je vnořená tabulka obsahující jména rodičů osoby pro kterou kreslíme strom. Další řádek tabulky obsahuje opět tabulky a v ní informace o osobě pro kterou kreslíme rodokmen. Dalším řádkem hlavní tabulky je informace o partnerovi, v případě více dalších partnerů se objeví v dalších buňkách tabulky. Do spodní části této tabulky je vložena ještě jedna, která zobrazuje informace o dětech. Na zobrazení detailních informací jsou vytvořeny další tabulky. Kód vykreslování rodokmenu je možné prohlédnout v příloze na CD-ROM. Vykreslování je prováděno podle získávání správných záznamů z polí. Nalezení správného indexu je nutné řešit pomocí více odkazů přes jiné záznamy a je velmi rozsáhlé, proto zde také není vloženo.

```
 $\$I[\$F[\$_GET['indi']][\$i][\$ii][0]$  – takto je indexován vykreslovaný záznam o dítěti  
     $\$_GET['indi']$  – ID osoby, pro kterou zrovna vypisujeme strom  
     $\$I[ ID osoby ][\$i]$  –  $\$i=11,12,\dots$  – získáme ID rodiny, ve které je osoba rodičem  
     $\$F[ ID rodiny ][\$ii]$  –  $\$ii=7,8,\dots$  – získáme ID dítěte  
     $\$I[ ID dítěte ][0]$  – získáme jméno dítěte
```

5.5 Ukázka souboru CSS s informacemi pro zobrazení

Popis vzhledu a vlastností tabulky se seznamem osob.

```
table.vypis {  
    border-top : 2px solid black;  
    border-left : 2px solid black;  
    border-right : 1px solid black;  
    border-bottom : 2px solid black;  
}
```

Popis vzhledu tabulky s výpisem manžela

```
td.I_on {  
    border-top : 2px solid black;  
    border-left : 2px solid black;  
    border-right : 1px solid black;  
    border-bottom : 2px solid black;  
    padding: 2px 4px 2px 4px;  
}
```

Popis vzhledu tabulky s výpisem manželky

```
td.I_ona {  
    border-top : 2px solid black;  
    border-left : 1px solid black;  
    border-right : 2px solid black;  
    border-bottom : 2px solid black;  
    padding: 2px 4px 2px 4px;  
}
```


6 Závěr

6.1 Zhodnocení výsledků

Standard GEDCOM je primárně určen pro přenos informací mezi ostatními programy pro zpracování genealogických dat, ale i mezi těmito programy a velkými databázemi a opačně. V dnešní době je však možné data mezi databázemi přenášet i ve formátu XML. Formáty GEDCOM a XML jsou si v některých částech velmi podobné i když je formát XML přesněji definován. Velký rozvoj jazyka XML pak také přispěl k vývoji nového standardu GEDCOM XML 6.0, jehož specifikace však v dnešní době (5/2006) ještě není konečná. Jedná se o betaverzi návrhu, ke které se mohou vyjádřit vývojáři případně i jiní odborníci.

Implementovaný program je použitelný zejména pro rychlé přehledné zobrazení jednoduchých GED souborů. Umožňuje intuitivní procházení záznamů, každé jméno v seznamu je odkazem, po jehož otevření se zobrazí strom pro vybrané osoby. Program počítá s možností více partnerů pro vybranou osobu. Ukázky vykreslování stromu jsou v příloze B.

6.2 Další možné rozšíření projektu

Bylo samozřejmě možné v projektu pokračovat a doplnit jej o řadu dalších funkcí včetně zápisu nových informací do souboru případně i zobrazování některých dalších informací obsažených v souborech. Ovšem pro základní informaci o vztazích je prohlížení na této úrovni dostatečné, pro lepší práci s GED soubory již dnes existuje mnoho jiných programů, případně lze formát GED naimportovat přímo do některého z volně šiřitelných programů. Pokud jde o grafickou prezentaci GED souboru, existuje již dnes systém, který je možné přímo nasadit na internetový server a provozovat jej. Proto si myslím že další vývoj tohoto programu není nutný a lepší možností by byla případná spolupráce na projektu PhpGedView. Tento projekt řeší všechny možné operace s GED soubory, umožňuje editaci a je dnes již plnohodnotným informačním systémem s překlady do mnoha jazyků. Na vývoji projektu PhpGedView se podílí řada lidí a stále se přidává podpora nových funkcí. Proto si myslím že mnou implementovaný program bude sloužit spíše pro uživatele, kteří si chtějí přehledně zobrazit strom a informace v GED souboru nechtějí mnoho měnit. Pro velký informační server pro amatérské genealogy, kde by mohli své soubory editovat, a ještě lépe s nimi pracovat je nejvhodnějším řešením nasazení již zmiňovaného programu PhpGedView.

Literatura

- [1] KOSEK J. PHP tvorba interaktivních internetových aplikací podrobný průvodce, Grada Publishing spol. s.r.o, 1999, ISBN 80-247-0441-2
- [2] Staniček P. CSS kaskádové styly kompletní průvodce, Computer Press a.s, 2003, ISBN 80-7226-872-4
- [3] Family and Church History Department - The Church of Jesus Christ of Latter-day Saints, The GEDCOM standard Release 5.5, dokument dostupný na URL:
<http://www.familysearch.org/GEDCOM/GEDCOM55.EXE> (Envoy Format)
- [4] Family and Church History Department - The Church of Jesus Christ of Latter-day Saints, The GEDCOM XML specification Release 6.0, dokument dostupný na URL:
<http://www.familysearch.org/GEDCOM/GedXML60.pdf>

Příloha A – Datové modely a diagramy

Datový model GEDCOM 5.5

4 Dec 1995
Prepared by:
Robert Booth

GEDCOM 5.5 DATA MODEL CHART

Page 1 of 2

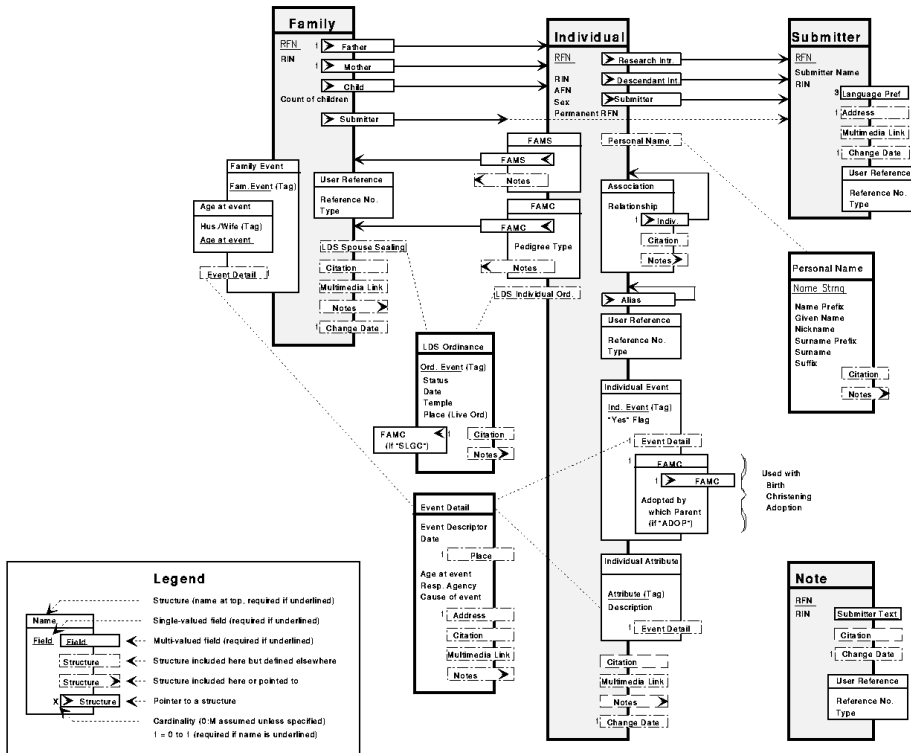
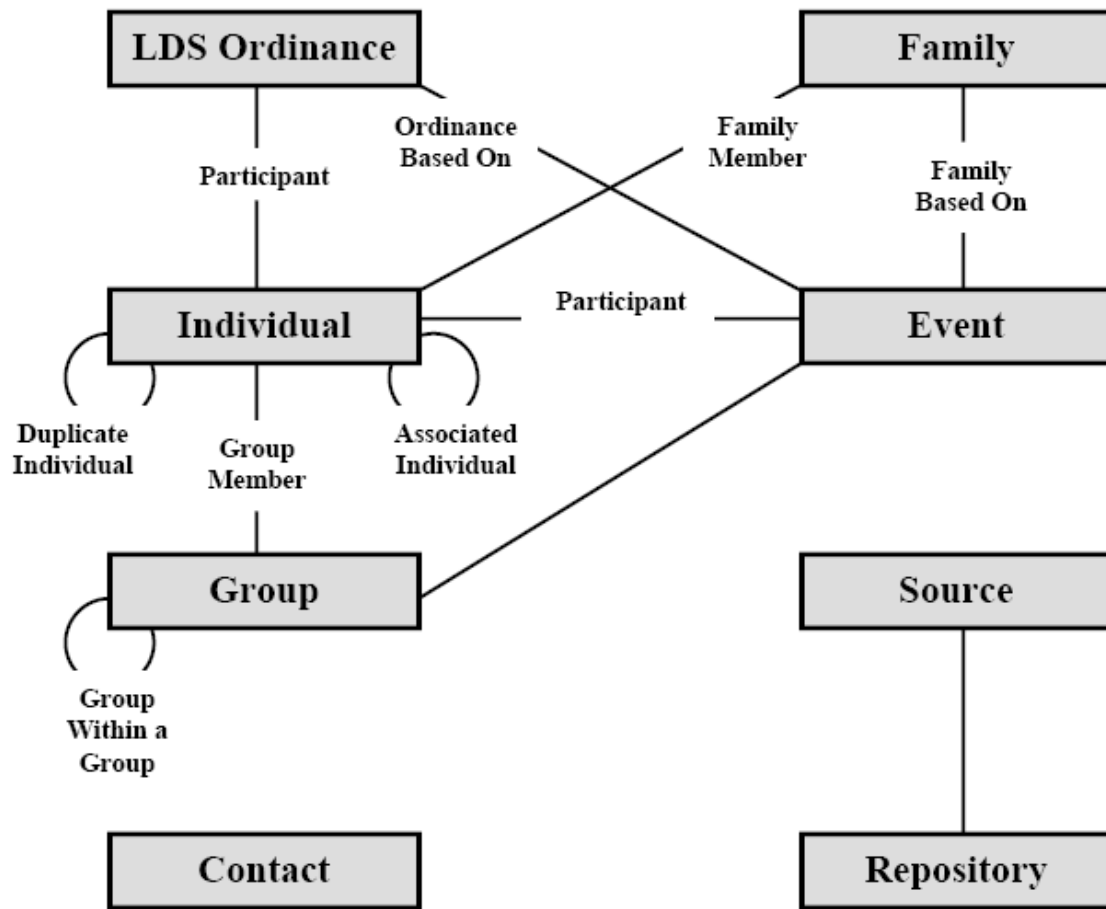


Diagram datových struktur GEDCOM XML 6.0



Příloha B - Ukázky prohlížeče

Rodokmen - Mozilla Firefox

Soubor Úpravy Zobrazit Přejít Zámky Nástroje Nápověda

http://a03-0530a/final_verzion/?zdroj=gedr1168.ged&indi=54

Mozilla Firefox Přehled zpráv

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resizer Tools View Source Options

[zpět na seznam](#)

John Wesley O'Donald/ (M)
* 18 JAN 1851
† 13 NOV 1927

Rebecca Ann CLINGMAN/ (F)
* 30 JAN 1858

Sarah Elizabeth O'Donald/ (F)
* 10 JUL 1876
† 1 JUN 1967

Henry Boyd McBride/ (M)
* 22 MAR 1868
x 1908
† 29 MAR 1946

John Henry/McBride/ (M)
* 22 SEP 1895

Charles Homer/McBride/ (M)
* 5 DEC 1897
† NOV 1967

Grace/McBride/ (F)
* 29 MAR 1900

Harry/McBride/ (M)
* 19 NOV 1902
† 7 FEB 1986

George/McBride/ (M)
* 11 AUG 1907
† 3 MAY 1989

Cecil Elmer/McBride/ (M)
* 26 JUL 1912

Bertha Hortense/McBride/ (F)
* 6 DEC 1914

Velma Eugene (Babe)/McBride/ (F)
* 23 AUG 1918

Images: 0/0 Loaded: 6 KB 54.55 KB/s Time: 0:00 Hotovo Adblock

Rodokmen - Mozilla Firefox

Soubor Úpravy Zobrazit Přejít Zámky Nástroje Nápověda

http://a03-0530a/final_verzion/index.php?zdroj=bush.ged&border=

Mozilla Firefox Přehled zpráv

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resizer Tools View Source Options

[zpět k mým souborům](#)

ID	jméno	rod	pohlaví	narozen
1	George Walker	/Bush/	M	6.7.1946
2	George Herbert Walker	/Bush/	M	12.6.1924
3	Prescott Sheldon	/Bush/	M	1895
4	Samuel Prescott	/Bush/	M	1863
5	Samuel Howard	/Fay/	M	1804
6	Susan	/Shellman/	F	
7	Clarrisa	/Montfort/	F	
8	Robert	/Munford/	M	
9	Samuel Prescott Phillips	/Fay/	M	
10	Harriet	/Howard/	F	1782
11	Jonathan	/Fay/	M	
12	Anna	/Lillie/	F	1760
13	John	/Fay/	M	
14	Elizabeth	/Wellington/	F	
15	Benjamin	/Wellington/	M	
16	Mary	/Pagrave/	F	
17	Richard	/Pagrave/	M	
18	Edward	/Pagrave/	M	
19	Thomas	/Pagrave/	M	
20	Robert	/Munford/	M	
21	John	/Lillie/	M	1728
22	Rebecca	/Bulkeley/	F	
23	Anne	/Cocke/	F	
24	Elizabeth	/Bolling/	F	
25	Clarissa	/Montfort/	F	
26	Hannah	/Hutchinson/	F	1669
27	Edward	/Bulkeley/	M	
28	Peter	/Bulkeley/	M	
29	Mary	/Aston/	F	

Images: 3/3 Loaded: 16 KB 78.82 KB/s Time: 0:00 Hotovo Adblock

Rodokmen - Mozilla Firefox

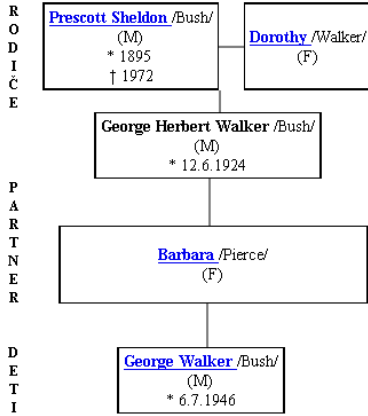
Šoubor Úpravy Zobražit Přejít Záložky Nástroje Nápožěda

http://a03-0530a/final_verzion/index.php?zdroj=bush.ged&indi=2

Mozilla Firefox Přehled zpráv

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options

[zpět na seznam](#)



Images: 0/0 Loaded: 4 KB 43.01 KB/s Time: 0:00 Hotovo Adblock