Západočeská univerzita v Plzni

Fakulta aplikovaných věd

Katedra informatiky a výpočetní techniky

Diplomová práce

Systém pro automatickou kontrolu samostatných prací vytvořených v MS Access

Plzeň, 2018 Vojtěch Kinkor

Zadání

1. Seznamte se s formátem souboru MS Access .accdb a možnostmi jeho čtení.
2. Seznamte se validátorem na portálu ZČU.
3. Vytvořte konfigurovatelný systém pro kontrolu samostatných prací vytvořených v MS Access se zaměřením na splnění zadání a plagiarismus.
4. Část systému pro kontrolu splnění zadání adaptujte pro validátor.
5. Celý systém pečlivě otestujte.

**< tato strana bude nahrazena originálním zadáním / kopií ! >**

Prohlášení

Prohlašuji, že jsem diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů.

V Plzni dne <TODO: 0. 0. 2018>

Vojtěch Kinkor

Abstract

**Systém pro automatickou kontrolu samostatných prací vytvořených   
v MS Access**

text

Abstrakt

**System for Automatic Checking of Student Works Created in MS Access**

text

Obsah

[1 Úvod 1](#_Toc510188371)

[2 Databázový software Microsoft Access 2](#_Toc510188372)

[2.1 Základní informace 2](#_Toc510188373)

[2.2 Prvky databáze 2](#_Toc510188374)

[2.2.1 Tabulky 2](#_Toc510188375)

[2.2.2 Vazby mezi tabulkami **Chyba! Záložka není definována.**](#_Toc510188376)

[2.2.3 Uložené dotazy 6](#_Toc510188377)

[2.2.4 Formuláře 6](#_Toc510188378)

[2.2.5 Sestavy 6](#_Toc510188379)

[2.3 Formát MDB a ACCDB 6](#_Toc510188380)

[2.3.1 Metadata 6](#_Toc510188381)

[2.4 Možnosti čtení souborů ACCDB 6](#_Toc510188382)

[2.4.1 ODBC 6](#_Toc510188383)

[2.4.2 Microsoft Office Interop 6](#_Toc510188384)

[2.4.3 MDBTools 6](#_Toc510188385)

[2.4.4 Jackcess 6](#_Toc510188386)

[2.4.5 Další možnosti (JDBC) 6](#_Toc510188387)

[3 Portál ZČU 7](#_Toc510188388)

[3.1 Základní informace 7](#_Toc510188389)

[3.2 Validátor studentských prací 7](#_Toc510188390)

[3.3 Vytvoření nové validační domény 7](#_Toc510188391)

[4 Analýza řešení 8](#_Toc510188392)

[4.1 Požadavky na řešení 8](#_Toc510188393)

[4.2 Případy užití 8](#_Toc510188394)

[4.3 Validace databáze 8](#_Toc510188395)

[4.4 Vyhodnocení plagiarismu 8](#_Toc510188396)

[5 Implementace 9](#_Toc510188397)

[5.1 Použité technologie 9](#_Toc510188398)

[5.2 Struktura aplikace 9](#_Toc510188399)

[5.3 Validace databáze 9](#_Toc510188400)

[5.4 Implementovaná validační pravidla 9](#_Toc510188401)

[5.5 Hledání podobností a detekce plagiarismu 9](#_Toc510188402)

[5.6 Grafické rozhraní 9](#_Toc510188403)

[5.7 Adaptace pro validátor portálu ZČU 9](#_Toc510188404)

[6 Testování 10](#_Toc510188405)

[6.1 Validační pravidla 10](#_Toc510188406)

[6.2 Konzolová aplikace pro validátor portálu ZČU 10](#_Toc510188407)

[6.3 Grafické rozhraní 10](#_Toc510188408)

[7 Závěr 11](#_Toc510188409)

[Reference 12](#_Toc510188410)

[Přílohy 12](#_Toc510188411)

[A Uživatelská příručka 13](#_Toc510188412)

[Spuštění a kompilace nástroje 13](#_Toc510188413)

[Obsluha nástroje 13](#_Toc510188414)

[B Obsah přiloženého média 13](#_Toc510188415)

# Úvod

# Databázový software Microsoft Access

## Základní informace

Microsoft Access je nástroj řadící se mezi takzvané systémy řízení báze dat (SŘBD či DBMS – *database management systém*). Jedná se o software, který umožňuje práci s relačními databázemi. Je součástí kancelářského balíku Microsoft Office, případně prodáván i samostatně.

Aplikace používá pro ukládání dat technologii Microsoft Jet Database Engine. Jednotlivé databáze jsou typicky uloženy v jediném souboru ve formátu MDB, nebo ACCDB.

Pro vytváření a správu databáze je uživateli dostupné přehledné grafické rozhraní.

## Prvky databáze

Dále jsou uvedeny různé prvky, které mohou být součástí databáze.

### Tabulky

Jedná se o stěžejní součást každé databáze. Tabulku lze definovat jako strukturovanou kolekci dat. Skládá se ze sloupců a řádků (též záznamů) a v rámci databáze má unikátní název.

Sloupce tabulky

Struktura je definována pomocí sloupců, které mají specifikovaný název (unikátní v rámci tabulky) a datový typ. Microsoft Access[[1]](#footnote-1) podporuje následující datové typy [[1](#_toc_1)]:

* **Automatické číslo** – typicky používáno jako primární klíč (viz dále), pro každý nový záznam se automaticky nastaví na následující hodnotu posloupnosti, nebo na náhodné číslo (dle nastavení).
* **Číslo** – rozsah a typ (celočíselné/s desetinnou čárkou) lze zvolit ve vlastnostech sloupce.
* **Krátký text** (dříve Text) – text do délky 255 znaků.
* **Dlouhý text** (dříve Memo) – text do velikosti 1 GB.
* **Datum a čas**.
* **Měna** – specializovaný případ číselného datového typu s fixní desetinnou čárkou (uchovává 4 desetinná místa).
* **Ano/ne** – uchovává hodnotu -1 (Ano) nebo 0 (Ne); v rámci Microsoft Accessu zobrazeno jako zaškrtávací pole (*checkbox*).
* **Hypertextový odkaz.**
* **Objekt OLE** – umožňuje vložit speciální objekty, například obrázek, jiný dokument, či odkaz na soubor.
* **Příloha** – umožňuje vložit libovolný soubor jako součást záznamu. Jedná se o univerzálnější možnost k předchozímu.
* **Počítané** – automatické vložení hodnoty vypočítané na základě zadaného vzorce.

Každému sloupci lze dále nastavit různé vlastnosti dle vybraného datového typu – typicky se jedná o ověřovací pravidla (validace vstupu od uživatele ještě před přidáním záznamu do databáze), výchozí hodnotu a dále nastavení zobrazení v tabulce (formátování, zarovnání, titulek po najetí myší, atp.).

Primární klíč

Tabulka může mít primární klíč – typicky se jedná o sloupec, jehož hodnoty jsou unikátní a vždy zadané (tzv. *not null*). V případě, že vytvoříme primární klíč pomocí více sloupců, nazýváme jej složeným primárním klíčem.

Primární klíč slouží pro odkázání na jeden konkrétní záznam v tabulce, čehož se využívá při vytváření dotazů nebo tvoření relací mezi tabulkami.

Pro vytváření primárních klíčů se obvykle využívá datový typ Automatické číslo, který každému záznamu přiřadí unikátní celé číslo. Často bývá takový sloupec pojmenován „ID“ (*Identification*).

Relace mezi tabulkami a cizí klíče

V případě, že chceme propojit více tabulek mezi sebou, využijeme tzv. relačních vazeb. Jedná se o situaci, kdy se záznam v tabulce odkazuje na konkrétní záznam (či záznamy) v druhé tabulce.

Existují tři druhy relačních vazeb.

* **Relace 1:1** – jednomu záznamu v tabulce A odpovídá žádný či právě jeden záznam v tabulce B. Pro referencování se využívají pouze primární klíče obou tabulek (mají tedy v obou tabulkách shodnou hodnotu).



Obrázek 1 – model relace 1:1

* **Relace 1:N** – k více záznamům v tabulce A lze přiřadit jeden záznam v tabulce B. To lze zajistit přidáním tzv. **cizího klíče** do tabulky A – sloupce, který bude obsahovat hodnoty primárního klíče z tabulky B (příp. skupiny sloupců, pokud se jedná o složený primární klíč). Jedná se o nejčastěji využívanou vazbu.



Obrázek 2 – model relace 1:N

* **Relace M:N** – k M záznamům v tabulce A lze přiřadit N záznamů v tabulce B. Relace se realizuje pomocí spojové tabulky (též mezitabulky) a dvojicí relací 1:N. Spojová tabulka obvykle obsahuje pouze sloupce cizích klíčů.



Obrázek 3 – model relace M:N

Relace mezi tabulkami mohou zajišťovat **referenčním integritu**. Cílem je zabránit odkazování na neexistující záznam či vzniku osiřelých záznamů, na který byly všechny reference zrušeny. Integritní pravidlo může zajistit kaskádovou aktualizaci polí – pokud se změní hodnota primárního klíče, změní se automaticky hodnota u všech záznamů, které na záznam odkazují. Dále může zajistit kaskádové odstranění souvisejících záznamů – v případě smazání záznamu budou smazány i všechny záznamy, které na tento záznam odkazovaly.

### Dotazy

### Formuláře

### Sestavy

## Formát MDB a ACCDB

### Metadata

## Možnosti čtení souborů ACCDB

### ODBC

### Microsoft Office Interop

### MDBTools

### Jackcess

### Další možnosti (JDBC)

# Portál ZČU

## Základní informace

## Validátor studentských prací

## Vytvoření nové validační domény

# Analýza řešení

## Požadavky na řešení

## Případy užití

## Validace databáze

## Vyhodnocení plagiarismu

# Implementace

## Použité technologie

## Struktura aplikace

## Validace databáze

## Implementovaná validační pravidla

## Hledání podobností a detekce plagiarismu

## Grafické rozhraní

## Adaptace pro validátor portálu ZČU

# Testování

## Validační pravidla

## Konzolová aplikace pro validátor portálu ZČU

## Grafické rozhraní

# Závěr

# Reference

[1] Data types for Access desktop databases. *Microsoft Office help and training - Office Support.* [Online] [Citace: 22. 3. 2018]. Dostupné z: <https://support.office.com/en-us/article/data-types-for-access-desktop-databases-df2b83ba-cef6-436d-b679-3418f622e482>

# Přílohy

## A Uživatelská příručka

### Spuštění a kompilace nástroje

### Obsluha nástroje

Typický postup práce s nástrojem je následující:

1. ...

## B Obsah přiloženého média

Součástí práce je přiložené paměťové médium (DVD) obsahující tyto adresáře a soubory:

* Project/ – adresář obsahující projekt vytvořené aplikace,
* Kinkor\_A16N0040P\_DP.pdf – text této práce ve formátu PDF,
* readme.txt – textový soubor obsahující popis struktury DVD.

Obsah přiloženého média (včetně aktuální verze nástroje) je možné najít též v repozitáři projektu v rámci služby GitHub na adrese:

<https://github.com/ikeblaster/access-validator/>

1. Aktuálně ve verzi 2016. [↑](#footnote-ref-1)