

## **Implementácia databáz knižnej grafiky**

Vladimír FILIP

Pracovisko:KMKF FHV UNIZA, Mail: vladimir.filip@fhv.uniza.sk

### **Abstrakt**

Článok sa zaoberá problematikou databázového spracovania údajov v kontexte projektu "Grafická výzdoba a typografia starých tlačí vydaných na Slovensku v kontexte vyvíjajúcej sa tlačiarskej techniky". Databázy sú využité na efektívnu manipuláciu s veľkým množstvom dát. Článok opisuje výhody databázového spracovania a využíva nástroje ako phpMyAdmin, WordPress a knižnice MariaDB Connector/J, Apache POI, iText Java a Apache PDFBox potrebné pre konverziu a manipuláciu s dátami. Implementácia je založená na adaptívnom a responzívnom dizajne, zabezpečujúc prístup a použiteľnosť pre rôzne typy zariadení.

### **Kľúčové slová**

grafická výzdoba, databázové spracovanie údajov, WordPress, databáza

### **Abstract**

The article deals with the issue of database processing in the context of the project "Graphic decoration and typography of old prints published in Slovakia in the context of evolving printing technology". Databases are used for efficient manipulation of large amounts of data. The article describes the advantages of database processing and uses tools such as phpMyAdmin, WordPress and the MariaDB Connector/J, Apache POI and Apache PDFBox libraries needed for data conversion and manipulation. The implementation is based on an responsive design, ensuring accessibility and usability for different types of devices.

### **Keywords**

graphic decoration, database processing, WordPress, database.

## Úvod

Predkladaný článok pojednáva o problematike databázového spracovania údajov v rámci projektu Grafická výzdoba a typografia starých tlačí vydaných na Slovensku v kontexte vyvíjajúcej sa tlačiarenskej techniky (*VEGA 1/0181/21*). V navrhovanom kontexte sa projekt zaoberá problematikou starých tlačí a prináša nový, kvalitatívne odlišný pohľad na túto tému. Hlavné ciele projektu definujú výstupy – prezentačnú vrstvu výsledkov identifikovaných v rámci jednotlivých fáz realizovaného výskumu. V rámci všeobecného prieskumu výsledkov podobne orientovaných projektov sme definovali ako výstup identifikovaných, resp. rozšírených poznatkov, databázový systém prezentácie.

Databázu vo všeobecnosti môžeme chápať ako organizovanú kolekciu dát resp. údajov, ktorá je štruktúrovaná tak, aby umožňovala efektívne ukladanie, spracovanie a následné získavanie informácií. Databáza slúži prioritne na uchovávanie väčšieho množstva dát a umožňuje rýchly a štruktúrovaný prístup a manipuláciu s nimi. Hlavným cieľom v podstate je poskytnúť spoľahlivý a efektívny spôsob ukladania a manipulácie s dátami pre rôzne aplikácie, resp. systémy (Janoušek 2023).

O získavaní informácií môžeme hovoriť ak informáciu chápeme ako množinu dát ktoré majú istú pridanú hodnotu v určitom kontexte. Táto hodnota spočíva predovšetkým v schopnosti prijímateľa, ktorý by jej mal porozumieť, a súčasne by mala samotná informácia vyvolať zmeny v prijímateľovom myslení alebo správaní a v konečnom dôsledku prispieť k zlepšeniu rozhodovacích procesov.

Keďže informácie môžu byť reprezentované v rôznych formách, vrátane textu, čísel, obrazov, zvukov alebo iných foriem. Ich efektívne využitie závisí na schopnosti správneho získavania, ukladania, spracovania a šírenia v rámci relevantného kontextu.

V roku 1970 začala história relačných databáz článkom pána Codd s názvom " A Relational Model of Data for Large Shared Data Banks ". Hlavným prínosom tejto teórie bolo, že dáta sa stali nezávislými od akéhokoľvek hardvéru alebo systému ukladania dát. V praxi to znamenalo, že dáta mohli byť organizované do tabuliek, ktoré boli poprepájané relačnými vzťahmi. Inak povedané vznikali vzťahy medzi rôznymi množinami informácií (Codd 1970).

Súčasné databázy sú kľúčovým nástrojom pre mnohé informačné systémy, webové aplikácie, podnikové aplikácie a ďalšie technologické riešenia, kde je potrebné efektívne spravovať a manipulovať s veľkým objemom dát (Janoušek 2023).

Dáta vložené do tabuliek sú organizované logicky na princípoch, ktoré

- minimalizujú duplicitu;
- redukujú anomálie dát;
- zabezpečujú integritu dát (Janoušek 2023).

Ako uvádza Hallová (2023) medzi výhody databázového spracovania dát môžeme zaradiť:

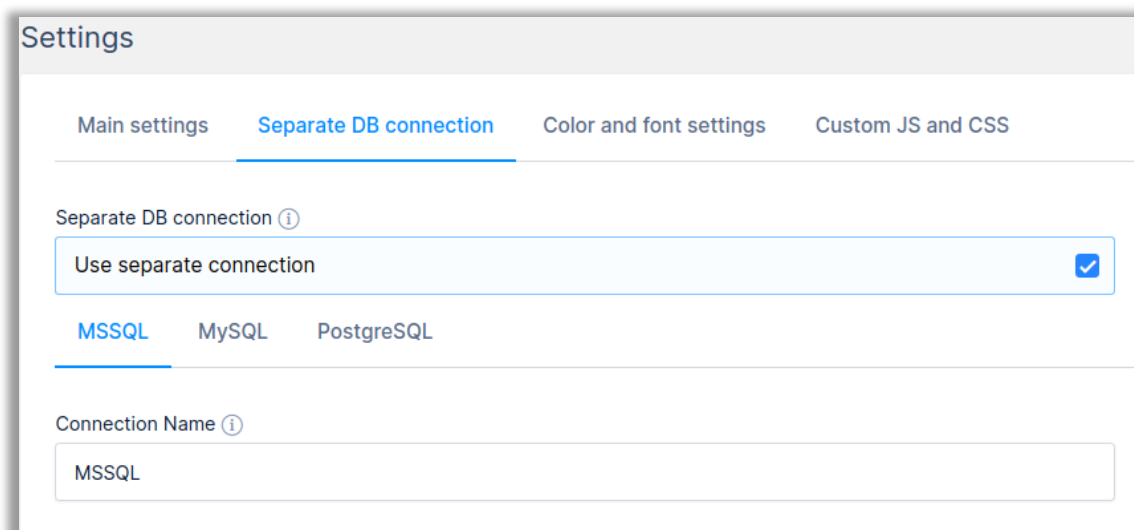
- štruktúry dátových súborov sú oddelené od aplikačných (*užívateľských*) programov;
- prístup k údajom je možný len prostredníctvom programov databázového systému a nie priamo;
- dáta je možné vyhodnotiť akýmkoľvek spôsobom;
- umožnený prístup viacerých užívateľov súčasne;
- zaistená ochrana údajov proti zneužitiu.

Databáza sa obvykle riadi systémom správy databáz (*DBMS*). Dáta spolu s DBMS a príslušnými aplikáciami tvoria celok označovaný ako databázový systém, často skrátené označovaný ako databáza. V súčasnosti sú údaje v bežných typoch databáz modelované v riadkoch a stĺpcoch usporiadaných do tabuliek, čo umožňuje efektívne spracovanie a vyhľadávanie údajov. Týmto spôsobom sú dáta ľahko prístupné, spravované, modifikované, aktualizované, kontrolované a organizované. Väčšina databáz využíva štruktúrovaný dotazovací jazyk (*SQL*) na zápis a vyhľadávanie údajov (Oracle 2023a).

Ako systém správy databáz sme zvolili phpMyAdmin (*vo verzii 5.2.1*). Tento softvérový nástroj napísaný v PHP, ktorý slúži na správu MySQL prostredníctvom webového rozhrania. phpMyAdmin podporuje rozsiahle možnosti manipulácie s databázami MySQL a MariaDB. Časté operácie, ako správa databáz, tabuliek, stĺpcov, vzťahov, indexov, užívateľov, oprávnení a podobne, sú realizovateľné pomocou užívateľského rozhrania. Okrem toho poskytuje možnosť priameho vykonávania ľubovoľného SQL príkazu.

Samotná prezentačná vrstva nášho riešenia je realizovaná v prostredí bezplatného softvéru CMS Wordpress (vo verzii 6.4.2. – verzia v čase písania článku). Podľa W3Techs (2023), WordPress poháňa 42,7 % všetkých webových stránok distribuovaných prostredníctvom internetu, vrátane tých, ktoré nemajú redakčný systém (CMS) alebo s CMS kódovaným na mieru. Ak obmedzíme bázu údajov iba na webové stránky, ktoré využívajú CMS, podiel WordPress na trhu sa stane ešte dominantnejším. V tom prípade má WordPress 62,5% podiel na trhu redakčných systémov. To predstavuje jeden z hlavných dôvodov prečo sme využili spomínaný CMS systém.

Pre samotnú vizualizáciu dát – údajov, je využívaný doplnok wpDataTables (plugin) v systéme Wordpress. Hovoríme o tabuľkovom doplnku, ktorý sa prioritne používa na rýchle vytváranie údajových tabuliek a tabuľkových grafov z rôznorodých zdrojov (Excel, CSV, XML, JSON, PHP, MySQL, ...).



Obrázok 1 Plugin wpDataTables, využívaný pri implementácii riešenia.

Veľmi kľúčovou úlohou v celom procese je samotné spracovanie dát, ktoré boli identifikované v rámci výskumu dotknutého projektu. Samotné spracovanie dát chápeme ako činnosti nutné pri manipulácii s dátami s cieľom tieto údaje transformovať v rámci zadanej úlohy. V praxi to znamenalo, všetky činnosti ktoré bolo potrebné vykonať s jednotlivými údajmi samotnými riešiteľmi projektu. Samotné spracovanie bolo spracovávané v tabuľkovom procesore Excel na základe interne definovanej metodiky. Po samotnom spracovaní bolo potrebné transformovať jednotlivé dáta (reprezentované tabuľkou) do samotnej databázy. Tu bolo nutné zdefinovať samotnú tabuľku, ktorá sú ekvivalentom už spracovaných tabuliek. Výstupom

z projektu sú dve databázy. Prvá pojednáva o sekundárnej literatúre a druhá, primárna, sa venuje jednotlivým grafickým záznamom, ktoré boli identifikované aj na základe sekundárnej literatúry.

Primárna databáza obsahuje prioritne textové záznamy. Základná báza atribútov, pomocou ktorých je možné prehliadať jednotlivé záznamy predstavuje nasledovné: (záznam, referencie, mesto, tlačiar, rok, ilustrátor, typ ilustrácie, obsah ilustrácie a kľúčové slová). Keďže pracujeme v prostredí textovej relačnej databázy vo webovom prostredí, obrázky sú reprezentované URL adresou daného obrázkového záznamu – URL môžeme chápať ako textový záznam (Mažútová a Filip 2023).

Záznam	Referencie	Mesto	Tlačiar	Rok	Ilustrátor	Typ ilustrácie	Obsah ilustrácie	Kľúčové slová	Obrázok
A Bdnös Embernek Meg-Terésére: Az igaz penitencia tartásnak Mógya Rövid summában elő-adatik A Nagy Szombati Jesus Társaságának Catechetica Bibliothekájától M.DCCXLVII. Nyomtatott. Nagy-Szombatban, az Akadémiai betűkkel, 1747. Esztendőben. 108 s.	BÚST 18, č. 1409	Trnava	Akademická tlačiareň	1747		signet	Signet Akademickej tlačiarne.	katolícka cirkev   morálka   náboženská   príručky	
Andächtige Übung; und Aufopferung des bitterm Leidens und Sterbens Unsers Heilands Jesu Christi, für die arme Seelen im Fegfeuer, bey der heiligen Meß, und auf alle Täg der Wochen ausgetheilet. Nebst angehängten Tagzeiten und Litaneu, zu Hülff deren Abgestorbenen Christgläubigen, Samt dem Kern aller Gebether. Pressburg: gedruckt und zu finden bey Franz Augustin Patzko, 1771. 228 s.: [1] obr.	BÚST 18, č. 358	Bratislava	Patzko, František Augustín	1771	Zherhofer, A.	frontispice		katolícka cirkev   modlitby	

Obrázok 2 Náhľad na primárnu databázu - výstup projektu.

Databáza sekundárnej literatúry obsahuje výlučne textové záznamy, kde základná báza atribútov pozostáva z atribútov (*stĺpcov*) popis, signatúra a kľúčové slová. Filtrovať jednotlivé údaje je možné na základe všetkých uvedených atribútov. Obe databázy sú prístupné prostredníctvom webového sídla na URL adrese: <https://grafikastarychtlaci.uniza.sk>.

Primárna grafika	Grafika s.I.
grafika_id : int(11)	grafika_id : int(11)
popis : text	popis : text
referencie : text	signatura : text
signatura : text	anotacia : text
klucove_slova : text	poznamka : text
mesto : text	klucove : text
tlaciar : text	
rok : text	
poznamky : text	
ilustrator : text	
typ_ilustracie : text	
obsah_ilustracie : text	
url : text	

Obrázok 3 Tabuľky databáz s jednotlivými atribútmi a ich dátovými typmi.

Pre samotnú transformáciu z už spracovaných údajov v Excel dokumentov bol využitý na mieru vytvorené algoritmické spracovanie vytvorené v programovacom jazyku JAVA. Samotná programy zodpovedné za úpravu a prenos dát, do výsledných tabuliek reprezentovaných stranou phpMyAdmin, bolo spracované vo vývojom prostredí NetBeans. Prostredie poskytuje nástroje pre vývoj Java webových, podnikových, desktopových a mobilných aplikácií. Ďalšou výhodou samotného nástroja je že ide o bezplatné a open source prostredie pre vývoj aplikácií v operačných systémoch Windows, Mac, Linux a Solaris (Oracle 2023b).

Keďže nie všetka funkcionálna je dostupná v rámci samotného prostredia a súčasne jej vývoj by znamenal relatívne vyššiu časovú náročnosť. Rozhodli sme sa využívať už dostupné knižnice. Prvou knižnicou, ktorú sme sa rozhodli použiť je MariaDB Connector/J (*Dostupná na URL adrese: <https://mvnrepository.com/artifact/org.mariadb.jdbc/mariadb-java-client>*).

MariaDB Connector/J je predstavuje ovládač, ktorý bol vyvinutý špeciálne ako jednoduchý konektor JDBC (*Java database connectivity*) na použitie s databázovými servermi MariaDB a MySQL. Pôvodne bol založený na kóde Drizzle JDBC s mnohými doplnkami a opravami chýb (MariaDB 2023).

Samotný ovládač je zodpovedný za všetky procesy týkajúce sa s manipuláciou s tabuľkami na strane servera kde sa nachádzajú samotné tabuľky databázy.

Obdobný prípad nastáva pri manipulácii s Excel súborami. Táto funkcionálna nie je obsiahnutá v bežnom balíčku základných nástrojov pre vývoj aplikácií JDK. Preto bolo nutné importovať knižnicu Apache POI, ktorá sprístupňuje metódy prístupu k súborom vo formátoch od spoločnosti Microsoft. (Dostupná na URL adrese: <https://mvnrepository.com/artifact/org.apache.poi/poi>)

Samotný algoritmus konverzie údajov spočíva na načítaní jednotlivých údajov v konkrétnej bunke v Excel dokumentu. Tieto údaje sú následne vkladane prostredníctvom knižnice MariaDB.jdbc do odpovedajúcej tabuľky konkrétnej databázy prostredníctvom príkazu INSERT (jazyk SQL).

```
public void vloz_do_DB_zaznam(String a, String b, String c, String d, String e)
{
    String insert = "INSERT INTO chronologia (popis,signatura,klucove,anotacia,poznamka) "
        + "VALUES ('"+a+"','"+b+"','"+c+"','"+d+"','"+e+"')";
    try {
        Statement stmt = (Statement) conn.createStatement();
        stmt.executeUpdate(insert);
        System.out.println("Insert sa podaril");
    } catch (SQLException e1)
    {
        System.out.println("Niečo sa pokazilo pri inserte");
    }
}
```

Obrázok 4 Kód využívaný pre samotné vkladanie dát do databázy.

Primárna databáza obsahuje grafickú prílohu (samotné grafické výstupy), ktoré sme sa rozhodli konvertovať do podoby súboru \*.pdf. Nakoľko niektoré záznamy obsahujú viacero čiastkových grafických výstupov. Z tohto dôvodu je dokument PDF vyhovujúci. Nakoľko samotných grafických súborov je relatívne veľké množstvo, bolo nutné tieto algoritmicky spracovať. Manuálne spracovanie by predstavovalo výraznú časovú náročnosť. Keďže základný balík JDK neobsahuje knižnice pre prácu s PDF súborami bolo nutné túto funkcionálnu importovať. Využili sme knižnicu iText Java ktorá umožňuje vytvárať a spravovať dokumenty PDF. (Dostupná na URL adrese: <https://mvnrepository.com/artifact/com.itextpdf/itextpdf/5.1.0>)

Pri skenovaní grafických formátov niektoré z inštitúcií, ktoré vykonávali samotný proces skenovania vyhotovili výsledné skenovanie do súboru formátu \*.pdf. Finálne prevedenie obsahuje grafický náhľad dokument, čo predstavuje problém pri PDF súbor, ktorý je nepoužiteľný ako náhľad. Z tohto dôvodu bolo nutné vykonať konverziu z formátu PDF do

rastrového formátu \*.jpg. Bolo nutné importovať knižnicu Apache PDFBox. Knižnica rozširuje funkcionality o vytváranie nových dokumentov PDF, manipuláciu s existujúcimi dokumentmi a možnosť extrahovať obsah z dokumentov. (Dostupná na URL adrese: <https://mvnrepository.com/artifact/org.apache.pdfbox/pdfbox-app/2.0.23>)

Ďalším z dôležitých cieľov kladených na samotné technické vypracovanie je zaručiť dobrú dostupnosť a použiteľnosť pre používateľov. Výsledná realizácia je implementovaná na základných paradigmách adaptívneho a responzívneho dizajnu. Keďže v súčasnosti bežný používateľ prístupuje na web prostredníctvom rôznorodých typov zariadení a rozlíšení obrazovky, je nanajvýš dôležité, aby bola zabezpečená istá komfortnosť a plynulosť z pohľadu ich práce so systémom (Mažútová a Filip 2023).

## **Záver**

Predkladané databázy sú neoddeliteľnou súčasťou a jedným z hlavných výstupov vyššie spomínaného projektu. Hlavné ciele projektu smerujú k prezentačnej vrstve výsledkov, identifikovaných v jednotlivých fázach výskumu. Súčasnú databázu sú nevyhnutné pre informačné systémy, webové a podnikové aplikácie, poskytujúce efektívne spravovanie veľkého objemu dát. Výber databázového systému, v tomto prípade phpMyAdmin, a redakčného systému WordPress, ukazuje praktickú implementáciu. Dôkladné spracovanie údajov v Excel a ich transformácia do databáz je kľúčovým krokom v projekte. Technická implementácia využíva nástroje, ako MariaDB Connector/J pre prácu s databázou a Apache POI pre manipuláciu s Excel súbormi. Knižnice iText Java a Apache PDFBox zabezpečujú konverziu a prácu s PDF súbormi. Rovnako je kladený dôraz na dostupnosť a použiteľnosť pre používateľov, implementovanú na princípoch adaptívneho a responzívneho dizajnu.



## Zoznam použitých zdrojov

CODD, E.F., 1970. A Relational Model of Data for Large Shared Data Banks. In: *Communications of the A. C. M.* [online]. New York: The Association for Computing Machinery. 1970, roč. 13, č. 6 [cit. 2023-16-12]. Dostupné na: <https://www.seas.upenn.edu/~zives/03f/cis550/codd.pdf>

JANOŤEK, Jaroslav, 2023. *Relačné databázy - pojem, história a vlastnosti* [online]. Košice: SPŠ elektrotechnická, 2023 [cit. 2023-16-12]. Dostupné na: <https://www.spseke.sk/tutor/prednasky/reldatab.htm>

HALLOVÁ, Marcela, 2023. Databázové spracovanie dát. *Databázové systémy* [online]. Nitra: SPU, 2023 [cit. 2023-16-12]. Dostupné na: <https://moodle.uniag.sk/mod/book/view.php?id=3032&chapterid=784&lang=en>

ORACLE, 2023a. *What Is a Database?* [online]. Redwood Shores: Oracle Corporation, 2023 [cit. 2023-16-12]. Dostupné na: <https://www.oracle.com/database/what-is-database/>

ORACLE, 2023b. *The Smarter and Faster Way to Code* [online]. Redwood Shores: Oracle Corporation, 2023 [cit. 2023-16-12]. Dostupné na: <https://www.oracle.com/tools/technologies/netbeans-ide.html>

W3TECHS, 2023. *Usage statistics and market share of WordPress* [online]. Rebengrund: Q-Success, 2023 [cit. 2023-16-12]. Dostupné na: <https://w3techs.com/technologies/details/cm-wordpress>

MARIADB, 2023. *MariaDB MaxScale* [online]. Redwood: MariaDB Foundation, 2023 [cit. 2023-16-12]. Dostupné na: <https://mariadb.com/kb/en/maxscale/>

MAŽGÚTOVÁ Veronika a Vladimír FILIP, 2023. Databázy ako výsledok projektovej činnosti v akademickom prostredí a ich význam pre výskum dejín knihy a knižníc. In: *Kniha 2023: Zborník o problémoch a dejinách knižnej kultúry*. Martin: Slovenská národná knižnica, 2023, s. 365 – 375. ISBN 978-80-8149-165-8.