

UNIVERZITA HRADEC KRÁLOVÉ  
UNIVERSITY OF HRADEC KRÁLOVÉ



FAKULTA INFORMATIKY A MANAGEMENTU  
FACULTY OF INFORMATICS AND MANAGEMENT

KATEDRA INFORMATIKY A KVANTITATIVNÍCH METOD  
DEPARTMENT OF INFORMATICS AND QUANTITATIVE METHODS

ANALÝZA DAT INFORMAČNÍHO SYSTÉMU  
INFORMATION SYSTEM DATA ANALYSIS

BAKALÁŘSKÁ PRÁCE  
BACHELOR'S THESIS

AUTOR PRÁCE  
AUTHOR

JAN HLADĚNA

VEDOUCÍ PRÁCE  
SUPERVISOR

prof. RNDr. HANA SKALSKÁ, CSc.

HRADEC KRÁLOVÉ 2015

## **ABSTRAKT**

Tato bakalářská práce se zabývá analýzou dat školního informačního systému Bakaláři. Navrhuje metodiky přístupu k analýze obecného informačního systému skrze identifikaci prostředí, ve kterém je systém provozován, a skrze poznání systému samotného. Na příkladě ilustruje rozdíly mezi širší (doménovou) znalostí a užší (pouze systémovou) znalostí informačního systému ve zvoleném prostředí. Podrobněji se zabývá prostředím českého školství, jako své zvolené znalostní domény, a administrativou a legislativou spojenou s vedením školní agendy. Detailně rozebírá populární školní informační systém Bakaláři a jím používané datové struktury. Součástí práce je případová studie s provedenou analýzou dat ve skutečném prostředí základní školy.

## **KLÍČOVÁ SLOVA**

analýza dat, školství, školní informační systém

## **ABSTRACT**

This bachelor thesis deals with data analysis of Czech school management information system Bakaláři. It proposes a methodology for a basic approach to data analysis of any given information system through identification it's surrounding environment, and through identification inner characteristics of the information system itself. Provided example illustrates the difference between the wider (domain) knowledge and narrower (system level only) knowledge in the chosen environment. This thesis specifically deals with the Czech educational system as it's main area of interest, and with the administration and legislation related to school management. It describes popular Czech school management information system Bakaláři and it's data structures in detail. The included case study applies described theoretical methodologies on a real environment of a primary school.

## **KEYWORDS**

data analysis, Czech educational system, school information management system

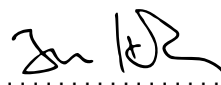
HLADĚNA, Jan: *Analýza dat informačního systému*: bakalářská práce. Hradec Králové: Univerzita Hradec Králové, Fakulta informatiky a managementu, Katedra informatiky a kvantitativních metod, 2015. 90 s. Vedoucí práce byla prof. RNDr. Hana Skalská, CSc.

## PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Analýza dat informačního systému“ jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., O právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4, Trestního zákoníku č. 40/2009 Sb.

Hradec Králové .....



.....  
(podpis autora)

# ZÁKLADNÍ ŠKOLA PARDUBICE - STUDÁNKA

POD ZAHRADAMI 317, 530 03 PARDUBICE



Na žádost studenta Jana Hladěny

Číslo jednací:

Vyřizuje: Mgr. Nejedlá

Pardubice dne: 14.4.2015

Potvrzuji, že Jan Hladěna, narozený 23.9.1987, bydl. Hraniční 723 Pardubice, ve školním roce 2014/2015 provádí analýzu dat školního informačního systému Bakaláři. Výzkum s daty bude sloužit jako podklad pro tvorbu bakalářské práce Jana Hladěny. Výzkum probíhá se souhlasem ředitelství školy a data jsou zanonymizovaná tak, aby nedošlo k porušení zákona č. 101/2000 Sb. v platném znění (zákon O ochraně osobních údajů).

V Pardubicích 14.4.2015

Ivana Nejedlá, ředitelka školy

Telefon / fax  
466651602-3

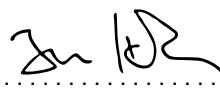
Bankovní spojení  
42534-561/ 0100

IČO  
48161276

## PODĚKOVÁNÍ

Děkuji prof. RNDr. Haně Skalské, CSc., za odborné vedení bakalářské práce a za doporučení vhodného nástroje pro zpracování dat. Dále děkuji ředitelce Základní školy Pardubice – Studánka, Mgr. Ivaně Nejedlé, za umožnění přístupu k pedagogické dokumentaci školy pro účely praktické aplikace teoretických metodik analýzy dat. Děkuji také Mgr. Jiřímu Mandysovi a Mgr. Zuzaně Virtové za konzultace a připomínky, které mi pomohly práci kompletovat.

Hradec Králové .....



.....  
(podpis autora)

# OBSAH

Úvod a cíle práce	1
<b>1 Informační systémy a data</b>	<b>2</b>
1.1 Využití informačních systémů	2
1.2 Analýza dat	3
1.2.1 Znalosti, informace a data	3
1.2.2 Identifikace prostředí a inventarizace systému	4
1.2.3 Extrakce, selekce, agregace a konsolidace dat	5
1.2.4 Dolování dat	6
1.2.5 Zpracování získaných dat a interpretace výsledků	7
<b>2 Specifická oblast - školství</b>	<b>9</b>
2.1 Vzdělávání jako výrobní proces	9
2.2 Informační systémy ve školství	10
2.2.1 Provozně-hospodářský úsek	10
2.2.2 Intranet, e-learning a interaktivní výuka	11
2.2.3 Prezentace školy a komunikace s veřejností	12
2.2.4 Informační systémy školní agendy	13
<b>3 Bakaláři - programy pro školní administrativu</b>	<b>19</b>
3.1 Rozšířenost systému	19
3.2 Stručný přehled systému a jeho modulů	19
3.2.1 Společné prostředí	20
3.2.2 Evidence žáků a zaměstnanců	20
3.2.3 Rozvrh hodin, generátor rozvrhů a suplování	21
3.2.4 Třídní kniha	21
3.2.5 Ostatní moduly a možnosti	22
3.2.6 Webová aplikace	22
3.2.7 Orientační ceník modulů pro školní rok 2014/15	23
3.3 Architektura programů a jejich provoz	24
3.3.1 Aplikační architektura programů	25
3.3.2 Provoz v síťovém prostředí	25
3.3.3 Provoz v jiných operačních systémech	30
3.3.4 Provoz webové aplikace	32
3.4 Podoba a formát uložených dat	37
3.4.1 Důležité tabulky	38
3.4.2 Datové typy FoxPro	39

3.4.3	Data programů Bakaláři . . . . .	39
3.4.4	Čtení a modifikace dat mimo prostředí programů . . . . .	40
3.5	Shrnutí . . . . .	41
<b>4</b>	<b>Případová studie</b>	<b>42</b>
4.1	Základní škola Pardubice – Studánka . . . . .	42
4.1.1	Historie nasazení IS Bakaláři . . . . .	42
4.1.2	Současný stav užívání IS Bakaláři . . . . .	43
4.2	Specifikace analýzy dat . . . . .	44
4.2.1	Anonymizace dat . . . . .	45
4.2.2	Transformace a konsolidace dat . . . . .	46
4.3	Výsledky a interpretace analýzy . . . . .	47
4.3.1	Míra využití IS Bakaláři . . . . .	47
4.3.2	Vývoj celkového hodnocení žáků . . . . .	47
4.3.3	Sledování nečíselných charakteristik . . . . .	49
4.3.4	Hledání kauzality v datech – vliv přítomnosti asistenta pedagoga na celkový výkon třídy . . . . .	49
	<b>Závěry a doporučení</b>	<b>51</b>
	<b>Literatura</b>	<b>53</b>
	<b>Seznam symbolů a zkratk</b>	<b>58</b>
	<b>Seznam příloh</b>	<b>62</b>

## SEZNAM OBRÁZKŮ

2.1	Diagram – vztah základních pedagogických dokumentů. Modře je vyznačen tok informací, červeně jejich tvorba. . . . .	13
2.2	Mapa – Geografická distribuce škol v České republice . . . . .	16
3.1	Diagram – obecná třívrstvá architektura typu klient–server . . . . .	24
3.2	Diagram – síťová komunikace programů Bakaláři v konfiguraci s daty uloženými na SQL serveru a webovým serverem pro elektronickou žákovskou knížku . . . . .	29
3.3	Graf – zastoupení jednotlivých webových serverů . . . . .	34
3.4	Diagram – zjednodušený relační model základní evidence celkové klasifikace pro každý školní rok . . . . .	37
4.1	Graf – evidované počty chorob daných kategorií . . . . .	49
B.1	Mapa – Bezkontextová geografická distribuce škol provozujících webovou aplikaci . . . . .	67
B.2	Diagram – prohledávání první úrovně . . . . .	70
B.3	Diagram – prohledávání druhé úrovně . . . . .	71
B.4	Diagram – prohledávání druhé úrovně – křížení odkazů . . . . .	72
B.5	Mapa – Geografický kontext škol s webovou aplikací . . . . .	73
B.6	Graf – rozdělení serverových technologií získané různými metodami . . . . .	77
B.7	Graf – rozdělení typů škol získané různými metodami . . . . .	78



## SEZNAM TABULEK

2.1	Celkové počty škol v ČR ke šk. roku 2014/2015 dle uvedených zdrojů	16
3.1	Ceník plné licence modulů v Kč vč. DPH platný od 1. listopadu 2014 [10]	23
3.2	Podíly typů škol využívajících webovou aplikaci podle výběru z číselníku AKDT – Druhy a typy škol a školských zařízení . . . . .	35
3.3	Zastoupení názvů v adresách URL . . . . .	36
B.1	Podíly typů škol získané metodou veřejného vyhledávání . . . . .	68
B.2	Podíly technologií získané metodou veřejného vyhledávání . . . . .	68
B.3	Podíly typů škol výběru z číselníku AKDT (Druhy a typy škol a školských zařízení) získané hloubkovým prohledáváním I. úrovně . . . . .	74
B.4	Podíly technologií získané metodou hloubkového prohledávání I. úrovně	75
B.5	Podíly typů škol výběru z číselníku AKDT (Druhy a typy škol a školských zařízení) získané hloubkovým prohledáváním II. úrovně . . . . .	75
B.6	Podíly technologií získané metodou hloubkového prohledávání II. úrovně	76
B.7	Celkové metriky použitých metod . . . . .	76
B.8	Porovnání metod – podíly serverových technologií . . . . .	77
B.9	Porovnání metod – podíly typů škol . . . . .	78
E.1	Vliv přítomnosti asistenta na celkový výkon třídy . . . . .	88
E.2	Vliv přítomnosti asistenta na výkon třídy – cizí jazyky . . . . .	88
E.3	Vliv přítomnosti asistenta na výkon třídy – společenskovědní předměty	89
E.4	Vliv přítomnosti asistenta na výkon třídy – přírodovědné předměty .	89

# ÚVOD A CÍLE PRÁCE

Předkládaná bakalářská práce se zabývá problematikou možností analýzy dat obecného informačního systému, ke kterému je přistupováno jako k uzavřené černé skřínce. Práce si klade za cíl dodržet vymezený postup, popsat zvolenou oblast a zvolený informační systém, provést samotnou praktickou analýzu dat a pokusit se o správnou reprezentaci a interpretaci získaných výsledků.

Úvodní, teoretická část, krátce pojednává o informačních systémech a dává ve filosofické rovině do souvislosti základní vztahy mezi znalostmi, informacemi, daty a přístupem samotných uživatelů k informačním systémům. Pokračuje obecnými doporučeními pro vytyčení metodiky přístupu k IS skrze poznání prostředí, ve kterém pracují, a skrze nutnost poznání systémů samotných. Závěr kapitoly zmiňuje obecné postupy pro přípravu dat k analýze a následnou interpretaci jejich výsledků, která musí být provedena v souladu se zvyklostmi zvoleného prostředí. Ačkoliv se jím tato práce nezabývá, ve stručnosti zde také otevírá téma metod pro dolování dat.

Následující, teoreticko-analytická část, se věnuje vybrané specifické oblasti, kterou je v tomto případě prostředí českého školství (kap. 2). Kapitola rozebírá jednotlivé podoblasti nasazení informačních systémů a zmiňuje důvody jejich zavádění. Pozornost je zde věnována vedení školní agendy a pedagogické dokumentace, jsou zde zmíněny některé charakteristiky systémů pro správu školní agendy.

Z dostupných systémů pro správu školní agendy je vybrán populární IS Bakaláři (kap. 3). První část kapitoly pojednává o systému z obecného pohledu, druhá se mu detailněji věnuje na úrovni provozování, v souladu s v úvodní kapitole zmiňovanými metodikami. Součástí je příloha B, která na jednom z modulů IS ilustruje rozdíl mezi systémovým a doménovým přístupem k analýze dat. Závěr kapitoly je orientován na popis fyzické struktury dat a na možnosti jejich samotné analýzy.

Ryze analytickou částí práce je konkrétní případová studie (kap. 4), kde je v souladu s popsanými metodikami a teoretickými pojednáními provedena samotná praktická analýza dat konkrétní instance informačního systému Bakaláři na základní škole Pardubice – Studánka.

Závěr je věnován shrnutí a zhodnocení výsledků práce, sebereflexi a doporučení dalšího možného využití dosažených výsledků.

# 1 INFORMAČNÍ SYSTÉMY A DATA

Ačkoliv pojem *informačních systémů* nelze zcela přesně vymezit, prostupují v současné době téměř všemi aspekty lidského bytí. Z velmi obecného pohledu jsou úzce či volně propojenými množinami lidí, procesů a dalších prostředků, pracujících s ekonomickými abstrakcemi fyzických vzorů v kontextově omezeném rozsahu oblasti zájmu své nosné domény. Jejich softwarová podoba je často nasazována v prostředích se složitými úlohami, které je nepříjemné, obtížné, nebo zcela nemožné zpracovat bez použití počítačů [35].

## 1.1 Využití informačních systémů

Informační systémy je možné rozdělit podle mnoha kritérií do různých kategorií, které ale nejsou nikdy definitivně uzavřené a navzájem se překrývají. Mohou jimi obecně být transakční systémy, automatizující zpracování úloh zabývajících se evidencí, rezervacemi či skladem, manažerské informační systémy určené primárně pro zpřístupnění přehledů v probíhajících konkrétních procesech, a usnadňující tak práci řídicím pracovníkům, systémy pro podporu rozhodování, jejichž úkolem je provádění analýz s grafickými výstupy, nebo informační systémy pro vrcholové řízení, poskytující důležité informace vrcholovému managementu dané oblasti nasazení, zvyšující tak efektivitu strategického řízení organizací [34].

Úlohou informačních systémů je podpořit úspěšnost v procesech probíhajících v oblastech jeho nasazení. Ideální informační systémy leží na průniku hodnoty poskytovaných informací a nákladů nutných pro jejich zpracování.

Mnohé informační systémy jsou primárně účelově orientovány pouze na úzce vymezenou oblast, danou aktuálním rozsahem zpracovávaných úloh, časovým ohraničením zpracování úloh, předepsaným legislativním rámcem, předáváním pouze souhrnných přehledových sestav a dalšími faktory. Ačkoliv mohou uchovávat – a ve většině případů uchovávají – data z minulých období, resp. jiných kontextových rámců, už jen výjimečně nabízí IS přímo ve svých rozhraních možnost širšího, historického pohledu napříč svou databází, umožňujícího uživatelům či managementu sledovat potenciálně důležité dlouhodobé ukazatele, plynoucí z využívání daného IS.

Významnou roli ve využití IS, a tedy v důsledku i v celkovém množství nashromážděných dat, hraje i osobní přístup jeho koncových uživatelů. Interakce s informačním systémem může být z mnoha důvodů omezená a uživatel nemusí důsledně

používat všechny jím nabízené možnosti, ať už je omezení dané povahou zpracovávané úlohy, nebo nevhodným návrhem uživatelského rozhraní, které ne zcela přehledně vystihuje význam informací, reflektovaných do vznikajících dat. To se týká především generického software, který si většinou klade v rámci své prodejnosti za cíl poskytnout možnosti pro co největší množství různých uživatelů, bez ohledu na jejich reálné využití v konkrétních aplikacích. Problematikou výstavby a životního cyklu informačních systémů se zabývá softwarové inženýrství [35].

## 1.2 Analýza dat

Cílem analýzy dat je získání konečného přehledu o informacích schraňovaných v analyzovaném informačním systému. Jejím výsledkem by mělo být shrnutí a interpretace vyvozených závěrů v kontextu prostředí, ve kterém je tento IS nasazen. Ačkoliv neexistuje žádná univerzální, přesto konkrétní metodika vedoucí k této analýze, mělo by k ní být přistupováno s dodržením alespoň obecných zásad.

### 1.2.1 Znalosti, informace a data

Předem je třeba vymezit si některé pojmy, se kterými informační systémy pracují, nebo které jsou jejich součástí. Následující hierarchie bývá často uváděna v opačném pořadí, avšak pro účely dekompozice, která je nezbytnou pro zodpovědný přístup k analýze, jsou uvedeny od systémově nejabstraktnějších po ty nejkonkrétnější.

**Znalosti** jsou souhrnem souvisejících poznatků a nabytých vědomostí o určitém prostředí a jeho zákonitostech, získané nejčastěji praktickými zkušenostmi nebo jako výsledek řízeného, opakovaného pozorování. Díky znalostem je možné předpokládat budoucí stavy prostředí, nebo též účinně ovlivňovat dění v něm.

**Informace** jsou strukturovanou, vnímatelnou a srozumitelnou zprávou o pozorovaném prostředí a jevech v něm. Znalost vzniká z informací procesem poznání, jehož úkolem je získat z konečné množiny informací jejich strukturovaným shrnutím odpovídající kognitivní model.

**Data** jsou už velmi konkrétním obrazem stavu pozorovaného prostředí, zaznamenaným pouze symboly nebo čísly. Díky tomu jsou ale velmi vhodná k přenosu a technickému zpracování, ovšem jako taková, mají data pouze explicitní lexikální význam na úrovni formálního jazyka. Vztah dat ke skutečným informacím lze tedy obsáhnout Platónovou tezí o jednotlivých elementech, které jsou osamoceně nepoznatelné,

ale jejich komplexy již poznatelné jsou [48]. Informace tedy vznikají z dat neoddělitelným sepnutím přiřazeného významu na základě znalostí nebo zkušeností.

### 1.2.2 Identifikace prostředí a inventarizace systému

Analýzu dat konkrétního informačního systému lze provádět pouze za předpokladu, že je alespoň v jeho rozsahu identifikováno a pochopeno chování prostředí, z něž vychází jím zpracovávané modely. Teprve poté je možné odpovědně a metodicky přistoupit k informačnímu systému samotnému.

#### Identifikace prostředí

Před přístupem k datům v samotné softwarové implementaci informačního systému je zcela nezbytně nutné znát celkovou oblast jeho nasazení a její specifický kontext, tzv. znalostní doménu, nesoucí konkrétní fyzické vzory, které jsou informačním systémem modelovány.

Výchozí specifická oblast může být často velice rozsáhlá a mít velké množství navzájem oddělených podoblastí zájmů s vlastními omezenými kontexty. Získání alespoň částečného rozhledu tohoto dělení, možných legislativních rámců vymezujících dění v jednotlivých podoblastech, studium historie oblasti v její celistvosti, případně poznání konkrétní praxe v některém z oborů oblasti, to vše významnou měrou přispívá možnosti získat co nejširší interpretaci výsledků prováděné analýzy.

Hlubší rozlišení základních a dodatečných dokumentů a jejich povahy v rámci konkrétních kontextově-orientovaných procesů oproti tomu umožní získat přehled o možnostech podoby jim odpovídajících datových modelů. Orientace ve specifické oblasti nasazení IS musí jednoznačně odpovídat na otázku, *jaká* data jsou evidována a zpracovávána.

#### Inventarizace systému

Proces poznávání informačního systému by měl být založen již na výše zmíněné identifikaci prostředí s ohledem na porovnání dostupných systémů určených pro analýzou zpracovávanou oblast.

Následná inventarizace, shromáždění všech možných poznatků o provozu a používání již konkrétního analyzovaného systému [49], jednoznačně stanoví rozsah a podobu zpracovávaných dat – orientace v cílovém informačním systému usnadňuje pochopit, *jak* jsou data evidována a zpracovávána. Do těchto poznatků je třeba

zahrnout i chování uživatelů, vyplývající z definice prostředí a zvolené specifické oblasti, které může mít silný vliv na způsob, jakým je konkrétní IS na konkrétním místě využíván. Kompletní poznání analyzovaného informačního systému je tedy vitální, ne však postačující okolností k provedení celkové analýzy.

### 1.2.3 Extrakce, selekce, agregace a konsolidace dat

Po shromáždění technických informací o IS následují fáze získání, výběru, čištění a celkové úpravy dat do podoby, se kterou bude výhodné pracovat při analýze.

Extrakce, neboli získání dat, označuje úkon vyjmutí dat z podoby uchovávané informačním systémem. Nejčastější podobou uložení dat v informačních systémech jsou relační databáze, jednotlivé datové soubory, případně kombinace obou uvedených možností.

Je-li to možné zařídit už během extrakce, je vhodné provést základní transformace a čištění dat. Cílem čištění je vypořádání se s chybějícími daty nebo odstraněním šumu [53], který by mohl způsobovat určitou nekonzistenci dat. Základními transformacemi jsou myšleny operace sjednocující například různé znakové sady uložených textových řetězců na univerzální, jednotnou (např. UTF-8), se kterou budou pracovat i analytické nástroje.

Agregace, tedy seskupení, je nezbytné v případě, že různé datové části pocházejí z různých zdrojů a bylo by žádoucí, kdyby byly uskupeny na jednom společném místě. Klíčem pro agregaci mohou být různé typy datových tabulek pocházející však z jednoho sledovaného období nebo jiného rámce, který bude následně analyzován jako jeden celek.

Jednotný formát by měl být dodržen napříč všemi získanými daty. Pokud je to možné zajistit, výsledná podoba pro analýzu shromážděných dat by měla být konsolidovanou relační databází, splňující požadované stupně normálních forem, v případě velkého množství nashromážděných dat pak datovým skladem sestaveným podle technik OLAP.

### Normální formy

Vznik relačních databází se datuje do 70. let 20. století, jako následek specifikace relačního modelu, který byl odpovědí na některé nedostatky dosavadních databázových, síťových a hierarchických modelů, související s realizací vazeb typu  $M : N$ . Schéma relačních databází je zajištěno technikami primárních a cizích klíčů, které

jednoznačně identifikují, resp. odkazují na jednoznačně identifikovaný záznam v odpovídající tabulce. Problematiku správného návrhu relační databáze definují *normální formy*, stanovující některá omezení a funkční závislosti sloupců v tabulkách.

Předpokladem pro **první normální formu** je atomicita všech údajů v tabulce. Lze-li některé údaje dále dělit na samostatné části, tabulka nespĺňuje první normální formu.

**Druhá normální forma** specifikuje tabulku první normální formy, ve které jsou všechny atributy přímo závislé na unikátním primárním klíči. Problematika druhé normální formy se obvykle týká tabulek se složenými primárními klíči z více sloupců, kde ostatní sloupce nemusejí mít závislost na celém složeném primárním klíči, ale pouze na některé z jeho částí.

Obdobná pravidla platí i pro **třetí normální formu** tabulek, jež zakazuje tranzitivní závislost sloupců, které nejsou primárním klíčem, na sloupcích, které na primárním klíči závislé jsou.

Ve **čtvrté normální formě** se nachází tabulka třetí normální formy právě tehdy, pokud každý sloupec v ní obsažený, který není primárním klíčem, popisuje pouze jediný jev, bez netriviální závislosti na jiném sloupci. Takováto nepovolená závislost je pouze implicitní a není funkčního typu.

Poslední, **pátá normální forma**, definuje tabulku čtvrté normální formy, ve které by při přidání jakéhokoliv nového sloupce muselo zákonitě dojít k dekompozici relačního schématu.

#### 1.2.4 Dolování dat

Termín *dolování dat* označuje souhrn vysoce specializovaných, netriviálních metod pro získávání implicitních, ale dříve neznámých znalostí z rozsáhlých databází [28]. Jádrem jejich základního principu staví na učení předsokratovské filosofie Démokrita či Anaxagora, vycházejícího z poznání, že pouze skrze fenomény lze spatřit to, co je nezjevné [47] – jejich cílem je aplikace konkrétních algoritmů pro detekci a extrakci opakujících se či vyčnívajících vzorů pozorovaných v datech, vedoucích k vytvoření statistických modelů dat. Po nalezení těchto vzorů je ale nutné provést zhodnocení jejich celkové užitečnosti [53].

Hlavní motivací k provádění dolování je v dnešní době obrovské množství sesbíraných dat z různých každodenních činností, zahrnujících nákupy, bankovní transakce, či změny obsahu webových stránek na Internetu. Současně také roste množství aplikací, které vyžadují zpracování dat co nejrychleji, ať už se jedná o data získaná z množství různých senzorů, nebo o vědecké simulace [28]. Dochází pak ke snadnému zahlcení daty, ale bez jejich analýzy a interpretace chybí tolik potřebné informace a znalosti [53]. Kvůli samotné povaze takto získaných dat je však zřejmé, že jejich analýza nemůže být prováděna tradičními analytickými metodami, ať už by byly vykonávány člověkem, nebo plně automatizovány [53, 28]. Metody dolování dat se tedy snaží vzniklou mezeru v nedostatku analytických prostředků zaplnit díky své mezioborové povaze, beroucí si inspiraci ze statistiky, umělé inteligence, databázových systémů a dalších.

Potenciál skrytý v možnostech dolování dat má nezměrný význam pro mnoho oblastí lidské činnosti, zejména v plánování marketingových strategií pro segmentaci trhu, formulování nových hypotéz v oblasti bioinformatiky [53], při plošném monitorování rozličných lidských aktivit, mající za úkol odhalit činnost zločinců a teroristů, nebo pro podporu obchodního rozhodování velkých firem. Velké firmy, vlastníci mnoho poboček a odvětví na různých geograficky oddělených místech, mají často vlastní centralizované datové sklady, schraňující všechna data z poboček za určitou dobu. Datové sklady jsou nejčastěji modelovány jako vícerozměrné databázové struktury, kde každý rozměr odpovídá sledovanému atributu a každá buňka obsahuje odpovídající agregované údaje [53].

Datové sklady velkých firem implicitně obsahují velmi cenné údaje a soudobí analytici, zabývající se dolováním dat, mohou být tedy v určitém smyslu s nadsázkou přirovnáni k horníkům ve skutečných diamantových dolech.

### **1.2.5 Zpracování získaných dat a interpretace výsledků**

Konečný výběr dat, nebo oblasti výběru dat, musí především obsáhnout prvky, které jsou relevantní a nezbytné pro účely dané analytické úlohy.

Výběr vhodného nástroje pro automatizovaná zpracování sledovaných ukazatelů v datech je plně závislý na typu zpracovávaných dat, příp. na charakteru datového typu, ve kterém se vybraný datový soubor nachází.

Číselné typy mohou být snadno zpracovány tabulkovým procesorem, nebo specializovaným statistickým softwarem, jejich vizualizace a grafy musí vhodně zvoleny



vzhledem k povaze skutečnosti, kterou číselné typy zachycují. Daleko složitější je zpracování textových řetězců, které se již neobejde bez použití znalostí získaných studiem oblasti, ve které IS pracuje. Na základě charakteru textových řetězců lze pak určit možnosti jejich zpracování, například po zběžném výpisu náhodného výběru z nich provést specifikaci požadovaných regulárních výrazů rozklad do tříd ekvivalence a s takto získanými daty dále pracovat.

Interpretaci analýzou získaných výsledků a informací je vždy nutné provádět s ohledem na specifické prostředí, ve kterém mohou nabývat konkrétního významu. Transformace získaných informací na znalosti je pak syntézou porozumění prostředí a způsobům, jakými jednotliví uživatelé s konkrétním informačním systémem pracují. Neumožňuje-li dostatečná znalost prostředí výsledky interpretovat, lze doporučit konzultaci s odborníkem a následně zvolit další postupy.

## 2 SPECIFICKÁ OBLAST - ŠKOLSTVÍ

Vzdělávání je jedním z nejdůležitějších pilířů kulturně vyspělé společnosti. Jeho hlavním a neodmyslitelným poskytovatelem a nositelem jsou školy, potažmo celá školská soustava, která, stejně jako celá společnost, musí v kontextu národní kultury odrážet potřeby moderního člověka, a být tedy založena na permanentních inovacích [33]. Ke konci minulého století se s raketovým nástupem výpočetní techniky začaly uplatňovat informační technologie pro usnadnění některých administrativních činností i ve školských zařízeních, přičemž se nejprve jednalo především o správu financí a o základní nástroje pro záznam výsledků žáků či studentů dané školy. Kvůli absenci hlubší integrace s jinými softwarovými nástroji bylo však v mnoha případech nutné zadávat data do programů opakovaně, přestože již byla zaznamenána na jiných místech [21].

### 2.1 Vzdelávání jako výrobní proces

Jelikož koncem 90. let započal pomocí poslanecké novely zákona o státní správě a samosprávě ve školství (č. 284/2002 Sb.) [3] proces přeměny, ve kterém vzdělávací instituce získaly vlastní právní subjektivitu [51], bylo najednou kromě samostatného rozhodování o finančních a rozpočtových záležitostech také nutné celkově řídit i vnitřní procesy a koordinovat tak celkovou činnost školy [54]. Vznikl tak prostor pro nezbytnou manažerskou podporu fungování škol, které se právě díky svým nově nabytým možnostem rázem dostaly do konkurenčního prostředí s vlastní odpovědností za přípravu budoucích odborníků nezbytných pro rozvoj kvartérního sektoru ekonomiky národního hospodářství. Management byl donucen naučit se rychle, pružně a efektivně v tomto prostředí jednat a vnímat svoji školu a její potenciál z čistě tržního pohledu, zahrnujícího i takové aspekty, jako například soutěžení škol o žáky nebo srovnávání výsledků a hodnocení škol [21]. Další prostor pro svůj rozvoj, sebeurčení a především lepší konkurenceschopnost [55] prostřednictvím ŠVP<sup>1</sup> získaly školy spolu s dalšími povinnostmi po kurikulární reformě na základě zákona č. 561/2004 Sb. [4]

Vhodné nasazení a využívání informačního systému v prostředí školy může hrát významnou roli právě pro usnadnění všech požadovaných administrativních činností, zároveň i pro podporu rozhodování managementu pomocí průběžného sběru výsledků a analýzy jejich rozdílů například mezi jednotlivými třídami nebo skupinami

---

<sup>1</sup>ŠVP (Školní vzdělávací program) je dokument vycházející z RVP (Rámcového vzdělávacího programu), umožňující škole se unikátně profilovat a odlišit se tak od ostatních škol

žáků. Též může být využit pro efektivní správu různých dokumentů a rovněž i jako rychlý a přehledný komunikační prostředek, ať už se jedná o předávání informací uvnitř školy, předávání informací státním úřadům, nebo o komunikaci školy s žáky, jejich rodiči či zákonnými zástupci a veřejností.

## 2.2 Informační systémy ve školství

V kontextu školství lze využívání informačních systémů chápat jako zcela specifickou oblast využití manažerských informačních systémů [21]. Je možné je definovat jako soubor lidí, metod a technických prostředků zajišťujících sběr, uchování, analýzu a prezentaci dat určených pro poskytování informací v oblasti vzdělávání [26].

S výjimkou určitých případů<sup>2</sup> nemají základní a střední školy zpravidla dostatek finančních, personálních a technologických prostředků k vývoji (internímu či externímu) vlastního informačního systému na míru pro svoje konkrétní potřeby, proto spíše volí nákup typových informačních systémů s možnostmi nastavení detailů pro co nejhodnější pokrytí těchto potřeb. Při jejich výběru je třeba pečlivě přihlížet k mnoha kritériím [45], počínaje deklarovanou funkcionalitou, požadovaným provozním prostředím, až po legislativní podmínky a poskytované služby.

V současné době lze rozdělit využití informačních systémů základními a středními školami do následujících hlavních oblastí:

- Provozně-hospodářský úsek
- Intranet, e-learning a interaktivní výuka
- Prezentace školy a komunikace s veřejností
- Školní agenda

### 2.2.1 Provozně-hospodářský úsek

Provozně-hospodářský, případně ekonomický úsek školy má za úkol spravovat finance a majetek školní organizace. Je odpovědný za dodržování rozpočtové kázně, inventury, vedení účetnictví, zpracování prvotních účetních dokladů a výkazů a zpracování mezd zaměstnanců. Vzhledem k povaze škol, nastíněné v úvodu této kapitoly, v tomto ohledu prakticky neexistuje významný rozdíl mezi školami a jinými právními subjekty. Povinnosti vedení účetnictví zachycujícího veškerou hospodářskou činnost

---

<sup>2</sup>například odborné školy zaměřené na výuku IT

vyplývají z legislativního rámce tvořeného především zákonem č. 563/1991 Sb. [1] a dalšími předpisy.

Na trhu je k dispozici velké množství softwarových produktů pro vedení účetnictví dle příslušného legislativního rámce, patří mezi ně například produkty společnosti Vema a.s., systémy HELIOS od Asseco Solutions, a.s., Pohoda, Money a mnoho dalších.

Pro úseky využívající ekonomické systémy obecně platí přísnější bezpečnostní pravidla a obvykle nebývá žádoucí, aby byly propojeny s jinými informačními systémy školy. Počítače využívané pro účetnictví bývají velmi často i zcela odděleny od školní sítě a disponují záložními nepřerušitelnými zdroji napájení.

### 2.2.2 Intranet, e-learning a interaktivní výuka

*Intranet* je v informatice označením služeb běžných internetových protokolů používaných pouze v rámci lokální (příp. virtuální privátní) sítě. Školy bývají uživateli malých a středně velkých sítí s desítkami až stovkami počítačů. Intranet nejčastěji zahrnuje vnitřní komunikační systémy a služby pro sdílení dokumentů a znalostí. Poskytují je například souborové, tiskové, poštovní a aplikační servery, které svou centrálně-orientovanou správou značně zjednodušují proces pro správu a rozšiřování počítačové sítě [38].

Pojem *e-learning* v sobě zahrnuje metody pro poskytování obsahu výuky prostřednictvím elektronických médií. Poskytuje studentům jeden nebo více on-line kurzů s učebními texty, příklady pro demonstraci probíraného učiva, které mohou být vhodně doplněny multimediálním obsahem, a také plně nebo částečně interaktivní testy získaných znalostí. Kurzy mohou být rovněž doplněny o diskusní fóra pro sdílení připomínek, námětů a dotazů. Přestože elektronickým médiem může být téměř jakýkoliv nosič, systémy e-learningu jsou nejčastěji realizovány jako komplexní webové<sup>3</sup> aplikace s nástroji LMS, dostupné buď pouze z vnitřní, nebo i vnější části sítě školy. Jedním z konkrétních příkladů e-learningového informačního systému může být otevřená<sup>4</sup> platforma Moodle.

Interaktivní výuka představuje progresivní metody učení s využitím potenciálu nástrojů ICT. Mezi ně spadá především používání dataprojektorů a interaktivních

---

<sup>3</sup>pro běh aplikace je vyžadován webový server, klient k ní přistupuje přes webový prohlížeč

<sup>4</sup>vyvíjená jako komunitní open-source software pod licencí GPL

tabulí, kde vyučující svým výkladem doprovází obsah promítané elektronické učebnice, která může být obohacena o multimediální ukázky a elektronické pracovní listy [46]. Prostřednictvím dotekových tabulí lze doplňovat poznámky a kolektivně zpracovávat cvičení. Smyslem interaktivní výuky je především probuzení zájmu žáků a zvýšení atraktivity vyučování v Komenského smyslu *školy hrou* [33], a to prostřednictvím moderních technologií. Elektronické učebnice jsou samy o sobě specializovaným softwarem, který může být provázán se školním e-learningovým prostředím. Na jejich správné používání a využívání při výuce ostatními pedagogy dohlíží odpovědný metodik vzdělávání.

### 2.2.3 Prezentace školy a komunikace s veřejností

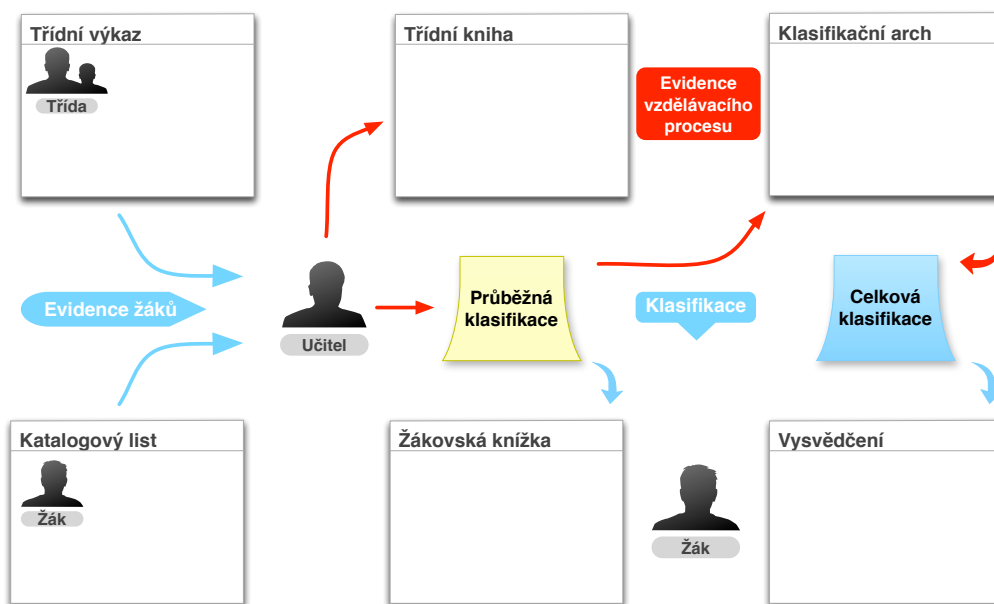
Moderní škola se dnes již neobejde bez dobré prezentace na Internetu, která je nezbytnou součástí její reprezentační strategie. Webové stránky mají za úkol informovat návštěvníky zejména o charakteru a poslání školy, nesmí chybět její přesný název a kontaktní údaje. Národní ústav pro vzdělávání se mezi svými doporučenými nástroji pro autoevaluaci škol věnuje i posuzování jejich internetových prezentací, kde je při hodnocení obsahu mimo jiné zohledňována i dostupnost kompletního znění ŠVP, výročních a inspekčních zpráv, harmonogram aktuálního školního roku, nabízená možnost komunikace veřejnosti se školou a další aspekty [36]. S technologickým rozvojem a dostupností hotových řešení bývají stránky škol často realizovány jako CMS, kde se na tvorbě jejich obsahu může podílet celý pedagogický sbor, informovat formou aktualit o dění ve škole, nabízet žákům rozšiřující studijní materiály a domácí úkoly a mnoho dalšího. Takovéto webové prezentace mohou být propojeny i s jinými informačními systémy a poskytovat jejich obsah. Například, nabízí-li daná škola stravování, může být zobrazován aktuální jídelníček, nebo kalendář připravovaných akcí.

Poměrně novým fenoménem pro komunikaci s veřejností bývá prezentace školy na sociálních sítích, kde je možné podávat rozšiřující nebo i neformální oznámení, rozesílat pozvánky na kulturní události pořádané školou přesné cílové skupině, publikovat reportáže v přímém přenosu apod. Protože se o řízení uživatelů z řad veřejnosti stará daná sociální platforma, nenarůstá škole nežádoucí nadbytečná režie na údržbu takovéto prezentace.

## 2.2.4 Informační systémy školní agendy

Jelikož je výchovně-vzdělávací činnost školy jejím základním komplexním procesem, jsou MIS školní agendy určitou analogií ERP systémů v jejich odpovídajícím ryze komerčním prostředí. K jejich rapidnímu nárůstu však nejvíce přispěla povinnost předávat údaje ze školní matriky pro MŠMT<sup>5</sup> v elektronické podobě, stanovená vyhláškou č. 364/2005 Sb. ve znění pozdějších předpisů [45, 5]. Rozsah dat, která škola musí povinně evidovat, je stanoven zákonem č. 561/2004 Sb. (a jeho novelou č. 472/2011 Sb.) a č. 364/2005 Sb. [4, 8, 5], přičemž musí být zachovány postupy podle zvláštního právního předpisu č. 101/2000 Sb. [2]

Kromě základního kamene administrativní agendy školní matriky, zabývající se evidencí žáků, tvoří významnou část další povinná a nepovinná pedagogická dokumentace. Pedagogickou dokumentací jsou téměř všechny úřední dokumenty spojené s činností školy, které musí být vedeny s náležitou péčí a opatřeny čitelným podpisem. Vztah základních dokumentů popisuje následující diagram.



Obr. 2.1: Diagram – vztah základních pedagogických dokumentů. Modře je vyznačen tok informací, červeně jejich tvorba.

<sup>5</sup>do 31.12.2011 byla data odevzdávána Ústavu pro informace ve vzdělávání, poté jeho agendy přebralo MŠMT

**Třídní kniha** zaznamenává průkazné údaje o poskytovaném vzdělávání a jeho průběhu. Přestože forma tohoto dokumentu není legislativně stanovena a je plně v kompetenci školy, povinné informace jsou jasně definovány v příslušných zákonech. V třídní knize musí být odpovědným učitelem zapsána každá rozvrhem daná hodina včetně stručné charakteristiky probíraného učiva daného předmětu a případné absence žáků. Za správnost údajů zapsaných v třídní knize odpovídá třídní učitel.

**Třídní výkaz** obsahuje jmenný seznam žáků třídy a při jeho změně je třídním učitelem aktualizován. Nese souhrnný přehled o počtech žáků, o jejich docházce, chování, klasifikaci a celkovém hodnocení, počtu absolvovaných let školní docházky a o přestupech z jiných školských zařízení nebo jiném způsobu vzdělávání. Třídní výkaz je zvlášť veden pro každý stupeň vzdělávání.

**Katalogový list** je základním dokumentem pro vedení školní matriky a uchovává všechny důležité skutečnosti o žákovi. Kromě běžných osobních informací o jméně, bydlišti, datu narození, rodném čísle, případně údajích o čísle občanského průkazu nebo pasu, a údajů souvisejících se studijním oborem, jsou v něm uvedeny informace o rodičích a zákonných zástupcích žáka včetně jejich kontaktních údajů. Dále kompletní katalogový list také udržuje veškerá výchovná opatření a zprávy pro ostatní úřady. Především jako přehledový nástroj pro třídního učitele či zastupujícího třídního učitele jsou zde evidovány informace o zdravotním a sociálním stavu žáka. Za úplnost a správnost údajů uvedených na katalogovém listě žáka odpovídá třídní učitel. Při přestěhování žáka a jeho přestupu na jinou školu se této škole zasílá pouze kopie katalogového listu, jeho originál zůstává na původní škole.

Třídní knihy, třídní výkazy a jednotlivé katalogové listy podléhají povinné archivaci po dobu 45 let.

**Průběžná klasifikace** je systematickým hodnocením dílčích výsledků a projevů žáka v jednotlivých vyučovacích předmětech a na čtvrtletních klasifikačních poradách v rámci výchovně-vzdělávacího procesu. Na průběžné klasifikaci se podílí učitel příslušného předmětu a žáků v souladu s vnitřním klasifikačním řádem školy.

**Klasifikační arch** slouží učiteli jako pomůcka pro přehled průběžné klasifikace v daném vyučovacím předmětu a třídě. Je nepovinným dokumentem usnadňujícím evidenci hodnocení žáků a nemá stanovenou podobu.

**Žákovská knížka** je nepovinným dokumentem zaznamenávajícím průběžné výsledky a hodnocení chování žáka a slouží zejména jako prostředek pro předávání informací školy zákonnému zástupci žáka [44].

**Celková klasifikace** a celkové hodnocení žáka je souhrnem hodnocení žáka na konci klasifikačního období. Stupeň hodnocení primárně nevychází z průměru odpovídající průběžné klasifikace, učitel musí přihlížet k průběžným výsledkům práce žáka.

**Vysvědčení** je úředním dokumentem potvrzujícím výsledky vzdělávání žáka na dané škole. Hodnocení žáka může být zapsáno slovně, nebo odpovídajícím klasifikačním stupněm v souladu se zákonem č. 561/2004 Sb. [4] Výsledný dokument je tiskopisem SEVT opatřeným vodoznakem, kulatým razítkem školy a podpisem ředitele školy a třídního učitele.

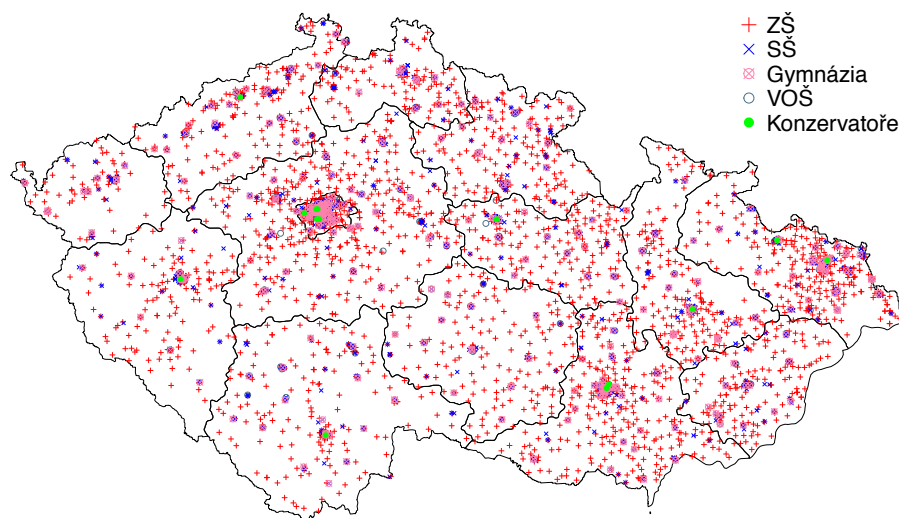
Vzhledem k velkému množství a náročnosti procesů spojených s celkovou administrativou je tato oblast klíčovým kandidátem pro zavedení vhodného informačního systému, který by umožnil celkovou práci nejen zefektivnit, ale poskytl by i souhrnné analytické informace sloužící jako podklady pro pravidelnou autoevaluaci – zpětnou vazbu ke korekci vlastní činnosti [52], případně dovolil nahlédnout na souhrnné informace získané za celou dobu vedení digitalizované pedagogické dokumentace.



Tato situace se týká v následující tabulce uvedeného počtu škol. Pro porovnání jsou kromě údajů z rejstříku škol a školských zařízení a výkonových ukazatelů statistické ročenky MŠMT pro rok 2015 uvedeny i údaje evidované soukromými subjekty, jmenovitě firmou just4web.cz s.r.o.<sup>6</sup>, která je vývojářem školního informačního systému *Etrádnice*, a P.F. art, spol. s r.o.<sup>7</sup>, orientující se na výrobu a distribuci učebních pomůcek pro školy v ČR.

Tab. 2.1: Celkové počty škol v ČR ke šk. roku 2014/2015 dle uvedených zdrojů

	MŠMT [12]	Seznam škol v ČR [11]	AtlasŠkolství.cz [14]
Základní školy	4 106 (2 707 2. stupeň)	4 151	4 136
Střední školy <sup>8</sup>	1 310	1 490	1 295
VOŠ	174	188	176
Celkem	5 590	5 829	5 607



Obr. 2.2: Mapa – Geografická distribuce škol v České republice<sup>9</sup>

<sup>6</sup><http://www.just4web.cz>

<sup>7</sup><http://www.pfart.cz>

<sup>8</sup>zahrnuty jsou střední školy, střední odborná učiliště, gymnázia a konzervatoře

<sup>9</sup>graficky je zachyceno pouze 82,09 % škol

## Obecná charakteristika dostupných systémů

Školní informační systémy zpracovávající povinná data školní matriky prošly za řadu let své existence významným vývojem, který umožnil obohatit jejich základní funkce o mnoho užitečných rozšíření umožňujících pokrýt téměř celou školní agendu [45]. Vžilo se pro ně souhrnné označení *programy pro školní administrativu*. Základní ideou bylo maximální využití centrálně řízené správy evidence žáků, zaměstnanců a jiných zdrojů školy, které zásadně snížilo množství administrativních úkonů spojených se zaváděním záznamů do docházkových systémů, evidence výpůjček knihovny, správy strážníků školní jídelny, nebo i organizaci suplování či přijímacích a závěrečných zkoušek, tvorbou rozvrhů a v mnoha dalších činnostech.

Z nabídky produktů dostupných na trhu si po technické stránce mohou základní a střední školy vybrat v zásadě ze dvou možných typů informačních systémů.

Prvním z nich jsou informační systémy založené na webových technologiích hostované v prostředí Internetu [24, 26]. Jedná se o cloudová řešení postavená na obchodním konceptu SaaS, kdy škole stačí pouze podepsat příslušnou smlouvu s dodavatelem [24] a zadat svá data. Tím se významně snižují personální a finanční náklady, jelikož se škola již nemusí zabývat technickými detaily zahrnujícími zajišťování provozu (hardware a síťovou infrastrukturu zajišťuje dodavatel ve svém hostingovém centru), nebo otázkami multiplatformní dostupnosti (k systému lze přistupovat prostřednictvím téměř libovolné platformy a webového prohlížeče). Mezi zástupce takovýchto produktů patří například Škola OnLine (ŠKOLA ONLINE a.s.), iŠkola (Computer Media s.r.o.), Etřídnice (just4web.cz s.r.o.) a další.

Druhým, historicky starším typem, jsou systémy provozované přímo v prostředí školy. Mají charakter modulárně uspořádaného uceleného souboru kooperujících programů využívajících společné databázové prostředí. Záznam a zpracování dat lze provádět na jednom či více počítačích v počítačové síti školy [24]. Provoz a chod programů si škola zajišťuje sama, z čehož pro ni vyplývají povinnosti zajistit potřebnou infrastrukturu, dostupnost a zabezpečení. Předplatitelům jsou pravidelně nabízeny nové verze vycházející z požadavků na potřeby školství. Typickými zástupci programů tohoto typu jsou robustní systémy Relax-KEŠ (Alis s.r.o.), SAS (MP-Soft, a.s.) a Bakaláři (Bakaláři software s.r.o.), pokrývající potřeby mnoha druhů základních, středních a vyšších odborných škol, gymnázií, konzervatoří a odborných učilišť.

Během výběru IS pro správu školní agendy se musí škola zamyslet nad mnoha klíčovými kritérii, mezi která patří jeho uživatelská podpora a školení, rozšiřitelnost

o další funkce, aktuálnost z hlediska legislativních požadavků, možnost si IS vyzkoušet před nasazením, možnost vzdáleného přístupu k datům a mnoha dalším [45, 35]. Při nasazování školního informačního systému je také neméně nutné zajistit jeho pozitivní přijetí [26, 35] členy pedagogického sboru, kteří mohou v určitých případech vykazovat neochotu [7, 32] se učit jeho používání či práci s moderními technologiemi.

## 3 INFORMAČNÍ SYSTÉM BAKALÁŘI

Programy Bakaláři jsou jedním z nejdéle používaných systémů pro správu školní agendy. Uvádí se, že patří k nejrozšířenějšímu software svého druhu v českých školách [24, 27, 30]. Jejich snahou je pokrytí co největší oblasti školní administrativy pro rozličné typy škol pomocí přizpůsobitelných modulů pro evidenci žáků a zaměstnanců, přijímací zkoušky, grafické zpracování klasifikace, rozpis maturit, suplování, hospitace, evidence úrazů, tematické plány, generátor rozvrhu hodin, inventarizaci majetku, knihovnu, rozpočet, plán akcí, třídní knihu a webové aplikace pro komunikaci a zobrazování údajů o klasifikaci žáků [10].

### 3.1 Rozšířenost systému

Jelikož žádný vývojář informačních systémů pro školní administrativu ve svých prezentacích přímo neuvádí svůj konkrétní podíl na trhu a přesné průzkumy v této oblasti téměř neexistují, lze se pouze pomocí extrapolace údajů omezeného, přibližně 1% reprezentačního vzorku [31] českých škol, a údajů o ICT vybavenosti škol dle výročních zpráv ČŠI<sup>1</sup> na tabulku 2.2.4 domnívat, že se celkový tržní podíl programu Bakaláři pohybuje pravděpodobně okolo  $(50,23 \pm 6,65 \cdot k) \%$ , kde  $k \geq 1$  představuje koeficient ještě dalšího možného zkreslení především v podobě poměru počtu školských stupňů nebo organizačních jednotek na jeden subjekt<sup>2</sup> se zakoupenou licencí programů.

Takto získaná výsledná aproximace přibližně koresponduje s údajem 3 500 škol [30] uváděným v oficiální základní příručce programů Bakaláři pro rok 2014.

### 3.2 Stručný přehled systému a jeho modulů

Systém Bakaláři není samostatnou monolitickou aplikací, jedná se o sadu kooperujících programů, tzv. modulů, využívajících společnou databázi. Každý z modulů rozšiřuje funkčnost systému o nové možnosti a k celkové licenci je dokoupitelný zvlášť.

---

<sup>1</sup>poslední údaje o evidenčním software jsou dostupné ve výroční zprávě ČŠI pro školní rok 2011/2012

<sup>2</sup>organizace zahrnující pod sebou více stupňů či typů škol vzdělávání, typicky víceletá gymnázia nebo sloučené SOŠ s SOU či s VOŠ

Přístup do prostředí jednotlivých modulů je zajišťován jednotným přihlašovacím dialogem, který obsahuje nabídku se jmény učitelů a poptává odpovídající heslo pro ověření totožnosti. Ke každému modulu a jeho možnostem může správce zvláště definovat přístupová oprávnění podle nastavených funkcí (ředitel školy, zástupce, ...). Samotný účet správce systému není spojen s žádným uživatelským účtem učitele a má neomezená oprávnění v celém systému.

### 3.2.1 Společné prostředí

Modul společného prostředí (v nabídce modulů také označovaný jako *Společné prostředí programů Rozvrh a Evidence*) je základním stavebním prvkem celkového informačního systému. Umožňuje definovat společná data pro všechny ostatní moduly, do kterých je během školního roku potřeba zasahovat jen zřídkka. Jedná se zejména o globální parametry školy, kterými jsou výčet budov, místností a sportovišť a všech vyučovaných předmětů. Dalšími, většinou začátkem školního roku měněnými parametry, jsou úvazky učitelů a nastavení tříd.

Společné prostředí je především odpovědné za realizaci zdánlivě netriviálních operací, které obecně představují definice různých množin. Zpracovávanými množinami mohou být seznamy místností vhodných, případně zcela nevhodných pro určitý předmět, definice skupin žáků napříč třídami (například pro společnou výuku cizích jazyků) a další možné kombinace včetně kontroly disjunktnosti, jež mohou pro celkový management provozu školy představovat značné ulehčení práce.

### 3.2.2 Evidence žáků a zaměstnanců

Sebepopisujícím názvem představuje modul *Evidence* základní seznam všech žáků a zaměstnanců rozšířený o nepřeberné množství sestav a pohledů umožňující s informacemi o žácích a zaměstnancích pracovat. Karta žáka je elektronickou alternativou katalogového listu nesoucí přehled o osobních, matričních, klasifikačních, rodinných, zdravotních a výchovných informacích o daném žákovi. Obdobně karta zaměstnance nese důležité informace o pedagogických a v některých případech i nepedagogických zaměstnancích školy, umožňuje přehledně zobrazit aktuální úvazky a další strukturované záznamy.

Modul evidence také slouží k zápisu průběžné a celkové klasifikace, záznamům o průběhu studia a zameškaných hodinách a především jako generátor sestav, kte-

rými je zejména vysvědčení a další přehledy hodnocení, výstupy pro MŠMT a údaje z matrik a jiná pedagogická dokumentace.

Modul je také používán pro manipulaci s číslicí, podporu práce na více počítačích mimo síť, aktualizaci dat zapsaných pouze ve webové aplikaci a pro definici přístupu jednotlivých skupin uživatelů k webové aplikaci. Takto rozsáhlé možnosti jediného modulu jsou známkou neodvratitelného zastarávání software [35] a snižují celkovou přehlednost uživatelského rozhraní.

### **3.2.3 Rozvrh hodin, generátor rozvrhů a suplování**

Obdobně jako u předchozího, název tohoto modulu napovídá jeho primárnímu určení. Tvorba rozvrhu je obecně, zvláště ale u škol většího rozsahu, velmi náročný proces, kdy je potřeba vzít v úvahu možné kombinace místností, časů pedagogů a výměny a přesuny vyučovacích hodin [30]. Rozvrh také může být generován s ohledem na různě stanovené týdenní nebo měsíční cykly a v kombinaci s modulem pro suplování umožní najít vhodné učitele k okamžitému zastoupení nebo prostor k výměně či nahrazení vyučovací hodiny. Generátor rozvrhu však počítá pouze s celoškolským homogenním začátkem a koncem časů vyučovacích hodin.

Dříve staticky generované rozvrhy a přehledy suplování ve formátu HTML jsou v novějších verzích nahrazeny dynamickým generátorem spadajícím pod webovou aplikaci.

### **3.2.4 Třídní kniha**

Třídní kniha si klade za cíl nahradit papírové třídní knihy [30] a zefektivnit proces evidence vzdělávacího procesu. Třídní kniha využívá dat poskytovaných generátorem rozvrhu hodin a suplování, čímž zrychluje samotný zápis hodiny učitelem předpřipravením příslušných položek. Výhodou elektronické třídní knihy je automatický výpočet absencí žáků [30], usnadňující další práci zejména třídním učitelům.

Oproti většině ostatních modulů má tento program v jednotném přihlašovacím dialogu ve výchozím stavu otevřenou grafickou klávesnici, indikující záměr přizpůsobit uživatelské prostředí a celkové používání programu na zařízení s dotekovým displejem.

### 3.2.5 Ostatní moduly a možnosti

Mezi další významné možnosti by bylo možné zařadit doplňky využívající modul *Bakalář*, vytvářející obecné prostředí pro jakoukoliv evidenci. Z placených doplňků jsou to aplikace pro zápis do 1. ročníku středních škol, *Knihovna*, s evidencí knih a výpůjček a možnosti načítání bibliografických informací o titulech z Národní knihovny ČR [30], evidenční program *Inventarizace* pro přehled o majetku a vybavení školy a shareware doplňky pro evidenci úrazů, hospitací a dalších<sup>3</sup>.

#### Programy pro správu a údržbu

Velmi důležitou součástí systému Bakaláři jsou nástroje pro správu systému samotného. Patří sem *Archivace* a *Deaktivace*, vytvářející, resp. obnovující zálohy aktuálního stavu celé databáze systému nebo jejích částí, *aktualizace z www*, která vyhledá nové verze programů a modulů a provede jejich stažení a aktualizaci z webových stránek vývojářů systému a v neposlední řadě také *Vzdálená podpora BakaTeam*, využívající program TeamViewer<sup>4</sup> k možnosti připojení pracovníka telefonické podpory na pracovní plochu uživatele, aby mohl pomoci vyřešit případné potíže s používáním programů přímo na daném počítači.

Zvláštní položkou programů pro správu jsou *Parametry webové aplikace*, sloužící, jak opět napovídá název, ke komplexnímu nastavení webové aplikace.

### 3.2.6 Webová aplikace

Webová aplikace bývá označována jako *elektronická žákovská knížka*. Umožňuje žákům a jejich zákonným zástupcům zobrazit detailní přehled o své klasifikaci prostřednictvím webového prohlížeče v prostředí Internetu. Kromě nahlížení na výsledky vzdělávání a absence poskytuje také komunikační rozhraní *Komens* sloužící k zasílání zpráv včetně omluvenek mezi žáky, rodiči a školou. Ve spolupráci s modulem *Knihovna* též umožňuje zobrazit seznam výpůjček z knihovny.

Učitelům poskytuje webová aplikace možnost zápisu klasifikace přes Internet podobně, jako desktopová aplikace modulu *Evidence*. V nabídce má i generování některých základních sestav se jmennými seznamy žáků třídy.

---

<sup>3</sup>další jsou dostupné na [http://www.bakalari.cz/sharewar\\_bak.aspx](http://www.bakalari.cz/sharewar_bak.aspx)

<sup>4</sup><https://www.teamviewer.com/cs/>

### 3.2.7 Orientační ceník modulů pro školní rok 2014/15

Následující orientační ceník převzatý z webových stránek<sup>5</sup> vývojářů systému Bakaláři ukazuje plné ceny licencí daných modulů pro školy určité velikosti (měřené v počtu žáků). Mateřské školy mají nárok na 50% slevu zakoupených programů, zdravotnické školy na 20% slevu. Každoročně prováděný automatický upgrade vyjde jakoukoliv školu na 20 % plné ceny modulů, slevy na upgrade ze starších verzí programů se úměrně snižují.

Tab. 3.1: Ceník plné licence modulů v Kč vč. DPH platný od 1. listopadu 2014 [10]

počet žáků <sup>6</sup>	do 100	do 200	do 300	do 500	do 800	do 1000
Společné prostředí	2 200	2 500	2 800	3 400	4 100	4 500
Evidence + Bakalář	2 000	3 500	5 300	8 700	12 300	14 700
Přijímací zkoušky, zápis	300	400	600	1 000	1 500	1 800
Gr. zpracování klasifikace	500	800	1 200	2 000	2 900	3 500
Rozpis maturit	600	1 000	1 500	2 500	3 600	4 300
Knihovna	600	1 100	1 700	2 700	3 900	4 700
Inventarizace	400	600	900	1 400	2 000	2 400
Rozpočet školy	300	400	600	1 000	1 500	1 800
Plán akcí	400	800	1 100	1 800	2 600	3 100
Rozvrh hodin	1 700	3 000	4 500	7 400	10 400	12 400
Suplování	700	1 200	1 800	3 000	4 300	5 200
Tematické plány	300	600	800	1 300	1 900	2 300
Třídní kniha	700	1 200	1 800	3 000	4 300	5 200
Webové aplikace	1 100	1 900	2 900	4 700	6 700	8 000

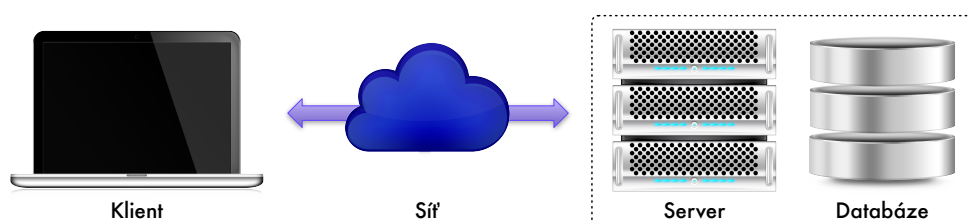
<sup>5</sup>plný ceník je k dispozici na <http://www.bakalari.cz/cenyprog.aspx>

<sup>6</sup>vybrány byly pouze některé velikosti, licence jsou k dispozici ve variantách pro školy o velikostech do 100, 200, 300, 400, 500, 600, 800, 1000, 1200, 1400, 1600, 1800 a 2000 žáků



### 3.3 Architektura programů a jejich provoz

Provozování programů Bakaláři z pohledu systémového nebo síťového správce může být stejně náročné, jako jejich plynulé používání z pohledu koncového uživatele, ale jeho organizace je velmi důležitá pro správné porozumění tokům zpracovávaných dat a struktuře jejich úložiště. Mnohá tvrzení, že programy Bakaláři pracují v rámci školní počítačové sítě na plnohodnotné architektuře typu klient–server [24], jsou však zavádějící a z neuživatelského pohledu, orientovaného právě na toky dat, jsou mýlkou.



Obr. 3.1: Diagram – obecná třívrstvá architektura typu klient–server

Plnohodnotná dvou a vícevrstvá architektura klient–server se zpravidla vyznačuje tím, že má jednoznačně oddělenou klientskou a serverovou část, komunikující navzájem po společné síti, a to v ideálním případně jediným vyhrazeným kanálem. Serverový program běžící na serverovém počítači je obvykle neustále spuštěný a vyřizuje požadavky občas spouštěných klientských programů, na kterých je jinak zcela nezávislý. Takovéto chování vychází ze samotné filosofie síťového prostředí.

Naproti tomu programy Bakaláři nebyly od svého vzniku koncipovány jako síťové a i dnes v principu představují pouze kolekci aplikací sdílejících ze svého pohledu lokální datové zdroje jediného počítače.

Instalace celého systému Bakaláři na jediném počítači v rámci celé školy představuje z hlediska nasazení, údržby a relativní bezpečnosti tu nejjednodušší možnou, ovšem v dnešní době již velmi nepravděpodobnou variantu. Vhodná může být především pro menší školy s minimem digitalizované agendy, kde jsou programy využívány jen úzkým okruhem odpovědných uživatelů či pouze vedením školy. Omezení se přitom nemusí vztahovat na jeden konkrétní počítač a souborový systém jeho pevného disku, aktivní kopie programů Bakaláři se spolu se svými daty může nacházet i na vhodném přenosném médiu. Pro umožnění práce více uživatelům je ale prakticky

nezbytné nasazení programů v síťovém prostředí. Instalační program nabízí dvě varianty síťového provozu, lišící se pouze v použitém principu ukládání dat.

### 3.3.1 Aplikační architektura programů

Aby bylo možné porozumět oběma nabízeným možnostem a jejich odpovídajícím výhodám a omezením, je nejdříve nutné objasnit aplikační architekturu programů. Jednotlivé programy systému Bakaláři byly vytvořeny v databázovém programovacím prostředí Microsoft Visual FoxPro. Jedná se o RAD nástroj předurčený k rychlému vývoji databázových desktopových aplikací, určených pro operační systémy Windows 32bitové platformy Intel x86 a primárně orientovaných především na úlohy generování sestav<sup>7</sup> [29].

Původní *FoxPro* je jako celek jazykem čtvrté generace, vycházejícím z jazyka *dBase*, pevně spjatým se stejnojmennou databázovou technologií. Její první primitivní víceuživatelská podpora práce s jednou databází byla představena v produktu *dBase II* v roce 1985 [25], dávno před akvizicí Fox Software firmou Microsoft v červnu 1992 [16]. Jako u mnohých moderních, ne pouze na souborových DSN založených (Microsoft Access, SQLite aj.), databázových systémů je problematika víceuživatelského přístupu k databázi v *dBase* ošetřena nasazením zámků.

**Zámek** je softwarová synchronizační konstrukce, informující ostatní procesy o aktuálním využívání daného zdroje. Tyto procesy by měly zámek respektovat a ke zdroji přistupovat pouze v takové míře, v jaké jim to příslušný zámek umožní.

### 3.3.2 Provoz v síťovém prostředí

Provoz systému Bakaláři v síťovém prostředí nespočívá v instalaci specializované serverové a klientské části, spoléhá na instalaci kompletní sady programů na sdílený disk souborového serveru prostřednictvím běžně dostupné technologie sdílení souborů a tiskáren systémů Windows, postaveném na obecném protokolu CIFS. Na „klientské“ počítače jsou pak kromě registrace podpůrných běhových knihoven<sup>8</sup> instalovány pouze sady ikon – zástupců – aplikací, vyžadující ke svému spuštění připojený síťový disk umístěný na souborovém serveru.

---

<sup>7</sup>mnohem výstižnějším termínem je původní anglické slovo *reports*, hlášení

<sup>8</sup>Microsoft Visual C++ Redistributable a další

Z historického hlediska bylo pro vývojáře systému Bakaláři takovéto zajištění síťové provázanosti logické a dostatečné, neboť nevyžadovalo žádné zvláštní, nebo vyžadovalo pouze minimální úpravy stávajících programů, a spoléhalo na zručnost správců příslušných škol při jejich nasazení. Před masivním rozvojem ICT ve školství, využívající dnes počítačové sítě jako běžný didaktický prostředek [38], byly také rozsáhlejší školní sítě spíše výsadou technických středních škol a gymnázií, používajících k centralizované správě uživatelů a síťových zdrojů především nástroje operačního systému Novell Netware. Tomuto systému je v nápovědě k programům Bakaláři věnována důležitá poznámka [19], více viz 3.3.3.

### **Instalace v síti s daty v DBF souborech**

První možnost síťové instalace spočívá v datech uložených v souborech formátu FoxPro na sdíleném disku CIFS serveru a jeho souborového systému. Jak již ale bylo řečeno v úvodu kapitoly 3.3, z pohledu aplikace jsou připojené síťové jednotky považovány za zdroje místního počítače a je s nimi podle toho zacházeno. V této konfiguraci hraje velkou roli samotné nastavení serveru pro sdílení souborů.

Implementace sdílení souborů v operačních systémech Windows, SMB a zejména verze SMB2 a vyšší, je optimalizována pro co nejrychlejší přístup ke sdíleným zdrojům a multimediálnímu obsahu, proto disponuje možností ukládání sdílených souborů do mezipaměti klientského počítače prostřednictvím tzv. *dobrovolného zamykání* vyšších úrovní. Jednou z možných výhod může být zobrazení obsahu vzdáleného serveru i v režimu offline, bohužel je tento přístup zcela nevhodný pro dokumenty a souborové databázové zdroje otevírané v režimu pro zápis z vícera klientských počítačů v síti v jednom okamžiku.

A právě takovými databázové soubory DBF formátu FoxPro jsou. V praxi to znamená, že nelze efektivně využít databázových zámků, neboť kopie otevřeného databázového souboru existuje pouze v mezipaměti daného klientského počítače a nepromítá se do skutečného souborového systému serveru. Dopadem je vzájemné přepisování nebo ztráta záznamů vytvořených více uživateli pracujícími s jedním datovým souborem najednou.

Jelikož v operačních systémech Windows použitých jako souborové servery CIFS neexistuje možnost vypnout dobrovolné zámkové pouze pro některé typy souborů, popřípadě od verzí operačního systému Vista a Sever 2008 (využívajících už pouze protokol SMB2) je tato možnost zcela vyloučena [40], je buď potřeba spoléhat se

na aplikační zámky<sup>9</sup> vyšších úrovní, které významně zpomalují běh programů, nebo použít jiný způsob pro sdílení databázových zdrojů (více viz 3.3.3).

Realizace zabezpečení dat uložených v DBF souborech je plně v kompetenci školy a jejího správce ICT, programy Bakaláři, mimo možnost šifrovat archivovaná data pomocí nástroje *Archivace*, nenabízí žádnou možnost jejich ochrany. Celkové odepření přístupu k adresářové struktuře není možné, protože některé funkce (např. generování XML pro elektronické předávání dat z matrik) nenabízí žádný dialog pro možnost výběru uložení výstupu uživatelem, ale ukládají výstup přímo do vnitřní adresářové struktury programů na sdíleném disku CIFS serveru.

Ačkoliv nápověda k programům Bakaláři uvádí určitá doporučení k nastavením zvláštních přístupových práv [19], počítá se specifickým nastavením počítačové sítě, které nemusí být ve všech případech pro danou školu vhodné. Protože je však adresářová struktura na první pohled pro nezkušené uživatele či potenciální útočníky značně nepřehledná, může být určitého stupně kvazi-zabezpečení dosaženo odepřením zobrazení obsahu<sup>10</sup> některých adresářů. Jedná se však o značně nespolehlivý koncept *security through obscurity*, kdy je záměrnou změnou výchozího a očekávaného chování takto měněného systému dosaženo pouhé iluze zabezpečení [50].

---

<sup>9</sup>přehled zámků aplikační úrovně lze získat z nabídky *Aktualizace – Správa modulů, uživatelů* nástroje *Aktualizace z www*

<sup>10</sup>pro server Samba pomocí direktivy `dont descend`

## Instalace v síti s daty na SQL serveru

Výše zmíněné problémy s nevyužitelností databázových zámek byly zřejmě jednou, nikoliv však jedinou z hlavních motivací vývojářů pro zavedení možnosti uložit data na jediný centrální SQL sever. V současné době je podporován pouze databázový server *Microsoft SQL Server* v libovolné verzi, včetně bezplatných variant *Express*. Podpora svobodného databázového systému *MySQL*<sup>11</sup> byla ukončena od verze Bakaláři 10/11 [18].

Data umístěná na databázovém serveru jsou také teoreticky mnohem lépe chráněna proti neoprávněnému přístupu pomocí definic uživatelských účtů a rolí přímo na straně daného serveru. Vývojáři programů Bakaláři však neposkytují žádnou podporu jeho instalace a správy a nechávají konkrétní implementaci zabezpečení a problematiku zálohování na odpovědných ICT správcích jednotlivých škol [17].

Možnost s uložením dat na SQL serveru s sebou nese z pohledu odpovědného ICT správce školy další komplikace. Kromě nutnosti disponovat patřičným databázovým serverem vyžaduje další nastavení klientských počítačů – vytvoření odpovídajícího datového zdroje ODBC, dostupného buď na úrovni uživatelského účtu, nebo na úrovni celého počítače. Toto nastavení vytvoří propojení klientského počítače s databázovým serverem v síti a nabídne ho aplikacím jako svůj lokální zdroj. V prostředí klientských počítačů s 64bitovým operačním systémem Windows musí být datové zdroje ODBC registrovány v 32bitové variantě, neboť programy Bakaláři jsou pouze 32bitové.

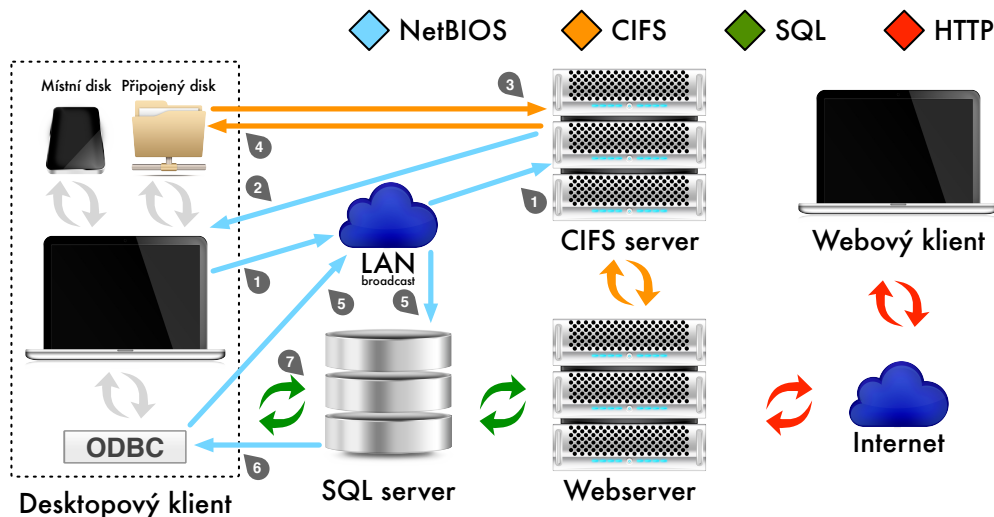
Kromě určitých nevýhod nutnosti manuálního rušení citlivých nastavení a uložených hesel při pravidelném odepisování a vyřazování hardware, či při běžné instalaci nových stolních počítačů, je tato varianta velmi nevýhodná zejména pro přenosné počítače používané i mimo školní síť (typicky služební notebooky), nebo pro školy se složitou síťovou topologií bez jakéhokoliv řízení v podobě *Active Directory*, nebo alespoň severu WINS, kde se může na dobách odezvy programů již projevit přetížení počítačové sítě způsobené reží překladačů jmen NetBIOS [23], nutného pro správnou práci sdílení souborů a tiskáren v dynamickém prostředí.

V simulovaném prostředí, předpokládajícím síť počítačů v pracovní skupině Microsoft Windows bez centrálně spravovaných adresářových a jmenných služeb, se tato možnost celkově ukázala až devítinásobně pomalejší oproti možnosti uložení dat v DBF souborech. Možnost instalace s daty uloženými na SQL serveru nelze zcela

---

<sup>11</sup><http://www.mysql.com>

doporučit na školách s počítačovou sítí bez adresářových a jmenných služeb Active Directory a jim podobných.



Obr. 3.2: Diagram – síťová komunikace programů Bakaláři v konfiguraci s daty uloženými na SQL serveru a webovým serverem pro elektronickou žákovskou knížku

Diagram se snaží zachytit nejhorší možnou výchozí variantu stavu sítě, kdy jsou jednotlivé služby hostované na různých serverech, jejichž názvy NetBIOS nejsou v mezipaměti klientských počítačů a v síti neexistuje žádná služba pro překlad jmen. Příklad neuvazuje další negativní okolnosti, například zátěž způsobenou komunikací webového serveru s klienty v prostředí Internetu, proudící od hlavní brány přes lokální síť.

1. Klientský počítač nenašel v mezipaměti ani v souboru `lmhosts`<sup>12</sup> název souborového serveru pro připojení síťového svazku. Odesílá tedy na *broadcast* lokální síť dotaz na NetBIOS jméno CIFS serveru.
2. Souborový server vyslyšel dotaz a posílá konkrétnímu klientskému počítači svou adresu.
3. Klientský počítač navazuje spojení pomocí protokolu pro sdílení souborů (CIFS) a žádá přístup ke sdílenému disku.
4. Souborový server ověřuje žádost klienta a poskytuje mu přístup k požadovanému zdroji.
5. Na klientském počítači je ze vzdáleného sdíleného disku spuštěn program Bakaláři, využívající knihovny uložené na lokálním disku. Program Bakaláři

<sup>12</sup>Local Manager Hosts – lokální soubor pro rozlišení doménových a NetBIOS jmen

- si vyžádá přístup ke zdroji dat ODBC.
6. Klientský počítač nenašel v mezipaměti ani v souboru `lmhosts` název databázového serveru a posílá na *broadcast* lokální síť dotaz na NetBIOS jméno SQL serveru.
  7. Databázový server vyslyšel dotaz a posílá konkrétnímu klientskému počítači svou adresu.
  8. Zdroj dat ODBC je ověřen a připojen pro potřeby programu Bakaláři.

SQL server je využíván pouze pro aktuální školní rok a data hlavních modulů. Po uzavření školního roku jsou všechna data převedena do archivů v podadresářích v `EVID\ABSOLV\` v podobě nezašifrovaných DBF souborů. Data používaná jinými moduly či podmoduly, například *Knihovnou*, nejsou na SQL server ukládána vůbec. V případě potřeby migrace serveru během reorganizace sítě je nutné pomocí archivačního nástroje získat data z SQL serveru zpět v podobě DBF souborů a na novém místě je opačným způsobem naimportovat do nového serveru.

Jelikož žádný z programů Bakaláři není spuštěn nepřetržitě (viz 3.3), je třeba pravidelné zálohování dat realizovat přímo na úrovni SQL serveru samotného, na což se nevztahuje žádná podpora ze strany vývojářů [17] a je vyžadována určitá znalost problematiky databází a jazyka SQL.

### 3.3.3 Provoz v jiných operačních systémech

Přestože jsou programy Bakaláři oficiálně určeny především pro 32bitové varianty operačních systémů Windows ve verzích 95, 98, NT, 2000, XP, 2003 Server, 2008 Server, Vista, Windows 7 a Windows 8 a 64bitové varianty verzí Windows 7 a pro Windows 8 [18], jejich architektura prakticky nevyklučuje možnost používání v jiných operačních systémech, jsou-li dodrženy určité podmínky.

#### V roli klientského počítače

Pro běh většiny desktopových aplikací vytvořených ve Visual FoxPro je nezbytná podpora aplikačního rozhraní `win32` na cílovém operačním systému. Na mnoha operačních systémech standardů POSIX lze této podpory dosáhnout pomocí překladové vrstvy WINE.

Pro potřeby programů Bakaláři je třeba do konfigurace WINE doplnit další svobodné a nesvobodné doplňky, zejména základní písma typu TrueType *Microsoft Core Fonts*, běhové prostředí jazyka *Microsoft Visual Basic 6.0* a novější, běhové prostředí jazyka *Microsoft Visual C++ 6.0 Service Pack 6* a novější, knihovny *Microsoft Data Access Components 2.8* pro přístup k některým datovým zdrojům a *Microsoft XML Core Services* pro zpracování souborů formátu XML. Po správném namapování cest k jednotlivým programům [27] již nic nebrání jejich spuštění v tomto operačním systému.

Protože však WINE není emulátor a programy Bakaláři vyžadují 32bitové prostředí platformy Intel x86, musí být procesorová architektura cílového počítače s jiným operačním systémem také typu Intel x86, nebo Intel x86\_64 s překladovými knihovnami umožňujícími běh 32bitových aplikací v 64bitovém prostředí. Operační systémy Windows 7 a 8 ve svých 64bitových variantách umožňují spuštění 32bitových aplikací [41] pomocí vlastní překladové vrstvy WoW64.

Využití operačního systému standardu POSIX v kombinaci s WINE ale určitou příležitost k práci s programy Bakaláři na jiných procesorových platformách nabízí, jsou-li v součinnosti s vlastním operačním systémem schopny poskytovat server X11 pro interpretaci grafického výstupu získaného z původního počítače, na kterém jsou programy Bakaláři pomocí vrstvy WINE spuštěny. Takto lze dosáhnout možnosti ovládat programy například na tabletech platformy ARM.

## V roli serveru

Situace jiných operačních systémů v roli serveru je o mnoho snazší a prakticky vychází z historického konceptu naznačeného v kapitole 3.3.2. Jako souborový server může sloužit dříve rozšířený operační systém Novell Netware [19], nebo populární implementace CIFS serveru Samba, určená pro mnohé operační systémy standardů POSIX.

Konfigurace serveru Samba pro sdílení souborů a poskytování programů umožňuje vypořádat se s potížemi s databázovými zámkami popsány v kapitole 3.3.2. Příklad konfigurace s omezením dobrovolných zámků pro databázové soubory je uveden v příloze A. Značnou výhodou serveru Samba je jeho otevřenost a prakticky žádné omezení v počtu souběžných připojení. Oproti tomu implementace sdílení souborů v systémech Windows umožňují přístup neomezenému počtu uživatelů pouze ve variantách Server, což může pro školu představovat další náklady spojené s poří-



zením odpovídající softwarové licence.

Nástroje běžně dostupné v jiných, svobodných operačních systémech, mohou být také s výhodou využity pro kompletní zálohování databáze v DBF souborech s pomocí automatizovaných úloh a jiných prostředků. Zatímco server Samba poskytuje sdílený obsah s programy Bakaláři na virtuálním souborovém systému čitelném z klientů Windows, samotný skutečný souborový systém serveru může být nasazen na samočinné sledování a případně vracení provedených změn (souborové systémy typu *copy-on-write*). Příkladem, i když pro rozsah programů Bakaláři možná poněkud přehnaným, může být souborový systém *HAMMER*, vyvíjený jako součást operačního systému DragonFly BSD<sup>13</sup>, poskytujícím vysoký výkon s vestavěnou možností přístupu k přesným historickým časovým okamžikům s velice jemně škálovatelnou granularitou.

Dříve podporovaný multiplatformní a otevřený databázový server MySQL byl nahrazen nutností použití Microsoft SQL Serveru [17]. S postupy popsány pro provoz klientské části programů v jiných operačních systémech lze však i Microsoft SQL Server spustit v prostředí unixového operačního systému [43].

### 3.3.4 Provoz webové aplikace

Webová aplikace Bakaláři, popisovaná též jako *elektronická žákovská knížka*, je historicky nejmladším modulem systému Bakaláři a používá zcela odlišný návrhový a uživatelský koncept oproti desktopovým aplikacím ostatních modulů.

Každá webová aplikace potřebuje k poskytování svého obsahu v síti Internet nebo sítích intranetu webový server a k běhu dynamické části aplikačního rozhraní, na které jsou požadavky webového serveru delegovány. Dřívější webová aplikace Bakaláři využívala dynamické skriptování realizované jazykem PHP, nové verze jsou k dispozici již jen pro rozhraní Microsoft ASP.NET<sup>14</sup>.

Doporučeným webovým serverem je Microsoft IIS obsažený přímo v instalaci systému Windows [20], avšak použít lze i jakýkoliv jiný, umožní-li vyřizovat požadavky přes nativní rozhraní ASP.NET 4.0. Webová aplikace využívá doplňků a frameworků třetích stran, mezi nimi konkrétně komponenty Obout<sup>15</sup>, knihovny

---

<sup>13</sup><https://www.dragonflybsd.org>

<sup>14</sup>od verze 08/09 ASP.NET 2.0, od verze 12/13 ASP.NET 4.0

<sup>15</sup><http://www.obout.com>

a MVC framework DevExpress<sup>16</sup>, nástroje pro podporu generování sestav Stimulsoft<sup>17</sup> a velké množství dalších.

Některé z těchto doplňků mají běhovou závislost na WPF, která není obvyklou či nutnou pro webové aplikace. Díky tomu je vývoji webové aplikace téměř znemožněno bez úprav některých konfigurací provozovat elektronickou žákovskou knížku na svobodných alternativách platformy .NET v jiných operačních systémech, například na platformě projektu Mono<sup>18</sup> v kombinaci s webovým serverem Apache.

Technologie webových aplikací jsou oproti zbytku desktopových programů Bakaláři připraveny na přímé propojení s databázovými servery. Jsou-li však programy síťově instalovány s daty uloženými v DBF souborech, webová aplikace s nimi též může pracovat. Naneštěstí není podpora databázových souborů Visual FoxPro zahrnuta v *Microsoft Data Access Components* od verze 2.6 [39], je tedy nutné na daném webovém provést dodatečnou samostatnou manuální instalaci databázových rozšíření.

Tato rozšíření však nejsou z důvodů ukončení podpory Visual FoxPro dostupná pro 64bitové systémy Windows, je tedy vyžadováno použití 32bitových technologií, které mohou mít za určitých podmínek pouze omezený přístup k hardwarovým prostředkům [22, 41].

---

<sup>16</sup><https://www.devexpress.com>

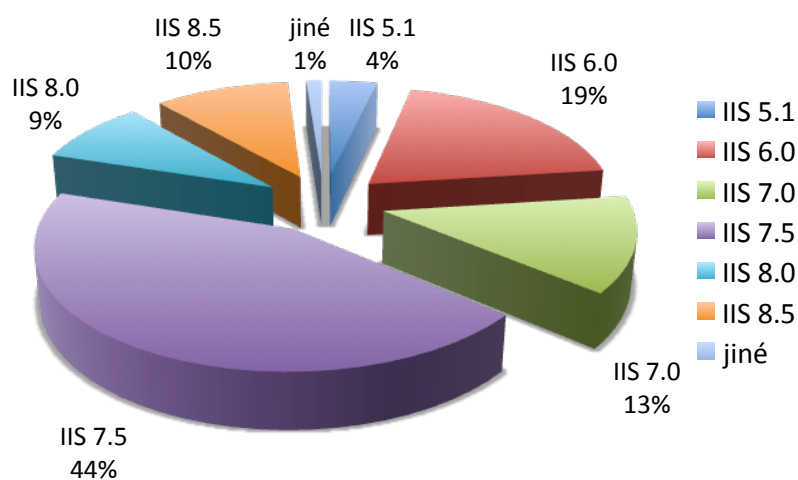
<sup>17</sup><http://www.stimulsoft.com>

<sup>18</sup><http://www.mono-project.com>

## Webové technologie nasazené v praxi

Díky průzkumu výskytu webových aplikací bylo možné pokusit se identifikovat některé technické detaily - použité serverové technologie a konkrétní verze rozhraní ASP.NET - ve skutečném produkčním prostředí škol.

Kromě několika vzácných výskytů jiných webových serverů ukazuje výsledný graf zastoupení jednotlivých verzí Microsoft IIS, a to z celkového počtu 761 nalezených aktivních instancí webové aplikace Bakaláři. Z těchto výsledků by bylo dále ještě možné odvodit pravděpodobnou verzi systému Windows, na kterém je služba IIS spuštěna, jelikož její jednotlivé nové verze vycházely souběžně s novou verzí jádra NT [13]. Z výsledků nebylo možné jednoznačně určit, je-li na dané škole použita instalace programů Bakaláři s daty uloženými na SQL serveru, nebo s daty v DBF souborech.



Obr. 3.3: Graf – zastoupení jednotlivých webových serverů

Položka *jiné* připadá na servery *Aprelium Abyss*<sup>19</sup> v bezplatné variantě *X1*, jednu instanci serveru *Apache* a na *nginx*. Webový server *nginx* je díky své vysoce škálovatelné architektuře populární především pro svůj vysoký výkon v režimu *reverse*

<sup>19</sup><http://www.aprelium.com/abyssws>

*proxy*<sup>20</sup>, kdy klientům předává statický obsah a zpracování dynamického obsahu deleguje na jeden či více jiných serverů či aplikačních rozhraní [37]. Je proto pravděpodobné, že díky tomu pouze změnil obsah výsledné HTTP hlavičky **Server** na svůj název.

Verze prostředí ASP.NET byla ve všech případech *4.0*, která je minimální vyžadovanou verzí, na které může aplikace 12/13 a vyšší pracovat. Přibližně pětina umístění webové aplikace umožňovala zabezpečené připojení pomocí SSL.

Následující tabulka uvádí konkrétní počty nalezených aktivních webových aplikací Bakaláři u škol daných typů a jejich podíly v rámci škol používajících tuto aplikaci a celkové podíly v rámci všech škol daných typů v ČR.

Tab. 3.2: Podíly typů škol využívajících webovou aplikaci podle výběru z číselníku AKDT – Druhy a typy škol a školských zařízení

Kód	Význam	Celkový počet	Počet w.ap.	Podíl	Celkový podíl
<i>Bxx</i>	Základní školy	4 155	267	31,94 %	6,42 %
<i>Cxx</i>	Střední školy a gymnázia	1 383	500	59,81 %	36,15 %
<i>C5x</i>	Pouze gymnázia	384	173	20,69 %	45,05 %
<i>Dxx</i>	Konzervatoře	18	1	0,12 %	5,55 %
<i>Exx</i>	Vyšší odborné školy	187	68	8,13 %	36,36 %
–	Celkem	5 471	761	–	13,90 %

Určitá chyba způsobená rozdělením druhů a typů škol zahrnuje školy s více stupni a typy vzdělávání využívající jedinou instanci aplikace.

<sup>20</sup>pro tento účel byl server původně vytvořen jako náhrada modulů pro akceleraci a rozložení zátěže serveru Apache

Dalším zajímavým pozorováním reálného provozu je označení stránky aplikace v názvu adresy URL. Názvy subdomén s aktivní aplikací byly rozděleny do uvedených ekvivalenčních tříd v kontextu oblasti školství. Velké množství škol nepoužívá žádné jméno, má server konfigurován pro přístup pouze pod IP adresou.

Tab. 3.3: Zastoupení názvů v adresách URL

Třída významu	příklad	podíl
Bakaláři	baka, bakalari, bakaweb, bkl	33,0 %
Klasifikace	znamky, zakovska, knizka	7,1 %
Označení školy	skola, zs	5,6 %
Brána nebo e-mailový server	fw, gw, mail, owa <sup>21</sup> , router	2,6 %
Informace pro skupiny	rodice, studenti	2,6 %
Informační systém a intranet	is, info, prihlasit	2,4 %
Přímá součást webových stránek školy	www.skola.cz/bakaweb	5,6 %
Jiné	ww2, www.isp.cz/skola/	10,1 %
Pouze IP adresa	–	33,0 %

Za zvláštní zmínku stojí jeden zaznamenaný výskyt subdomény sas, pod kterou se webová aplikace Bakaláři dané školy nachází. Může vycházet ze zkratky SAS, označující konkurenční informační systém pro školy společnosti MP-Soft a.s.

Přestože zcela jistě nebyly nalezeny úplně všechny výskyty webových aplikací, jejich celkové počty se ani vzdáleně neblíží číslu 1 400, uváděném v příručce systému Bakaláři pro rok 2014 [30].

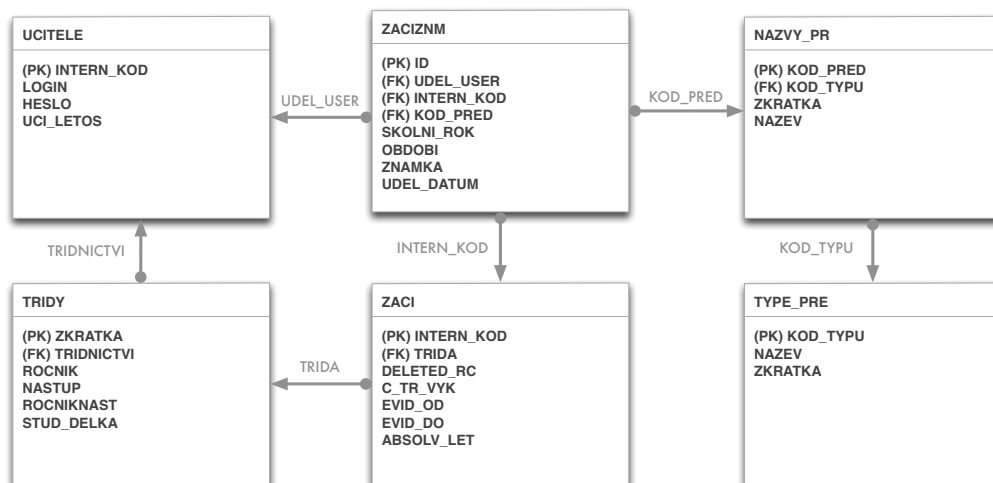
Metodiky prováděného průzkumu jsou podrobně popsány v příloze B. Údaje zde uvedené vycházejí z výsledků metody popsané v B.2.1.

<sup>21</sup>Microsoft Exchange – Outlook Web Access

### 3.4 Podoba a formát uložených dat

Programy Bakaláři zachovávají všechna dříve zadaná data. V běžném režimu ale pracují pouze s aktuálním školním rokem (do jiného školního roku se lze přepnout například v případě nutnosti dodatečného tisku vysvědčení) a historickými údaji pouze současného souboru žáků. Aktuální školní rok se na SQL serveru nachází v rozsahu uvedeném v kap. 3.3.2, v případě používání DBF souborů v adresáři EVID\data\dbf\. Kompletní data již uzavřených období jsou archivována v DBF souborech v adresáři EVID\ABSOLV\absYYYY\FullYear.zip, kde YYYY označuje rok uzavření odpovídajícího školního roku. Součástí adresáře jsou také textové soubory nesoucí čitelný údaj s datem, ke kterému byl daný školní rok uzavřen.

Relační databáze je realizována pouze na aplikační úrovni samotného prostředí Visual FoxPro, společného pro všechny programy a moduly. V případě uložení dat na serveru Microsoft SQL Server jsou označeny pouze primární klíče jednotlivých tabulek, databáze pouze z pohledu SQL serveru relační není. Názvy jednotlivých tabulek uložených na SQL serveru a názvy DBF souborů s odpovídajícími tabulkami jsou shodné, přičemž není rozlišována velikost písmen. Z instalace na SQL server tedy neplynou žádné výhody z pohledu realizace relační databáze.



Obr. 3.4: Diagram – zjednodušený relační model základní evidence celkové klasifikace pro každý školní rok

Kompletní databáze obsahuje velké množství tabulek, pro zjednodušení zde jsou uvedeny pouze některé z nich (a pouze s omezeným výběrem sloupců), využívané

*Společným prostředím a Evidencí.* Názvy tabulek jsou s výjimkou globálních nastavení programů a protokolů o zápisech klasifikace uvedeny v češtině bez diakritiky buď celým slovem, nebo zkratkou, aby nepřekročily limity 8.3 názvů systému DOS. Číselníky a definované číselníky zpravidla začínají symbolem podtržítka. Význam zkratek je možné za použití doménových znalostí intuitivně odvodit, případně je lze v omezené míře dohledat v integrovaném datovém slovníku, spouštěném pomocí parametru hlavního programu – `bakasql.exe slovník`. Slovník mimo popis struktury a významu jednotlivých tabulek a sloupců popisuje i organizaci samotné adresářové hierarchie programů a doporučená bezpečnostní nastavení pro zajištění optimální bezpečnosti přístupu.

Databáze není důsledně normalizovaná a mnoho shodných údajů je uloženo na více různých místech. Je to však dáno jednak historickým vývojem, ale také i potřebou zachovat lidsky-čitelnou strukturu pro relativně volná rozšíření o další položky a funkce.

### 3.4.1 Důležité tabulky

Tabulka ZACI je v programech označována jako „data celé školy“ a je určitou kombinací *katalogového listu* a části *třídního výkazu*. Jsou v ní uvedeny všechny osobní údaje o žácích aktuálního školního roku, včetně informací o jejich zákonných zástupcích, přiřazeném vzdělávacím programu, nebo textová pole o zdravotním či rodinném stavu, sloužící především pro třídní učitele.

Pro záznam *celkové klasifikace* a jako podklad pro generování výsledných sestav nutných pro tisk *vysvědčení* slouží tabulka ZACIZNM, nesoucí současné a historické údaje o udělené celkové klasifikaci žáků, evidovaných v aktuálním školním roce. Kromě cizího klíče odpovídajícího předmětu, ze kterého byl žák klasifikován, obsahuje každý záznam datum a čas zanesení údajů do systému, pořadí klasifikace a cizí klíč učitele, který hodnocení provedl. Prováděl-li hodnocení správce systému, je jako cizí klíč uvedena hodnota \*, případně je buňka zcela prázdná.

Struktura, uvedená v diagramu výše, je však platná pouze pro programy verze 05/06 a vyšší. Ve starších verzích nebyly tabulky v současné úrovni normálních forem. Údaje o klasifikaci, datu klasifikace, předmětech, celkových průměrech, celkovém prospěchu, počtech omluvených a neomluvených hodin a historických třídách a číslech v třídním výkazu byly všechny součástí jedné tabulky ZACI. Programy jsou ale zpětně kompatibilní, umí v případě otevření staršího roku s těmito údaji korektně pracovat. Protože však legislativa neumožňuje zasahovat do starších, již uzavřených

období, není možné přímo za použití programů převést takto strukturované záznamy do současně používané formy.

### 3.4.2 Datové typy FoxPro

Databáze FoxPro, resp. její datové soubory DBF, umožňují záznam základních typů jako u většiny moderních databází. Čísla bývají označena jako NUMBER, společně s údajem o své šířce a počtu uchovávaných desetinných míst, data a časy jako DATE, resp. DATETIME, umožňující uložit datum či datum a časový údaj, používaný v IS Bakaláři například pro účely protokolování zápisů klasifikace, logické hodnoty LOGICAL, nabývajících hodnot *pravda* či *nepravda* zapsaných jako .T., resp. .F., textové poznámky MEMO, často párované na externí soubory s příponou .MEM a v neposlední řadě textové řetězce CHARACTER, nesoucích ještě údaj o své šířce.

Visual FoxPro ale není silně typovaným jazykem, proto ve svém prostředí nabízí ještě obdobné datové typy tzv. *polí*, používaných v grafických ovládacích prvcích pro reprezentaci datových typů uložených v databázi [42].

### 3.4.3 Data programů Bakaláři

Základním poznatkem o fyzické organizaci dat, uchovávaných programy Bakaláři, je fakt, že téměř všechny pro analýzu důležité hodnoty jsou uloženy jako textové řetězce, používající pevnou šířku, na kterou bývají doplněny odpovídajícím počtem mezer.

Důvodů používání textových řetězců může být mnoho, jedním z hlavních je využití ne zcela zřejmých cizích klíčů i jako přímých nositelů hodnot. Například v celkové klasifikaci, v tabulce ZACIZN, je typ sloupce ZNAMKA textovou hodnotou, referující na aplikační úrovni na číselník základních a dodatečných typů klasifikace. Mimo očekávaných hodnot 1, 2, 3, 4 a 5 lze ještě zapsat mnoho jiných označení, zejména *N*, označující stav *neklasifikován(a)*, *U* pro *wolněn(a)*, či *S*, napovídající o zvláštním záznamu se slovním hodnocením, nikoliv číselným hodnocením. Protože jsou programy Bakaláři určeny pro české prostředí operačního systému Windows, používají pro ukládání textových informací znakovou sadu windows-1250.



### 3.4.4 Čtení a modifikace dat mimo prostředí programů

Protože je ověřování oprávnění uživatelů pro práci s určitými informacemi a jejich sestavami realizováno pouze na aplikační úrovni a ochrana samotných dat je delegována na nižší vrstvy souborového systému, resp. SQL serveru, není nezbytně nutné disponovat platným uživatelským účtem v rámci celého IS.

Formáty souborů DBF, vycházejících z *dBase*, které používá i prostředí Visual FoxPro, nejsou v programech Bakaláři nijak zašifrované a existuje celá řada softwarových nástrojů pro jejich čtení, například včetně tabulkového procesoru LibreOffice. Ne však všechny z nich mohou podporovat zápis DBF souborů novějších verzí *dBase 7*, které používá právě Visual FoxPro, případně může hrozit poškození indexů daného DBF souboru. Při instalaci s daty uloženými na SQL serveru je možné k celé aktivní databázi bez jejího narušení přistupovat pod účtem správce SQL serveru, nebo řídicím účtem programů Bakaláři, vytvořeným během instalace a konfigurace zdrojů ODBC.

Z hlediska bezpečnosti se nabízí otázka, je-li možné na úrovni datových souborů nebo SQL serveru neoprávněně modifikovat určitá data mimo programy Bakaláři v produkčním prostředí školy. Odpověď závisí pouze na motivaci případného útočníka. Při pomnutí závažných rizik spojených s odcizením osobních údajů (kde ale není nutné žádné oprávnění pro zápis) a záměrným poškozením celé databáze (proti kterému by měla být chráněna pravidelnými zálohami), přichází v úvahu pouze záměrná změna průběžné nebo celkové klasifikace samotnými žáky [15].

Údaje o klasifikaci jsou však na úrovni aplikace chráněny různými redundantními kontrolami, spojenými s protokoly zápisů klasifikace. Evidence průběžné klasifikace také používá na první pohled nepřehledné systémy [50] cizích klíčů z číselníků vah jednotlivých typů hodnocení, vytvořených buď správcem, nebo každým z učitelů zvlášť, a orientace v nich vyžaduje časově mnohem náročnější analytický přístup.

V tomto ohledu je datově zcela korektní modifikace klasifikace zajistitelná pouze po ověření oprávnění učitele na aplikační úrovni. Za důslednou ochranu svých přihlašovacích údajů nese, jako v případě uživatelů jakéhokoliv jiného informačního systému, plnou odpovědnost každý učitel.

## 3.5 Shrnutí

Přestože je na celém systému Bakaláři díky četným nekonzistencím v ovládacích prvcích uživatelského rozhraní či organizační struktuře dat patrná jeho degradace a opotřebení [35], uplatněná vývojová metodologie rapidního prototypování s použitím vhodných vývojových nástrojů (Visual FoxPro) se díky jeho celkové rozšířenosti a více než patnáctileté době provozu na českých školách ukázala jako velice úspěšná pro informační systém tohoto typu a rozsahu. Významnou roli zcela nepochybně sehrál i fakt, že mnozí, ne-li všichni z původních i současných vývojářů, byli či jsou původním povoláním pedagogové, zcela seznámení s oblastí školství, tvořící komplexní znalostní doménu zpracovávanou tímto informačním systémem.

Technologie použitá při vývoji, spoléhající se na databázové prostředí aplikační úrovně Visual FoxPro, je v současné době již zastaralá a je překážkou systému Bakaláři v jeho dalším šíření na trhu. Neefektivní využívání síťových prostředků, modulární roztržitost, problematický provoz na přenosných zařízeních, platformní omezení na architekturu Intel x86 a praktická nemožnost zcela spolehlivě ochránit data, to jsou aktuálně největší slabiny celé sady programů.

Vzhledem ale k tomu, že ze strany společnosti Microsoft skončila podpora produktu Visual FoxPro v roce 2007, je budoucnost dalšího vývoje programů ve stejné podobě nejistá. Je pravděpodobné, že se programy budou integrovat do postupně se rozšiřující webové aplikace tak, aby časem obsáhla veškerou funkcionalitu dosud poskytovanou pouze provozem desktopových aplikací. Této možnosti kromě výhod úplné klient-server architektury s odděleným datovým úložištěm nahrává i skutečnost, že webové prohlížeče v roli tenkých klientů již dlouho standardně existují na různých platformách, včetně tabletů a jiných přenosných dotekových zařízeních, která mají již dnes v prostředí škol ideální uplatnění.

## 4 PŘÍPADOVÁ STUDIE

Teoretické poznatky, popsané v předchozích kapitolách, byly v rámci případové studie aplikovány v reálném prostředí základní školy Pardubice – Studánka, využívající informační systém Bakaláři pro správu své školní agendy.

### 4.1 Základní škola Pardubice – Studánka

Základní škola Pardubice – Studánka je úplnou základní školou s 1. až 9. ročníkem pro 720 žáků a pedagogickým sborem čítajícím na 70 vyučujících, vybavena je vlastní školní družinou a jídelnou. Je příspěvkovou organizací – samostatným právním subjektem, zřízeným Magistrátem města Pardubic v roce 1994. Její historie však sahá do jejího založení v roce 1963. Geograficky je škola umístěna v tříhektarovém parkovém areálu se sportovišti a sportovními halami, jídelnou, budovou školní družiny a jednotlivými pavilony s učebnami, rovnoměrně rozmístěnými v areálu [6].

Rozmístění jednotlivých školních budov je podřízena i fyzická topologie počítačové sítě, která je typu asymetrické rozšířené hvězdy. Rozvaděče strukturované kabeláže jsou u pavilonů s kmenovými třídami umístěny v prvním patře ve společných učitelských kabinetech a v počítačových učebnách budovy odborných učeben, která tvoří pomyslný hlavní uzel sítě. Budova odborných učeben je s ředitelstvím, kde se nachází hlavní brána poskytovatele připojení, propojena optickou linkou. Největší slabinou sítě, spojenou s nutností násobného větvení mezi různými budovami, jsou pro logickou topologii hvězdy kapacitně ne zcela postačující aktivní prvky – switche.

#### 4.1.1 Historie nasazení IS Bakaláři

Z rozhovorů s odpovědnými zaměstnanci vyplývá, že prvotním impulsem pro nasazení programů Bakaláři od školního roku 2002/03 bylo, kromě očekávaného snížení administrativní zátěže, především sjednocení vzhledu výsledných vysvědčení, které byly do té doby všechny psané ručně, každé jiným, odpovědným třídním učitelem. Tehdejší vedení školy nebylo digitalizaci agendy příliš nakloněno.

Prvotní instalace byla provozována na jednom počítači v počítačové učebně a s programem pracoval pouze ICT správce, který si občas nosil aktivní kopii na USB disku domů, aby na ní mohl pracovat a rozšiřovat dosud papírovou evidenci do digitální podoby i mimo pracovní dny. Slučování změn provedených na více místech bylo značně obtížné.

K nasazení síťové instalace došlo v roce 2004, kdy byl pro požadované účely k dispozici vyhrazený již nepoužívaný serverový počítač s operačním systémem Windows z projektu MŠMT *Internet do škol*. Byla tedy zvolena varianta instalace s daty uloženými na SQL serveru. Tuto instalaci provedli technici poskytovatele internetového připojení školy. V tomto stavu často docházelo k přetěžování počítačové sítě, jak je popisováno v kap. 3.3.2. Zálohování dat probíhalo v definovaných intervalech na kazety s magnetickou páskou. Provoz varianty s daty na SQL serveru byl ukončen v roce 2011.

V roce 2004 byly poprvé z programů vytištěny katalogové listy žáků (pro 1. a 6. ročníky) a kromě evidence žáků se užívání programu ujalo i pro evidenci přihlášek žáků na střední školy.

#### 4.1.2 Současný stav užívání IS Bakaláři

V současné době je provozována instalace programů Bakaláři ve variantě s daty uloženými v DBF souborech, v roli serveru je počítač s 64bitovým operačním systémem *Debian GNU/Linux 8.0 Jessie*, poskytující sdílené soubory programů pomocí software SAMBA. Zabezpečení přístupu je zajištěno kontrolou povolených počítačů jednotlivých pedagogů. Adresářová struktura je na serveru izolována, správce ICT získává výstupy pro odeslání dat školní matriky pomocí symbolických odkazů ve svém domovském adresáři operačního systému GNU/Linux.

Pro případ havárie používá server dvojici pevných disků v konfiguraci RAID 1, přírůstkové zálohování probíhá pomocí automatizovaných plánovaných úloh denně na lokálních discích a krátce po začátku každého pololetí jsou prováděny zálohy plné, odesílané na úložiště hostingového střediska, na kterém se nachází webová prezentace školy. Obě varianty záloh jsou snadno obnovitelné a jsou asymetricky šifrované pomocí klíčů PGP.

Škola vlastní licenci programů *Společného prostředí*, modulu pro *Evidenci žáků a zaměstnanců*, *Suplování*, *Plány akcí*, a *Rozvrh a generátor rozvrhu*. Modul pro rozvrhy však není používán, neboť ve škole platí jiné časy začátků hodin pro 1. stupeň, jiné časy pro 2. stupeň bez odpoledního vyučování a s odpoledním vyučováním. Programy Bakaláři s tímto uspořádáním nepočítají, proto není možné rozvrhy vytvářet. Z tohoto důvodu zatím nejsou elektronicky vedeny ani třídní knihy, neboť jejich odpovídající modul je závislý na správně vytvořeném rozvrhu. Moduly *Knihovna* a *Webová aplikace* pro realizaci elektronické žákovské knížky byly pořízeny až v letošním školním roce 2014/15. Vedení školy v současné době (duben 2015) mezi rodiči

žáků ověřuje, zda a do jaké míry chtějí tohoto nového způsobu komunikace využívat.

O běh webové aplikace se stará server Microsoft IIS, spuštěný ve virtualizovaném prostředí 32bitového systému Windows 7. Virtualizaci pod linuxovým počítačem realizuje hypervisor KVM, požadavky z Internetu jsou delegovány na virtuální systém z webového serveru Apache. Limity a celkové zabezpečení zajišťuje firewall `iptables`, resp. jeho framework `netfilter`, který je součástí linuxového jádra.

## 4.2 Specifikace analýzy dat

Vzhledem k relativně omezenému rozsahu, ve kterém je na této škole systém Bakaláři používán, se analýza orientuje převážně na dlouhodobé vyhodnocení celkové klasifikace žáků. Přestože od tohoto školního roku, 2014/2015, již dochází k zápisu průběžné klasifikace a její zveřejňování v elektronické žákovské knížce, nebyla zde analýza prováděna. Důvodem je v kontrastu s celkovou dobou provozování IS ve škole příliš krátké používání průběžné klasifikace, navíc omezené tím, že mnozí pedagogové se ještě teprve učí s tímto nástrojem a jeho možnostmi pracovat.

- Testováno je celkové využívání samotné evidence žáků a zaměstnanců pro zvýraznění postupného zaplňování databáze, ke kterému docházelo v začátcích nasazení systému od roku 2002 do roku 2004.
- Dále je provedena analýza a zhodnocení vývoje celkové klasifikace žáků za posledních 13 let, kdy je IS Bakaláři na škole používán (zahrnuto je i první pololetí školního roku 2014/2015).
- Dle v úvodu zmíněných metodik jsou ilustrativně analyzovány některé nečíselné charakteristiky.
- Je hledána korelace nebo kauzalita v evidovaných datech.

Analytickým nástrojem, použitím pro všechna zpracování specifikovaných charakteristik, bylo zvoleno svobodné statistické prostředí  $R^1$ , použité na platformách operačních systémů GNU/Linux a Apple OS X.

---

<sup>1</sup><http://www.r-project.org>

### 4.2.1 Anonymizace dat

Jelikož se tato případová studie zabývá analýzou skutečných dat, získaných z informačního systému Bakaláři provozovaného více než deset let v prostředí Základní školy Pardubice – Studánka, byly pro účely analýzy a zveřejnění podkladů na CD, přiloženém k této práci, určité klíčové osobní údaje převedeny do anonymní podoby tak, aby nedošlo k porušení zákona 101/2000 Sb., O ochraně osobních údajů [2].

**Jména a příjmení** byla v evidenci žáků a učitelů změněna na *Jan Novák*, resp. *Jana Nováková* v závislosti na pohlaví. Jména a příjmení všech evidovaných zákonných zástupců žáků byla změněna na *Alena Nováková*, resp. *Dalibor Novák* a *Jakub Novák*.

U **rodných čísel** a **dat narození** byl zachován pouze odpovídající rok a měsíc, ostatní části a části za lomítkem včetně kontrolní číslice<sup>2</sup> byly nahrazeny náhodnými číselnými sekvencemi.

**Adresy** a názvy ulic byly změněny na hodnoty *Ulice 321* u žáků, na *Matčina 123*, *Otcova 123* a *Zástupná 324* u zákonných zástupců.

Veškerá **telefonní čísla** a **čísla občanských průkazů** a **pasů** byla změněna na zcela náhodné devítimístné číselné sekvence.

Údaje o **zaměstnání** zákonných zástupců žáků byly změněny na obecné *pracovník – Firma s.r.o.*, resp. *pracovnice – Firma s.r.o.*

**E-mailové adresy** byly ve všech výskytech změněny na *user@mail.cz*.

Všechny **tituly**, byly-li uvedeny v evidenci zaměstnanců, byly změněny na *Ing.*, resp. *Ph.D.* za jménem.

**Přihlašovací jména** a šifrované podoby **hesel** v evidenci zaměstnanců byla nahrazena údaji *user123* a *XXXXXXXXX*.

---

<sup>2</sup>anonymizovaná data nemají validní formát umožňující jejich zpětný import do systému

## 4.2.2 Transformace a konsolidace dat

Výchozí stav, ve kterém se datové soubory systému Bakaláři nacházejí, není zcela vhodný pro účely přímého automatizovaného zpracování. Nezbytné je sjednotit velké množství samostatných datových souborů DBF do uzavřených celků, představujících jednotlivé školní roky, a převést kódování českých znaků jednotlivých položek do mezinárodní znakové sady UTF-8 kvůli možnostem snazšího zpracování na moderních operačních systémech.

Pro účely seskupení byl zvolen databázový formát SQLite3. Hlavním důvodem byla jeho snadná přenositelnost, podpora plnohodnotných relačních databází, podpora mezinárodních znakových sad a také vlastnost, že obdobně jako FoxPro není silně typově orientovaný.

Proces převodu a transformace dat byl mimoděk zahrnut již při anonymizaci osobních údajů, ze které ale vyplynulo několik úskalí. V rámci snahy zachovat originální, ale plně anonymizovanou databázi, bylo provedeno několik pokusů o přímou modifikaci souborů DBF pomocí skriptovacího jazyka PHP s DBI rozšířením poskytovaným repozitářem PEAR. Zápis však nebyl možný pro tabulky verze dBase 7. V rámci procesu anonymizace tedy došlo rovnou k převodu na formát SQLite3 a kódování UTF-8, zachovány však byly pevné mezery, nacházející se ve všech původních polích.

Transformace relačního schématu z aplikační úrovně Visual FoxPro nebyla provedena a není v tomto případě nutná, neboť s tabulkami lze intuitivně pracovat v použitém analytickém nástroji i bez databází spravovaných relací.

Provedena byla konsolidace formátu starších tabulek z let 2003 – 2005, který byl výchozí v programech verze 04/05 a nižších (viz 3.4.1). Nově vytvořená tabulka ZACIZNM obsahuje pouze záznamy z jí odpovídajícího školního roku, nebyla přenesena celá historie celkového hodnocení. Takto konsolidované databáze byly označeny 2003c.db, 2004c.db a 2005c.db.

## 4.3 Výsledky a interpretace analýzy

Za použití doménové znalosti o oblasti školství byly navzájem odděleny skupiny žáků dle odpovídajících stupňů, pro 1. stupeň jsou to žáci 1., 2., 3., 4. a 5. ročníků, na 2. stupni pak žáci 6., 7., 8. a 9. ročníků.

Jednotlivé vyučovací předměty byly intuitivně rozděleny na *cizí jazyky*, předměty *přírodovědné*, kam byly zařazeny všechny vyučovací předměty vycházející z exaktních věd (matematika, fyzika, přírodopis, ...), a na předměty *společenskovědní*, do kterých spadají ostatní nauky (dějepis, hudební výchova, občanská výchova, ...). Byl k nim přiřazen i *český jazyk*, který má jako vyučovací předmět národně-kulturní význam.

### 4.3.1 Míra využití IS Bakaláři

Využívání systému bylo hodnoceno celkovým počtem evidovaných záznamů v každém roce jeho provozu. Absolutní počty žáků v jednotlivých ročnících zcela neodpovídají tvrzení, že byl systém v počátečních letech naplňován postupně od 1. a 6. ročníků, naopak je patrné, že byl plně využíván pro celý 2. stupeň a postupně byly pouze doplňovány údaje žáků 1. stupně.

Protože rušení záznamů v tabulce učitelů pomocí příznaku DELETED\_RC je dodržováno až od roku 2013, nejsou patrné přesné změny v počtu pedagogických pracovníků. Počáteční nulové údaje způsobuje celková absence tabulky UCITELE pro zálohy jednotlivých uzavřených období ve verzích programů 04/05 a starších.

Detailní přehled zachycuje příloha C. Na první pohled je patrná nekonzistence v dílčích počtech záznamů v období školního roku 2009/2010, kde zcela chybí 1. ročníky, ale 9. ročníky jsou zastoupeny ve dvojnásobném počtu. To může poukazovat na chybu se špatným uzavřením školního roku, nebo mnohem pravděpodobněji na špatně obnovenou zálohu pořízenou během uzavírání školního roku a migrace dat z SQL serveru na data uložená v souborech DBF.

### 4.3.2 Vývoj celkového hodnocení žáků

Výsledky analýzy vývoje celkového hodnocení žáků byly vizualizovány pomocí krabicových grafů, které jsou k dispozici v přílohách D. V popisné statistice jsou krabicové grafy velmi přehledným nástrojem, zobrazujícím především medián – střední hodnotu statistického souboru, neovlivněnou extrémními hodnotami – a odpovídající



variační rozpětí mezi prvním a třetím kvartilem, díky kterému je i bez nutnosti výpočtů snadno patrná šikmost a špičatost sledovaného rozdělení.

Vyhodnocení každého školního roku probíhalo výpočtem odpovídajících průměrů známek každého žáka zvlášť. Kromě celé školy byly odděleně analyzovány 1. a 2. stupeň, přičemž u 2. stupně byly ještě odděleny typy jednotlivých předmětů.

Z pohledu na získané výsledky je zřejmé, že v posledních letech nepatrně klesá medián a zužuje se i variační rozpětí, přibývá však počtu odlehlých bodů. Z pohledu vzdělávacího procesu to znamená, že celková úroveň žáků sice stoupá a jejich výsledky jsou lepší, v posledních letech ale roste i též počet individuálních případů, které se očekávanému rozpětí vymykají. Patrné je to zejména u výsledků žáků 1. stupně, kde jednotlivé odlehlé body mohou zachycovat případy nesnadné adaptace žáků na postupně se zvyšující náročnost probírané látky. Přesné vysvětlení tohoto jevu by ale vyžadovalo názor odpovědných metodiků vzdělávání pro 1. stupeň, majících na starost příslušný ŠVP.

Naopak výsledky žáků 2. stupně školy ukazují dlouhodobě zlepšující se charakter, a to zejména v přírodovědných předmětech. Lze se domnívat, že se na tom určitou měrou podílí zapojení školy do projektu *Podpora přírodovědného a technického vzdělávání v Pardubickém kraji*<sup>3</sup>, jehož cílem je zkvalitnění podmínek pro výuku přírodovědných a technických předmětů a navázání hlubší spolupráce se středními školami a gymnázii, které mohou pořádat pravidelné výukové akce v prostorách svých moderně vybavených dílen a laboratoří.

Nejmarkantnější změny v hodnocení jsou u 2. stupně v jazykových předmětech, kde se celkové hodnocení skokově zlepšuje po školním roce 2008/2009. Uspokojivé vysvětlení tohoto pozorování není zřejmé a opět by záviselo na konzultaci s odpovědnými metodiky vzdělávání.

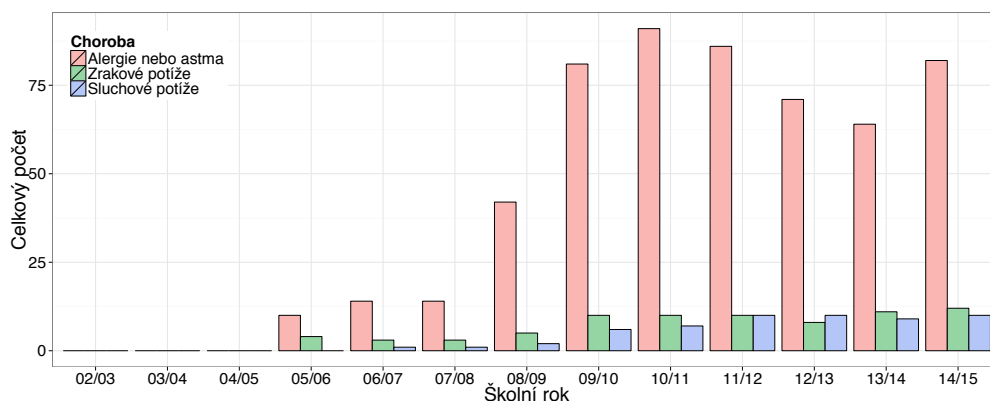
Sadu grafů uzavírají přehledy adaptace žáků z 1. na 2. stupeň základní školy, resp. jejich projevy v celkovém závěrečném hodnocení všech, nebo jen přírodovědných předmětů. Z těchto přehledů vyplývá patrně správně nastavené škálování náročnosti probírané látky od nasazení současného vzdělávacího programu *Studánecká cesta* v roce 2008.

---

<sup>3</sup>reg. č.: CZ.1.07/1.1.00/44.0012

### 4.3.3 Sledování nečíselných charakteristik

Pro ilustraci možností použití rozkladu charakteristik do ekvivalenčních tříd byly analyzovány celkové počty některých typů evidovaných zdravotních potíží u žáků. Výběry byly realizovány aplikací jednoduchých regulárních výrazů na textový sloupec CHOROBY, kde jsou nejčastěji zdravotní problémy uvedeny.



Obr. 4.1: Graf – evidované počty chorob daných kategorií

Do kategorie *Alergie nebo astma* spadají všechny položky vyhovující výrazu `alergie|astma|pyl|atop`, pro *Zrakové potíže* výrazu `očn|zrak` a pro *Sluchové potíže* obdobně `nedosl|sluch|uch`. Výrazy odpovídají nejčastějším částem slov specifikujících konkrétní zdravotní skutečnost.

Výsledný graf ukazuje skokový nárůst a kolísající trend alergických a astmatických zdravotních potíží, také ale znázorňuje pomalu rostoucí počty žáků s potížími zrakovými i sluchovými.

### 4.3.4 Hledání kauzality v datech – vliv přítomnosti asistenta pedagoga na celkový výkon třídy

Asistent pedagoga je dospělou osobou, běžně přítomnou ve třídě během vzdělávacího procesu, zajišťující nezbytné personální služby žákům se zdravotním postižením, nebo žákům se sociálním znevýhodněním.

K dispozici byl pouze omezený vzorek testovaných ročníků, neboť Základní škola Pardubice – Studánka elektronicky v IS Bakaláři eviduje třídy s asistentem pedagoga teprve od šk. roku 2013/14, přestože služeb asistentů využívá mnohem déle.

Zajímavým ukazatelem je ale pozorovaná závislost celkového výkonu třídy na přítomnosti či absenci asistenta pedagoga. Během průzkumu vlivu přítomnosti asistenta pedagoga na celkový výkon, případně výkon v určitých typech předmětů, shrnutý v přehledu v příloze E, bylo prokázáno, že přítomnost asistenta hraje u mnohých ročníků statisticky významnou, nebo dokonce statisticky velmi významnou roli.

Zejména na druhém stupni se projevuje dosahováním lepších závěrečných výsledků třídních kolektivů, což je podle konzultací se školní psycholožkou dáno prostou přítomností další dospělé osoby během vzdělávacího procesu, eliminující nežádoucí nadbytek rušivých vlivů, nepřímo tak garantující vyšší soustředění jednotlivých žáků na vzdělávací proces.

Ačkoliv je vliv asistenta zdánlivě zanedbatelný na 1. stupni v 1. – 4. ročnících a u některých tříd 6. ročníku, velmi významně přispívá lepším výsledkům v 5. ročnících, především v přírodovědných předmětech. Odůvodnění tohoto zdánlivého paradoxu, proč není většího rozporu pozorováno mezi 5. a 6. ročníky, ale je ho právě mezi 4. a 5. ročníky, lze najít ve vzdělávacích plánech ŠVP. Především předmět *Matematika* je koncipován tak, aby vyššího nárůstu náročnosti látky bylo dosaženo mezi 4. a 5. ročníkem [9, s. 151, s. 157, s. 439], umožňující tak žákům určité odlehčení v náročnosti probírané látky ve prospěch jejich snazší adaptace z 1. stupně na 2. stupeň základní školy, který je součástí 6. ročníku.

## ZÁVĚRY A DOPORUČENÍ

Cílem této bakalářské práce bylo pomocí obecných postupů a metodik zodpovědně přistoupit k informačnímu systému ve zvolené specifické oblasti, provést analýzu jeho dat z širšího pohledu a vhodným způsobem interpretovat takto získané výsledky.

Během práce bylo důsledně dodržováno postupu a obecných metodik nastíněných v úvodní teoretické kapitole. Ve zvolené oblasti českého školství byla popsána podoblast administrativy školní agendy a vedení pedagogické dokumentace s definicemi vztahů mezi jednotlivými dokumenty. Tím byla vymezena rámcová představa o informacích, se kterými školní agenda pracuje, a pro které je vhodné jejich digitální zpracování některým z dostupných informačních systémů.

Na zvolený konkrétní informační systém Bakaláři bylo nahlíženo z více různých pohledů, pro získání kompletního přehledu o jeho rozšířenosti, nabízených službách pro vedení školní agendy a pro podmínky nutné k jeho provozu z hlediska manipulace s daty.

V rámci vyšetřování celkové rozšířenosti systému byla v příloze B různými automatizovanými metodami zkoumána rozšířenost jednoho z jeho modulů, webové aplikace. Každá z použitých metod využívala jiné úrovně počáteční znalosti. Přestože nashromážděné celkové výsledky si byly velmi podobné, vyplynuly na povrch nedostatky jedné z metod, související s příliš mělkou úrovní znalostí prostředí, na čemž byla demonstrována nepřesnost interpretace výsledků z ní pramenící.

Následně byla provedena příprava k samotné analýze dat. Jelikož během procesu kompletního poznávání systému došlo k odhalení sebepopisujícího charakteru datových tabulek a funkce slovníku (viz 3.4), nebylo nutné provádět žádnou významnou dokumentaci datových souborů. Z podobného důvodu nedošlo ani ke konkrétnímu návrhu relačního modelu, neboť by nebyl žádným způsobem nezbytný pro snadné získání souvisejících dat použitým analytickým nástrojem.

Na konsolidované formě dat byla provedena analýza dle specifikovaného rozsahu. Byla obsažena analýza nečíselných dat dle metodik popsaných v úvodu práce, při souhrnu změn v hodnocení žáků bylo použito vhodných grafických reprezentací, začleněných v příloze C. V rámci hledání kauzalit a korelací byla pozornost zaměřena na vliv přítomnosti asistenta pedagoga na celkové výsledky žáků, souhrn provedených testů je k dispozici v příloze E.

Jelikož byly během analýzy vlivu asistenta pedagoga na výkon třídy (viz 4.3.4) patrně objeveny implicitní, opakující se vzory, mající svou jednoznačnou interpretaci vycházející z ŠVP, bylo by možné ve spolupráci s odpovědnými metodiky vzdělávání najít konkrétní podobu ukazatele míry adaptace žáků z 1. na 2. stupeň základní školy a sestavit specifikace parametrizovaných testů pro kvantitativně-kvalitativní verifikaci ŠVP, které by se mohly v budoucnu stát součástí sady autoevaluačních nástrojů školy.

Vzhledem k omezenému využívání informačního systému Bakaláři na konkrétní základní škole, kde byla analýza prováděna, nebylo možné zatím získat jiný charakter výsledků. Ačkoliv je v daném systému velké množství informací, které mají svou hodnotu, metody jejich získávání jsou si velmi podobné těm, které byly zahrnuty v této práci. I přesto by však bylo možné využít automatizovaných úloh k provádění popsaných analýz pro rozličné množství kombinací sledovaných charakteristik a pro předání jejich výsledků buď vedení školy, nebo odpovědným pedagogům.

Celkový cíl práce byl naplněn potvrzením, že za dodržení metodického a zodpovědného přístupu, souvisejícího s poznáním nebo se znalostí prostředí, je možné provést analýzu dat téměř libovolného informačního systému a využít získaných výsledků k vyvození závěrů vymezených kontextem znalostní domény.

## LITERATURA

- [1] Zákon č. 563/1991 Sb., o účetnictví. 1991.
- [2] Zákon č. 101/2000 Sb., o ochraně osobních údajů. 2000.
- [3] Zákon č. 248/2002 Sb., Zákon, kterým se mění zákon č. 564/1990 Sb., o státní správě a samosprávě ve školství, ve znění pozdějších předpisů, a některé další zákony. 2002.
- [4] Zákon č. 561/2004 Sb., o předškolním, základním, středním, vyšším a jiném vzdělávání. 2004.
- [5] Vyhláška č. 364/2005 Sb., o vedení dokumentace škol a školských zařízení a školní matriky a o předávání údajů z dokumentace škol a školských zařízení a ze školní matriky (vyhláška o dokumentaci škol a školských zařízení). 2005.
- [6] Webová prezentace školy. *ZŠ Pardubice – Studánka*. [online]. 2008- [cit. 2015-04-10].  
URL <http://www.zs-studanka.cz>
- [7] Program Bakaláři – Diskusní fórum. *Metodický portál RVP*. [online]. 2009-05-27 [cit. 2014-12-17]. ISSN 1802-4785.  
URL <http://diskuze.rvp.cz/viewtopic.php?f=329&t=343>
- [8] Zákon č. 472/2011 Sb., kterým se mění školský zákon. 2011.
- [9] Školní vzdělávací program Studánecká cesta. *ZŠ Pardubice – Studánka*. [online]. 2013- [cit. 2015-04-07].  
URL <http://www.zs-studanka.cz/svp.html>
- [10] Bakaláři software, s. r. o.: Webová prezentace projektu. [online]. [cit. 2015-01-10].  
URL <http://www.bakalari.cz>
- [11] SeznamŠkol.eu. *Seznam škol v ČR*. [online]. just4web.cz s.r.o., 2014 [cit. 2015-03-17].  
URL <http://www.seznamskol.eu>
- [12] Výkonové ukazatele. *Statistická ročenka školství*. [online]. Ministerstvo školství, mládeže a tělovýchovy, odbor analyticko-statistický, 2015 [cit. 2015-03-17].  
URL <http://toiler.uiv.cz/rocenka/rocenka.asp>

- [13] Internet Information Services. *Wikipedia, The free encyclopedia*. [online]. 2001- [cit. 2015-03-18].  
URL [http://en.wikipedia.org/wiki/Internet\\_Information\\_Services](http://en.wikipedia.org/wiki/Internet_Information_Services)
- [14] AtlasŠkolství.cz. *Vysoké školy, střední školy, ZŠ a jazykové školy v celé ČR*. [online]. P.F. art, spol. s r.o., 2015 [cit. 2015-03-17].  
URL <http://www.atlasskolstvi.cz>
- [15] HackForum - Bakalaři. *SOOM.cz - Hacking & ICT security*. [online]. SOOM.cz. 2014-10-19 [cit. 2015-03-21]. ISSN 1804-7270.  
URL <http://www.soom.cz/hack-forum/44564--Bakalari>
- [16] ALVARES, F.: *The History of FoxPro*. [online]. [cit. 2015-03-15].  
URL <http://www.foxprohistory.org>
- [17] BAKALÁŘI software, s. r. o.: Jak zprovoznit systém Bakaláři s uložením dat na SQL serveru. *Nápověda systému Bakaláři*. [online]. [cit. 2015-01-10].  
URL [https://skola.bakalari.cz/napoveda/index.html?ss\\_jak\\_zprovoznit\\_system\\_bakalari.htm](https://skola.bakalari.cz/napoveda/index.html?ss_jak_zprovoznit_system_bakalari.htm)
- [18] BAKALÁŘI software, s. r. o.: Podporované operační systémy. *Nápověda systému Bakaláři*. [online]. [cit. 2015-01-10].  
URL [https://skola.bakalari.cz/napoveda/index.html?ss\\_podp\\_os.htm](https://skola.bakalari.cz/napoveda/index.html?ss_podp_os.htm)
- [19] BAKALÁŘI software, s. r. o.: Použití v počítačové síti. *Nápověda systému Bakaláři*. [online]. [cit. 2015-02-12].  
URL [https://skola.bakalari.cz/napoveda/index.html?ss\\_pouziti\\_v\\_pocitacove\\_siti.htm](https://skola.bakalari.cz/napoveda/index.html?ss_pouziti_v_pocitacove_siti.htm)
- [20] BAKALÁŘI software, s. r. o.: Webová aplikace - Systémové požadavky. *Nápověda systému Bakaláři*. [online]. [cit. 2015-03-17].  
URL [https://skola.bakalari.cz/napoveda/index.html?wa\\_instal\\_systpoz.htm](https://skola.bakalari.cz/napoveda/index.html?wa_instal_systpoz.htm)
- [21] BASL, J.: Informační systémy škol - specifická oblast využití manažerských informačních systémů. *Ikaros, elektronický časopis o informační společnosti*. [online]. Ikaros. 2006, ročník 10, číslo 12 [cit. 2015-01-07]. ISSN 1212-5075.  
URL <http://www.ikaros.cz/node/3736>
- [22] BERGER, T.: WOW64: Memory mapping of 32bit apps running on a 64bit Windows. *The Citrix Blog*. [online]. 2011-10-13 [cit. 2015-03-30].  
URL <http://blogs.citrix.com/2011/10/13/wow64-memory-mapping-of-32bit-apps-running-on-a-64bit-windows/>

- [23] BOGUE, R. L.: How NetBIOS name resolution really works. *TechRepublic*. [online]. 2003-03-11 [cit. 2015-03-18].  
URL <http://www.techrepublic.com/article/how-netbios-name-resolution-really-works/>
- [24] CALDOVÁ, J.: *Informační systém střední školy*. Závěrečná práce. Mendelova univerzita v Brně, Institut celoživotního vzdělávání. Brno. 2007.
- [25] CARRABIS, J.-D.: *dBASE III Plus Networking and Multi-User Systems*. Portland, Oregon: Management Information Source. 1987. ISBN 978-09-435-1826-8. 270 s.
- [26] DOSTÁL, J.: *Školní informační systémy*. Olomouc: Univerzita Palackého v Olomouci. 2011. ISBN 978-80-244-2784-3. 68 s.
- [27] DVOŘÁK, D.: Bakaláři na Wíně. *Linux ve škole - web pro linuxové učitele*. [online]. 2012-02-12 [cit. 2015-01-02].  
URL <http://linuxveskole.cz/bakalari-na-wine/>
- [28] GULLO, F.: From Patterns in Data to Knowledge Discovery: What Data Mining Can Do. *Physics Procedia*. [online]. 2015, vol. 62, s. 18-22 [cit. 2015-04-02]. doi:10.1016/j.phpro.2015.02.005.  
URL <http://dx.doi.org/10.1016/j.phpro.2015.02.005>
- [29] HICKS, T. E.: *Creating A FoxPro Application*. Tutorial. Trinity University, Computer Science Department. San Antonio, Texas. 2003.  
URL <http://www.cs.trinity.edu/~thicks/Tutorials/FoxPro-G-CreateApplication/FoxPro-G-CreateApplication.html>
- [30] JELÍNEK, L.: *Školní agenda*. [online]. Základní příručka. Bakaláři Software s.r.o.. Hradec Králové. 2014 [cit. 2015-03-22].  
URL <http://www.bakalari.cz/ucebnice/zakladniPrirucka.pdf>
- [31] JÁGR, M.: *Informační systémy školy a jejich možnosti nasazení*. Bakalářská práce. Univerzita Karlova v Praze, Pedagogická fakulta. Praha. 2013 [cit. 2015-03-18].  
URL <https://is.cuni.cz/webapps/zpp/detail/133014/13755126>
- [32] KALHOUS, Z.; OBST, O.: *Školní didaktika*. Praha: Portál. 2002. ISBN 978-80-717-8253-7. 447 s.
- [33] KASPER, T.; KASPEROVÁ, D.: *Dějiny pedagogiky*. Praha: Grada Publishing a.s.. 2008. ISBN 978-80-247-2429-4. 224 s.



- [34] KOCAN, M.: Co vlastně je informační systém a jak souvisí s řízením? *Živě.cz*. [online]. 1998-11-08 [cit. 2015-04-10]. ISSN 1212-8554.  
URL <http://www.zive.cz/clanky/co-vlastne-je-informacni-system-a-jak-souvisi-s-rozenim/sc-3-a-4436/default.aspx>
- [35] KOČÍ, R.; KŘENA, B.: *Úvod do softwarového inženýrství*. Studijní opora. Vysoké učení technické v Brně, Fakulta informačních technologií. Brno. 2007.
- [36] LUKÁŠ, J.; MAREŠ, J.: *Internetová prezentace školy: posuzovací arch*. Praha: Národní ústav pro vzdělávání. 2012. ISBN 978-80-87063-58-3. 32 s.
- [37] LUSTFIELD, M.: Nginx Community. [online]. 2015-02-21 [cit. 2015-03-04].  
URL <http://wiki.nginx.org/Main>
- [38] MALÍK, M.: *Využití služby Active Directory pro správu počítačové sítě ve škole*. [online]. Diplomová práce. Univerzita Palackého v Olomouci, Pedagogická fakulta. Olomouc. 2012 [cit. 2015-01-12].  
URL <http://theses.cz/id/8jfx6w/>
- [39] MICROSOFT: Visual FoxPro ODBC Driver not included in MDAC 2.6 and later. *Microsoft Support*. [online]. 2007-03-29 [cit. 2015-02-19].  
URL <http://support.microsoft.com/en-us/kb/277772/en-us>
- [40] MICROSOFT: Configuring opportunistic locking in Windows. *Microsoft Support*. [online]. 2011-04-04 [cit. 2015-02-08].  
URL <http://support.microsoft.com/en-us/kb/296264>
- [41] MICROSOFT: Memory Limits for Windows and Windows Server Releases. *Microsoft Developer Network*. [online]. 2015-03-20 [cit. 2015-03-30].  
URL <https://msdn.microsoft.com/en-us/library/aa366778.aspx>
- [42] MICROSOFT: Visual FoxPro Data and Field Types. *Microsoft Developer Network*. [online]. 2015-03-20 [cit. 2015-04-17].  
URL [https://msdn.microsoft.com/en-us/library/aa975357\(v=vs.71\).aspx](https://msdn.microsoft.com/en-us/library/aa975357(v=vs.71).aspx)
- [43] MIHAILESCU, E. I.: MS-SQL Server on Linux? Yes, it works! *My \*nix world*. [online]. 2012-12-22 [cit. 2015-01-02].  
URL <http://mynixworld.info/2012/12/22/ms-sql-server-on-linux-yes-it-works/>
- [44] NEJEDLÁ, I.: Vnitřní klasifikační řád. *Pravidla pro hodnocení žáků*. [online]. ZŠ Pardubice - Studánka. 2005-08-31 [cit. 2015-04-15].  
URL <http://www.zs-studanka.cz/file/d2/klasifikacni-rad.pdf>

- [45] NEUMAJER, O.: Školní informační systémy. *Metodický portál RVP*. [online]. 2010-03-17 [cit. 2014-12-18]. ISSN 1802-4785.  
URL <http://clanky.rvp.cz/clanek/c/Z/8019/skolni-informacni-systemy.html>
- [46] NEUMAJER, O.: Ideál elektronické učebnice. *Česká škola - portál pro ZŠ a SŠ*. [online]. 2013-05-02 [cit. 2015-01-09]. ISSN 1213-6018.  
URL <http://www.ceskaskola.cz/2013/05/ondrej-neumajer-ideal-elektronicke.html>
- [47] PATOČKA, J.: *Evropa a doba poevropská*. Praha: Lidové noviny. 1992. ISBN 80-7106-017-8. 192 s.
- [48] PLATÓN: *Theaitétos*. 4. opr. vyd. Překlad František Novotný. Praha: Oikoymenh. 2007. Platónovy dialogy, sv. 11. ISBN 978-80-7298-294-3. 117 s.
- [49] SCAMBARY, J.; MCCLURE, S.; KURTZ, G.: *Hacking bez tajemství*. Praha: Computer Press. 2001. ISBN 80-7226-549-0. 592 s.
- [50] STUTTARD, D.: Security & obscurity. *Network Security*. [online]. 2005, vol. 2005, issue 7, s. 10-12 [cit. 2015-03-11]. doi:10.1016/s1353-4858(05)70259-2.  
URL [http://dx.doi.org/10.1016/s1353-4858\(05\)70259-2](http://dx.doi.org/10.1016/s1353-4858(05)70259-2)
- [51] TRIADA, s. r. o.: Právní subjektivita škol. *Deník veřejné správy*. [online]. 17.9.2002, OF 4/2010, [cit. 2015-01-07]. ISSN 1213-6336.  
URL <http://www.dvs.cz/clanek.asp?id=5134461>
- [52] TUPÝ, J.; JEŘÁBEK, J.; aj.: *Rámcový vzdělávací program pro základní vzdělávání*. Praha: MŠMT. 2007.
- [53] ZENDULKA, J.; aj.: *Získávání znalostí z databází*. Studijní opora. Vysoké učení technické v Brně, Fakulta informačních technologií. Brno. 2006.
- [54] ŠTEFFLOVÁ, J.: Školství má staronového strašáka: právní subjektivitu. *Učitelství noviny: týdeník pro učitele a přátele školy*. [online]. GNOSIS. Roč. 2002, č. 35 [cit. 2015-01-07]. ISSN 0139-5718.  
URL <http://www.ucitelskenoviny.cz/?archiv&clanek=2079>
- [55] ŠTEFFLOVÁ, J.: Konkurenční prostředí není vždycky plus. *Učitelství noviny: týdeník pro učitele a přátele školy*. [online]. GNOSIS. Roč. 2006, č. 18 [cit. 2015-01-07]. ISSN 0139-5718.  
URL <http://www.ucitelskenoviny.cz/?archiv&clanek=4240>

## SEZNAM SYMBOLŮ A ZKRATEK

učitel	pedagogický zaměstnanec základní nebo střední školy, gymnázia, obchodní akademie, konzervatoře nebo vyšší odborné školy; mistr odborného výcviku středního odborného učiliště
žák	dítě povinné školní docházky navštěvující základní školu; žák navštěvující střední školu, gymnázium, obchodní akademii, žák konzervatoře; učeň středního odborného učiliště; student vyšší odborné školy
správce ICT	pedagogický nebo nepedagogický zaměstnanec školy odpovědný za vedení a správu ICT prostředků školy
AD	Active Directory – implementace adresářových služeb společnosti Microsoft v prostředí systémů Windows
API	Application Programming Interface – rozhraní pro programování aplikací
CIFS	Common Internet File System – standardizovaný protokol aplikační vrstvy pro sdílení souborů v síti Internet. Je podporován systémy Microsoft Windows, UNIX a unixovými systémy, DEC, VMS a dalšími.
CMS	Content Management System – nástroj pro správu obsahu
ČŠI	Česká školní inspekce - státem zřízená organizace podléhající MŠMT zajišťující hodnocení vzdělávací soustavy ČR
DBF	označení databázových souborů typu <i>dBase</i> , s příponou <i>.dbf</i> a jim odpovídajících indexových souborů <i>.cdx</i> , <i>.idx</i> a textových a dočasných informací <i>.fpt</i> , <i>.mem</i>
DSN	Data Source Name – množina atributů popisujících přístup ke zdroji dat
DoS	Denial of Service – záměrné nebo nezáměrné přetížení služby velkým množstvím požadavků
ERP	Enterprise Resource Planning – informační systém integrující procesy související s relevantní činností podniku

GNU	GNU's Not UNIX – svobodný operační systém projektu GNU používající jádro GNU Hurd, případně Linux a některá další
GPL	GNU General Public Licence – licence pro svobodný software, umožňující jeho bezplatné používání, šíření a modifikaci za předpokladu, že i odvozený software bude šířen bezplatně
HTML	HyperText Markup Language - značkovací jazyk používaný pro tvorbu webových stránek
HTTP	HyperText Transfer Protocol – internetový protokol určený k přenosu hypertextových dokumentů HTML
ICT	Information and Communication Technologies – IKT, informační a komunikační technologie
IEEE	Institute of Electrical and Electronics Engineers – Institut pro elektrotechnické a elektronické inženýrství
IIS	Internet Information Services – Internetová informační služba, implementace HTTP serveru společnosti Microsoft
IS	Informační systém
LDAP	Lightweight Directory Access Protocol – protokol pro ukládání a přístup k datům ve stromové struktuře na adresářovém serveru
LMS	Learning Management System – systém pro řízení výuky
MIS	Manažerský informační systém
MVC	Model-View-Controller – třívrstvý návrhový model pro oddělení aplikační vrstvy od uživatelského rozhraní
MŠMT	Ministerstvo školství, mládeže a tělovýchovy
NetBIOS	Network Basic Input Output System – programové rozhraní poskytující služby relační síťové vrstvy, určené k zpřístupnění dat na vzdálených počítačích
ODBC	Open Database Connectivity – standardizované programové rozhraní pro přístup k databázovým systémům
OLAP	Online Analytical Processing – technologie uložení velkého množství dat v databázích, které je výhodné a snadno srozumitelné pro účely sledování trendů a výsledků

PGP	Pretty Good Privacy – standardizované rozhraní pro asymetrickou kryptografii
POSIX	Portable Operating System Interface – IEEE 1003 specifikace jednotného standardního aplikačního a uživatelského rozhraní zajišťující přenositelnost programů mezi různými unixovými systémy a hardwarovými platformami
RAD	Rapid Application Development – přístup pro rychlý vývoj software pomocí iterativního prototypování s důrazem na potřeby firemních procesů; vývojová prostředí podporující RAD velmi často zahrnují generátory grafického uživatelského rozhraní a generátory programového kódu
RAID	Redundant Array of Independent Disks – množina technik kombinujících jednotlivé pevné disky počítače v propojených polích, sloužících v závislosti na typu pro akceleraci nebo redundanci a zvýšení celkové spolehlivosti úložiště
REDIZO	Resortní identifikátor právnické osoby – označení identifikačního čísla ředitelství školy
RVP	Rámcový vzdělávací program – kurikulární dokument státní úrovně stanovující rámec pro konkrétní stupeň vzdělávání; je závazným dokumentem pro tvorbu ŠVP
SAS	Systém agent pro školy – informační systém pro správu školní agendy společnosti MP-Soft a.s.
SMB	Server Message Block – síťový komunikační protokol aplikační vrstvy sloužící ke sdílenému přístupu k souborům, tiskárnám, sériovým portům a pro další formy komunikace mezi uzly v síti, jedná se o konkrétní realizaci protokolu CIFS
SOU	Střední odborné učiliště
SOŠ	Střední odborná škola
SQL	Structured Query Language – strukturovaný dotazovací jazyk používaný pro práci s relačními databázemi
SSL	Secure Socket Layer – protokol poskytující zabezpečení šifrováním přenosu mezi transportní a aplikační vrstvou TCP/IP

SaaS	Software as a Service – software jako služba
SŠ	Střední škola
ŠVP	Školní vzdělávací program
URL	Uniform Resource Locator – řetězec znaků s definovanou strukturou sloužící k přesné specifikaci adresy umístění dokumentu nebo služby v Internetu
VOŠ	Vyšší odborná škola
WINE	Wine Is Not an Emulator – software poskytující binárním programům platformy Microsoft Windows rozhraní překládající systémová volání aplikačního rozhraní systému Windows na systémová volání cílového POSIX-kompatibilního systému
WINS	Windows Internet Name Service – implementace jmenného serveru v prostředí NetBIOS
WPF	Windows Presentation Foundation – součást frameworku Microsoft .NET pro tvorbu uživatelského rozhraní desktopových aplikací
WiFi	Wireless Fidelity – Označení bezdrátových sítí podle několika standardů IEEE 802.11
WoW64	Windows on Windows 64 – rozhraní systémů Windows umožňující spouštění 32bitových aplikací v 64bitovém prostředí
X11	X Window System – standardní software vytvářející grafické rozhraní unixových operačních systémů
XML	eXtensible Markup Language – obecný volně rozšiřitelný značkovací jazyk
ZŠ	Základní škola

# SEZNAM PŘÍLOH

<b>A</b>	<b>Příklad konfigurace serveru SAMBA</b>	<b>63</b>
<b>B</b>	<b>Průzkum rozšířenosti webové aplikace Bakaláři</b>	<b>65</b>
B.1	Veřejné vyhledávání . . . . .	65
B.1.1	Výhody a nevýhody metody veřejného vyhledávání . . . . .	66
B.1.2	Výsledky získané metodou veřejného vyhledávání . . . . .	68
B.2	Hlubkové prohledávání . . . . .	69
B.2.1	Víceúrovňové prohledávání . . . . .	71
B.2.2	Výhody a nevýhody hloubkového prohledávání . . . . .	73
B.2.3	Výsledky získané metodou hloubkového prohledávání . . . . .	74
B.3	Porovnání metod . . . . .	76
B.3.1	Porovnání rozdělení serverových technologií . . . . .	77
B.3.2	Porovnání rozdělení typů škol . . . . .	78
<b>C</b>	<b>Využití systému Bakaláři</b>	<b>79</b>
<b>D</b>	<b>Vývoj změn v hodnocení žáků</b>	<b>80</b>
<b>E</b>	<b>Vliv přítomnosti asistenta pedagoga na výkon třídy v ročníku</b>	<b>88</b>
<b>F</b>	<b>Obsah přiloženého CD</b>	<b>90</b>

## A KONFIGURACE SERVERU SAMBA

Následující nastavení zapsaná v konfiguračním souboru `smb.conf` mohou sloužit jako příklad pro využití POSIXového počítače v síti jako souborového serveru Samba poskytujícího programy Bakaláři uživatelům náležícím do předem definované skupiny `bakalari`. Ukázka si však neklade za cíl popsat správu oprávnění uživatelů či připojení tohoto počítače do existující domény, demonstruje především některé možnosti přispívající k bezproblémovému spouštění a běhu programů Bakaláři na připojených systémech Windows různých verzí.

```
[global]
    security = user
    workgroup = SKOLA
    # server bude dostupný i pod názvem \\BAKALARI
    netbios aliases = BAKALARI
    server role = standalone server
    ...
    deadtime = 25 # odpojení klientů po 25 minutách nečinnosti

[bakalari]
    comment = Bakaláři
    path = /server/bakalari
    follow symlinks = yes
    wide links = no
    available = yes
    browseable = yes
    public = yes
    read only = no
    writable = yes

    # Dobrovolné zamykání je povoleno všem souborům
    # s výjimkou používaných k zápisu aplikací FoxPro
    oplocks = yes
    level2 oplocks = yes
    veto oplock files = /*.mem/*.dbf/*.txt/*.log/*.fpt/
                       *.cdx/*.idx/*.err/
                       *.xml/sqltran2.*/*
```



```
# Nastavení velikosti písmen a práce s názvy souborů
case sensitive = no          # nerozlišovat velikost
default case = lower        # výchozí velikost - malá
preserve case = yes         # zachovat velikost
short preserve case = yes

# Povolí přístup k souborům s dlouhým názvem
# pod názvy 8.3 systému DOS
mangled names = yes

store dos attributes = yes # atributy systému DOS
map archive = yes         # archivní bit DOS

write list = bakalari, @bakalari
force user = bakalari
force group = bakalari

create mask = 0660
directory mask = 0770
force create mode = 0660
force directory mode = 0770
directory_mode: parameter = 0770

# ACL - přidá oprávnění pro spouštění
# na virtuálním svazku NTFS
acl allow execute always = yes
```

## B ROZŠÍŘENOST WEBOVÉ APLIKACE

Míra rozšířené webové aplikace Bakaláři byla zkoumána dvěma různými automatizovanými metodami, které se liší především kategoriemi uplatněných znalostí nutných pro získání primárních zdrojů, kterými jsou v tomto případě URL adresy konkrétních instancí webové aplikace na jednotlivých školách.

Po získání těchto adres je pak zpracování v obou uvedených metodách podobné a vychází z celkového shromáždění metrik (protokol, adresa, port, verze webového serveru, verze prostředí ASP.NET a názvu školy) pomocí programu `curl`<sup>1</sup>.

Obě metody k automatizaci využívají převážně standardní nástroje řízené pomocí skriptů pro unixový shell `bash`. Kvůli mnohým zjednodušením byla využita specifická GNU-rozšíření některých standardních nástrojů, a tedy skripty nejsou přímo bez úprav přenositelné na jiné systémy nevyházející primárně z GNU.

### B.1 Veřejné vyhledávání

První metoda využívá pouze konkrétních technických poznatků o informačním systému a služeb nabízených veřejnými internetovými vyhledávací podporujícími rozšířené a filtrované vyhledávání pomocí speciálních operátorů.

Mezi speciální vyhledávací operátory patří hledání podle obsahu titulku v hlavice webové stránky (HTML značka `<title>`). Konkrétním hledáním umožňujícím odhalit alespoň přibližný počet škol využívajících webovou aplikaci je specifikace titulku přihlašovací stránky webové aplikace Bakaláři, kterým je přesný výraz `Bakaláři - přihlášení`.

V aktuálním období, ve druhém pololetí školního roku 2014/2015, nabídly na takto specifikovaný dotaz vyhledávače Google a Seznam přibližně stejný počet výsledků. Zpracováním těchto výsledků je možné získat adresy aktivních webových aplikací a ověřit následně jejich funkčnost a parametry.

Pro strojové zpracování byly využity výsledky poskytnuté vyhledávačem Google, neboť i bez jakéhokoliv API umožňují pomocí velmi jednoduché lexikální analýzy vyjmout cílové adresy vyhledávaných přihlašovacích stránek, nacházející se vždy v osamocených blocích `<cite>` na každé stránce s výsledky vyhledávání. Je však

---

<sup>1</sup><http://curl.haxx.se>

nutné dodat, že cílová adresa URL nesmí být příliš dlouhá, jinak bude ve výsledcích zkrácena.

### B.1.1 Výhody a nevýhody metody veřejného vyhledávání

Hlavní výhodou metody využívající vyhledávání Google je především její přímočará lineární složitost a relativně rychlý zisk konkrétních adres přihlašovacích stránek instancí hledané webové aplikace.

Proces automatizovaného sběru výsledků vyhledávání je ale do jisté míry komplikován ochranou systémů Google proti útokům DoS, znemožňující okamžité získání všech výsledků najednou. Pozorováním bylo zjištěno, že lze získat přibližně 17 stránek výsledků vyhledávání v krátkém časovém sledu, který musí být následován určitou odmlkou, jež pro zajištění maximální spolehlivosti metody představuje alespoň deset minut. Toto omezení značným způsobem znemožňuje paralelizaci sběru cílových adres, je-li prováděno pouze z jedné veřejné adresy.

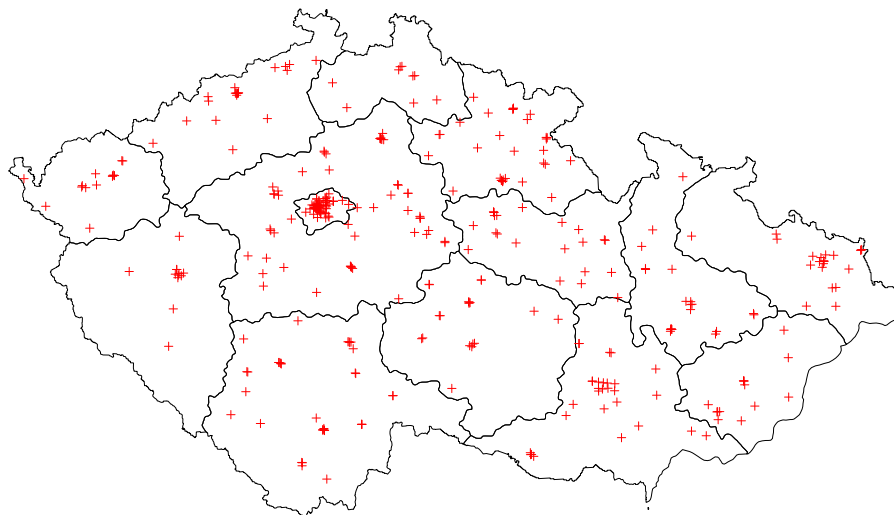
Metoda je také primárně nepřesná už jen tím, že v první řadě předpokládá, že hledaná přihlašovací stránka bude vyhledávačem indexována, čemuž může správce snadno předejít například dodatečným přidáním informačního souboru `robots.txt` do kořenového adresáře webové aplikace. Bakaláři se záznamem zakazujícím její automatické procházení indexovacími roboty. Dalším z řady úskalí je též možná neaktualnost indexovaných stránek a jí způsobená nadbytečná zátěž pro zpracování.

Dalším omezením, které by bylo třeba uvažovat v případě většího rozšíření webové aplikace mezi školami, je limitovaný počet výsledků vyhledávání, který může vyhledávač Google nabídnout. Tento limit představuje pouze prvních 1 000 výsledků vyhledávání, což by v ideálním případě (se zanedbáním nerelevantních výsledků) poskytlo maximální výsledek 18% rozšířenosti aplikace.

Bohužel samotný princip této metody, využívající pouze poznatky o konkrétním informačním systému a ne kontextová omezení stanovená jeho znalostní doménou, nepracuje s žádnou pevně danou výchozí množinou, tudíž by i ideální získané výsledky bylo možné interpretovat pouze tvrzením, že *existuje alespoň 1000 aktivních aplikací*, bez možnosti určení konkrétního podílu.

Tuto situaci ilustruje následující slepá mapa, do které byla zasazena část výsledků vyhledávání. Bez celkové představy o absolutních počtech škol a jejich geografickém

rozložení v České republice nemá tato mapa prakticky žádnou výpovědní hodnotu.



Obr. B.1: Mapa – Bezkontextová geografická distribuce škol provozujících webovou aplikaci<sup>2</sup>

Geolokační informace byly získány pomocí API Mapy.cz společnosti Seznam, dotazováním se na souřadnice názvu školy uvedeného v její odpovídající webové aplikaci Bakaláři.

---

<sup>2</sup>graficky je zachyceno pouze 54,63 % získaných výsledků

### B.1.2 Výsledky získané metodou veřejného vyhledávání

Následující tabulky shrnují výsledky získané metodou veřejného vyhledávání. Celkem bylo nalezeno 648 unikátních instancí webové aplikace Bakaláři. Celkový počet a relativní četnosti typů škol však není možné pouze podle získaného názvu přesně určit.

Tab. B.1: Podíly typů škol získané metodou veřejného vyhledávání

Typ školy	Celkový počet	Relativní četnost
Základní školy	204	28,29 %
Střední školy	301	41,75 %
Gymnázia	155	21,50 %
Konzervatoře	2	0,28 %
Vyšší odborné školy	59	8,18 %
Celkem	721 (634)	– %

Tab. B.2: Podíly technologií získané metodou veřejného vyhledávání

Server	Počet	Relativní četnost
IIS 5.1	26	4,01 %
IIS 6.0	133	20,56 %
IIS 7.0	86	13,27 %
IIS 7.5	283	43,67 %
IIS 8.0	59	9,10 %
IIS 8.5	53	8,18 %
jiné	8	1,23 %
Celkem	648	100 %

## B.2 Hlubkové prohledávání

Daleko pragmatičtější přístup, čerpající především ze získaných poznatků oboru, představuje metoda hlubkového prohledávání. Může sloužit i jako velmi jednoduchý příklad konkrétní aplikace dolování dat. Vstupním datovým souborem pro tuto metodu je adresář škol a školských zařízení, dostupný na serverech statistického odboru MŠMT<sup>3</sup> ve formátu tabulky Microsoft Excel. Tabulka obsahuje všechny relevantní záznamy o školách, včetně plného a zkráceného názvu, charakter školy podle resortního číselníku AKDT (Druhy a typy škol a školských zařízení), kontaktních údajů na ředitelství školy a uvedly-li školy adresu své webové prezentace, je tam též přítomna.

Zcela skrytou, ačít implicitní informací, je v tomto případě hledaná adresa URL webové aplikace Bakaláři dané školy. Uplatněnou oborovou znalostí je předpoklad existence odkazu na tuto aplikaci buď na první stránce webové prezentace školy, nebo v globálním navigačním oddílu celé webové prezentace. Tento předpoklad nepřímo vyplývá z doporučení Národního ústavu pro vzdělávání a jeho posuzovacího archu pro internetovou prezentaci školy (více viz 2.2.3, [36]).

Metoda využívá systematického procházení všech webových stránek uvedených v poskytnutém adresáři škol, pátrajíc přitom po určitém vzorci. Tento hledaný vzorec, projev přítomnosti webové aplikace Bakaláři, je dle technických poznatků o webové aplikaci cílový odkaz končící `login.aspx`, který je po stažení svého obsahu dále testován na přítomnost odpovídající HTML značky `<div id='nazevskoly'>`. Je-li test pozitivní, byla nalezena přihlašovací stránka webové aplikace. Její adresa, uvedený název školy v příslušné značce a technické informace o serverovém prostředí jsou doplněny do původního vstupního souboru.

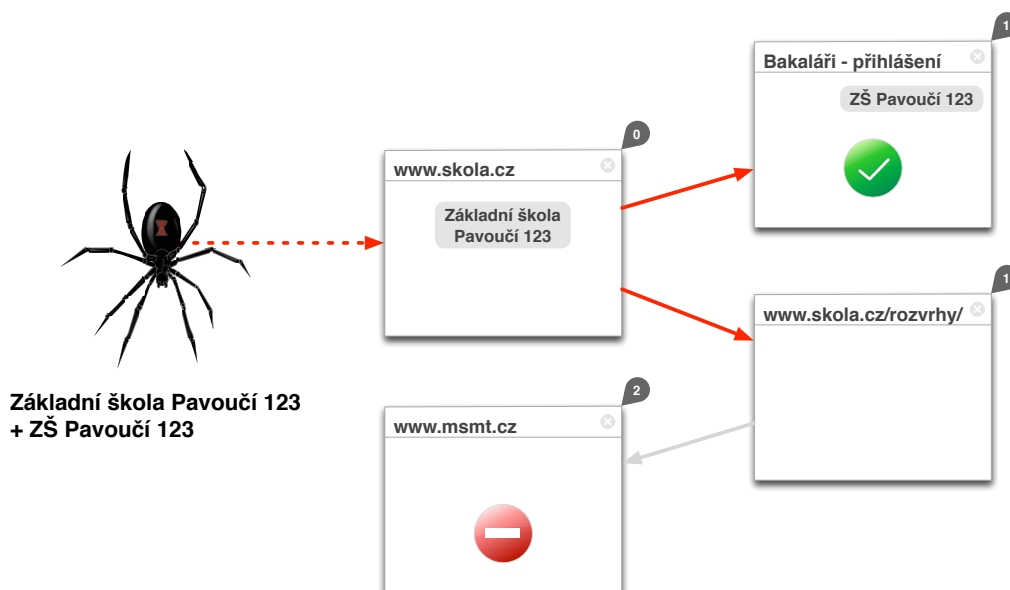
Pro snazší strojové zpracování byla tabulka formátu Microsoft Excel převedena na textový soubor CSV, oddělující textové hodnoty jednotlivých sloupců na každém řádku čárkami, v adresách URL byly opraveny některé chyby způsobené překlepy (např. `http`, `http.:`). URL adresy webových prezentací škol a požadovaná hloubka prohledávání jsou předány k procházení programu `wget`<sup>4</sup> v režimu robota, agenta kontrolujícího existenci či neexistenci obsahu na požadované adrese URL. Tento robotický režim má mnoho označení – např. *crawler*, *web spider*, *indexing bot* – protože se však jeho aktivace v programu `wget` provádí přepínačem `--spider`, bude zde označován českým slovem *pavouk*.

---

<sup>3</sup><http://stistko.uiv.cz/registr/vybskolrn.asp>

<sup>4</sup><https://www.gnu.org/s/wget>

Základní funkcí každého samostatného *pavouka* je projít všechny hypertextové odkazy, uvedené v předané adrese URL, až do požadované úrovně. V první úrovni prohledávání je kontrolována vstupní stránka samotná a všechny stránky, na které vedou odkazy ze vstupní stránky, další odkazy už procházeny nejsou. Je-li nalezen odkaz vyhovující vyhledávacím podmínkám (konečný název cílové adresy po všech přesměrováních protokolu HTTP končí řetězcem `login.aspx`), je předán programu `curl` pro další vyhodnocení obsahu a technických parametrů webového serveru.

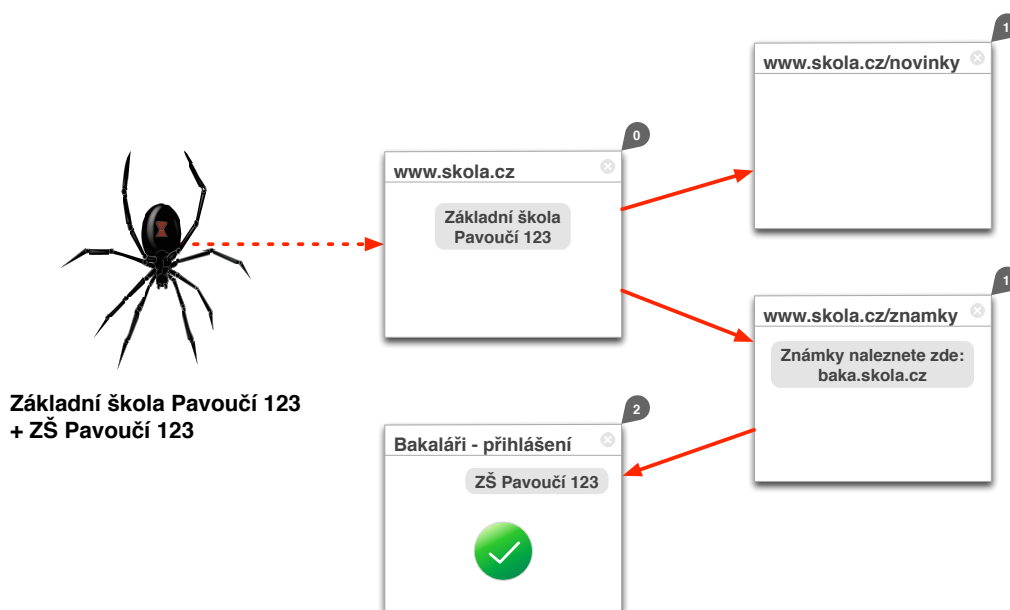


Obr. B.2: Diagram – prohledávání první úrovně

Během vyhledávání v první úrovni hloubky může docházet k určitému zkreslení, poskytuje-li škola více odkazů na jednu aplikaci ve variantách bez zabezpečení a se zabezpečením SSL. Toto zkreslení v konkrétním výsledku způsobuje celkovou chybu 0,6 % v určení typů škol, je však považováno za korektní informaci při analýze metrik použitých technologií.

## B.2.1 Víceúrovňové prohledávání

Víceúrovňovým prohledáváním je myšlena hloubka prohledávání větší, než první externí odkaz na předané počáteční adrese URL webové prezentace školy. Je tak možné najít adresy aktivní aplikace, na kterou nevedou odkazy přímo z hlavní stránky webové prezentace, případně mezistránky, které používají k přesměrování adresy internetového prohlížeče JavaScript, který *pavouk* neumí interpretovat.

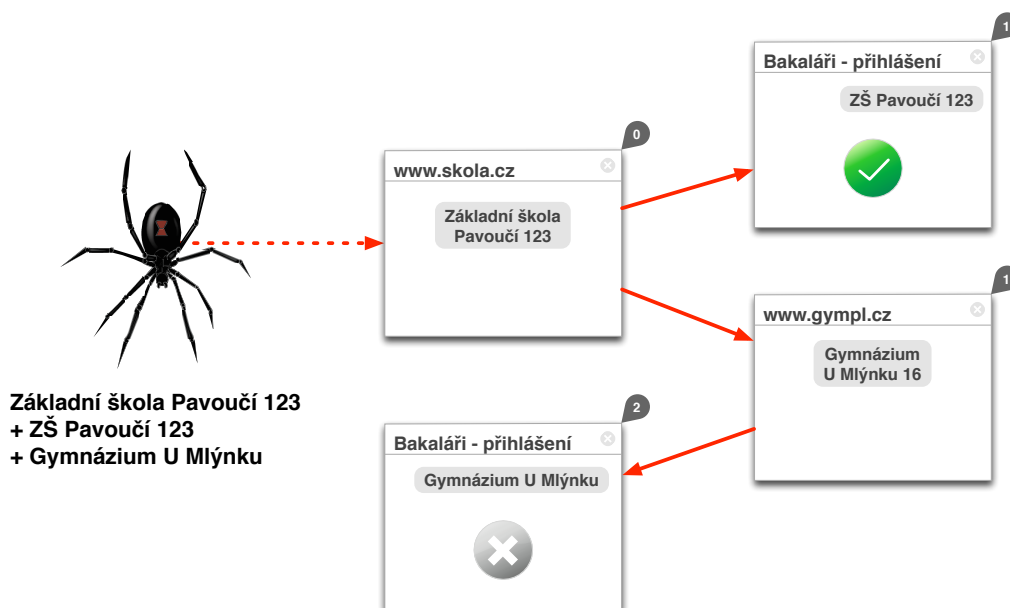


Obr. B.3: Diagram – prohledávání druhé úrovně

Daní za prohledávání vyšších úrovní je kromě dalšího rozměru časové složitosti také riziko křížení odkazů z jiných škol. Jedná se především o situace, kdy jsou jednotlivé školy zapojeny ve společném projektu, na který odkazují, nebo situace, kdy jedna škola odkazuje na jinou, partnerskou školu. Jelikož je výchozí seznam počátečních adres URL velmi rozsáhlý, nelze ho předat v dodatečném parametru pro vyloučení určitých domén prohledávání. Pracují-li navíc *pavouci* paralelně a každý z nich je v jiné fázi různých částí vstupního seznamu, nemají prakticky žádnou možnost, jak se doménám jiných škol vyhnout.



Situaci zachycuje následující diagram.



Obr. B.4: Diagram – prohledávání druhé úrovně – křížení odkazů

Projevem křížení odkazů je vznik více záznamů s duplicitním číslem REDIZO ve výsledném doplněném adresáři škol. Tyto záznamy jsou dodatečně zpracovány a podrobeny porovnání oficiálního názvu školy s názvem, který poskytla webová aplikace Bakaláři v odpovídající HTML značce. Dostatečně účinnou a rychlou metodou pro tuto porovnání se ukázalo měření Levenshteinovy vzdálenosti.

Levenshteinova vzdálenost je metrika podobnosti dvou textových řetězců na základě počtu provedených změn v jednotlivých znacích (vlození, mazání, nahrazení), nutných pro dosažení jejich úplné shodnosti. Skóre podobnosti těchto řetězců je pak poměr nezměněných znaků a celkové délky řetězce.

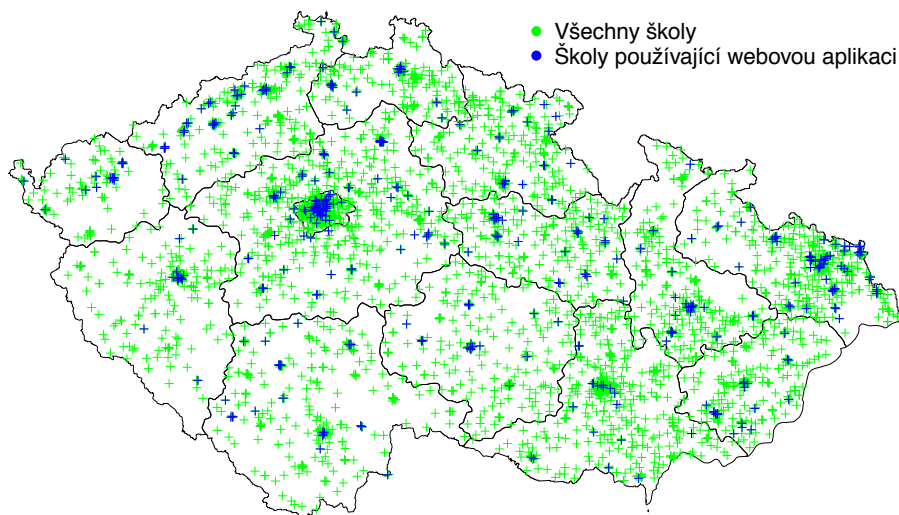
Protože instance webových aplikací Bakaláři převážně uvádí plný nebo zkrácený název školy (ať už automaticky získaný z registračních údajů licence programů Bakaláři, nebo změněný správcem do jiné podoby), je prováděno porovnání oficiálních údajů z rejstříku škol s názvem poskytnutým nalezenou instancí webové aplikace.

Empiricky bylo zjištěno, že při 73% podobnosti s plným ( $L_I \geq 0,73$ ), nebo při 63% podobnosti se zkráceným ( $L_{II} \geq 0,63$ ) názvem, případně jsou-li obě hodnoty podobnosti alespoň 60 % a rozdíl mezi nimi je menší než 0,9 % ( $(L_I \geq 0,6 \wedge L_{II} \geq 0,6) \wedge \Delta L \leq 0,009$ ), je shoda oficiálních a webovou aplikací poskytovaných názvů

dostatečná. Samotný výpočet Levenshteinovy vzdálenosti je realizován volně dostupnou knihovnou `fstrcmp`<sup>5</sup>.

## B.2.2 Výhody a nevýhody hloubkového prohledávání

Hlavní výhodou metody hloubkového prohledávání je přesně vymezený kontext, ve kterém pracuje, určený obsahem adresáře škol. Oproti metodě veřejného vyhledávání lze tedy získat přehled o míře využívání webové aplikace Bakaláři mezi všemi školami. Díky adresářem poskytovaným poštovním adresám je navíc usnadněno jejich přesné geografické označení.



Obr. B.5: Mapa – Geografický kontext škol s webovou aplikací

Další z nabízených výhod je velice snadná paralelizace prohledávání. Protože adresy jednotlivých škol spolu jen zřídka souvisí, je omezení maximálního počtu *paravouků*, pracujících na rozděleném vstupním seznamu, stanoveno pouze technickými parametry počítače, ze kterého jsou úlohy spouštěny, a celkovou propustností internetového připojení.

Naopak velkou nevýhodou je exponenciální časová složitost, neboť vyhledávání není přímočaré, ale musí obsáhnout všechny adresy na dané úrovni, včetně prodlev způsobených nefunkčními hypertextovými odkazy, nebo hypertextovými odkazy vedoucími do privátních adresních rozsahů. V závislosti na množství paralelních vláken

<sup>5</sup><http://fstrcmp.sourceforge.net>

může prohledávání první úrovně trvat desítky minut, prohledávání druhé úrovně hodiny až desítky hodin, prohledávání další úrovně už téměř týden. Při vyšších úrovních už navíc není možné dostatečně spolehlivě předcházet či dodatečně filtrovat křížení odkazů. Jediný způsob, který by umožnil tuto nevýhodu eliminovat, by bylo začlenění strojově čitelného čísla REDIZO do stránek webové aplikace Bakaláři.

## B.2.3 Výsledky získané metodou hloubkového prohledávání

### Výsledky prohledávání I. úrovně

Shrnutí výsledků z první úrovně prohledávání je uvedeno v následujících tabulkách, klasifikujících typy jednotlivých škol podle resortního číselníku AKDT (Druhy a typy škol a školských zařízení). Kódy z číselníku jsou součástí vstupního adresáře škol.

Chyba, způsobená duplicitními odkazy pro varianty bez a se zabezpečením SSL, činila 0,6 % z celkového množství nalezených výsledků.

Tab. B.3: Podíly typů škol výběru z číselníku AKDT (Druhy a typy škol a školských zařízení) získané hloubkovým prohledáváním I. úrovně

Kód	Význam	Celkový počet škol	Počet web. ap.	Podíl
<i>Bxx</i>	Základní školy	4 145	181	4,36 %
<i>Cxx</i>	Střední školy a gymnázia	1 325	322	24,30 %
<i>C5x</i>	Pouze gymnázia	360	107	29,72 %
<i>Dxx</i>	Konzervatoře	18	1	5,55 %
<i>Exx</i>	Vyšší odborné školy	175	37	21,14 %
–	Celkem	5 402	496	9,18 %

Tab. B.4: Podíly technologií získané metodou hloubkového prohledávání I. úrovně

Server	Počet	Relativní četnost
IIS 5.1	15	2,99 %
IIS 6.0	90	17,93 %
IIS 7.0	60	11,95 %
IIS 7.5	220	43,82 %
IIS 8.0	51	10,16 %
IIS 8.5	54	10,76 %
jiné	6	1,23 %
Celkem	496	100 %

### Výsledky prohledávání II. úrovně

Obdobně jsou shrnuty výsledky prohledávání druhé úrovně. K určité chybě, popsané u případu prohledávání první úrovně, je připočtena chyba špatné korekce názvu školy. Celková chyba činí 0,91 %.

Tab. B.5: Podíly typů škol výběru z číselníku AKDT (Druhy a typy škol a školských zařízení) získané hloubkovým prohledáváním II. úrovně

Kód	Význam	Celkový počet škol	Počet web. ap.	Podíl
<i>Bxx</i>	Základní školy	4 155	267	6,42 %
<i>Cxx</i>	Střední školy a gymnázia	1 383	500	36,15 %
<i>C5x</i>	Pouze gymnázia	384	173	45,05 %
<i>Dxx</i>	Konzervatoře	18	1	5,55 %
<i>Exx</i>	Vyšší odborné školy	187	68	36,36 %
–	Celkem	5 471	761	13,90 %

Tab. B.6: Podíly technologií získané metodou hloubkového prohledávání II. úrovně

Server	Počet	Relativní četnost
IIS 5.1	27	3,51 %
IIS 6.0	146	18,99 %
IIS 7.0	99	12,87 %
IIS 7.5	335	43,56 %
IIS 8.0	67	8,71 %
IIS 8.5	79	10,27 %
jiné	8	1,04 %
Celkem	761	100 %

### B.3 Porovnání metod

Obě popsané metody používají odlišné úrovně znalostí k získání vstupních dat pro zpracování. Z absolutních i relativních četností sledovaných znaků poskytují přibližně stejné výsledky, velmi se však liší v šířce získaných a interpretovatelných informací.

Tab. B.7: Celkové metriky použitých metod

Ozn.	Metoda	Složitost	Počet výsledků	Chyba
$G$	Vyhledávání	lineární $O = c \cdot n$	648	neznámá
$P_I$	Prohledávání I	exponenciální $O = t_I \cdot c \cdot n^m$	496	0,60 %
$P_{II}$	Prohledávání II	exponenciální $O = t_{II} \cdot c \cdot n^{m \cdot p}$	761	0,91 %

$c$  – průměrná doba navazování spojení, získávání a zpracování obsahu,  $n$  – počet vstupních adres,  $m$ ,  $p$  – počet větví I., resp. II. úrovně,  $\frac{1}{t}$  – počet paralelních vláken

Pro následující provedené Pearsonovy chí-kvadrát testy byly výsledky získané metodou prohledávání druhé úrovně, označené jako  $P_{II}$ , stanoveny jako referenční pro porovnání distribuce pravděpodobnosti výskytu sledovaných znaků. Metody  $P_I$  i  $G$  poskytly shodné výsledky rozdělení technických parametrů – serverového software. Při sledování rozdělení typů škol se však výsledky metody  $G$  od referenčních významně liší, neboť pro správnou detekci typu školy neposkytuje tato (pouze na systém orientovaná) metoda dostatek rozlišitelných informací.

### B.3.1 Porovnání rozdělení serverových technologií

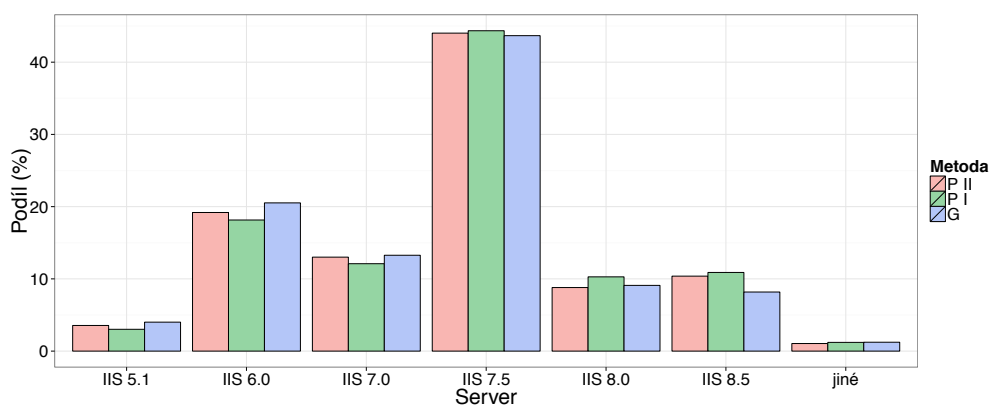
Hypotéza  $H_0$ : Rozdělení technologií nalezených danou metodou odpovídá vztažnému rozdělení metody  $P_{II}$ .

Hypotéza  $H_1$ : Rozdělení technologií nalezených danou metodou se liší od vztažného rozdělení metody  $P_{II}$ .

Tab. B.8: Porovnání metod – podíly serverových technologií

Metoda/IIS	5.1	6.0	7.0	7.5	8.0	8.5	jiné	$\Sigma$
$P_{II}$ ref.	27	146	99	335	67	79	8	761
$P_{II}$ %	3,55 %	19,19 %	13,01 %	44,02 %	8,80 %	10,38 %	1,05 %	100 %
$P_I$	15	90	60	220	51	54	6	496
$G$	26	133	86	283	59	53	8	648
Oč. $P_I$	17,60	95,16	64,53	218,34	43,67	51,49	5,21	–
Oč. $G$	22,99	124,32	84,30	285,26	57,05	67,27	6,81	–
$\tilde{\chi}_{P_I}^2$	0,38	0,28	0,32	0,01	1,23	0,12	0,12	2,46
$\tilde{\chi}_G^2$	0,39	0,61	0,03	0,02	0,07	3,03	0,21	4,35

Stupňů volnosti je  $N - 1 = 6$ , pro hladinu významnosti  $\alpha = 5 \%$  je kritická hodnota  $\chi^2 = 12,59$ . U obou porovnávaných metod,  $P_I$  i  $G$ , nelze hypotézu  $H_0$  zamítnout, získaná rozdělení serverových technologií jsou si u všech metod podobná, vzorky jsou reprezentativní.



Obr. B.6: Graf – rozdělení serverových technologií získané různými metodami

### B.3.2 Porovnání rozdělení typů škol

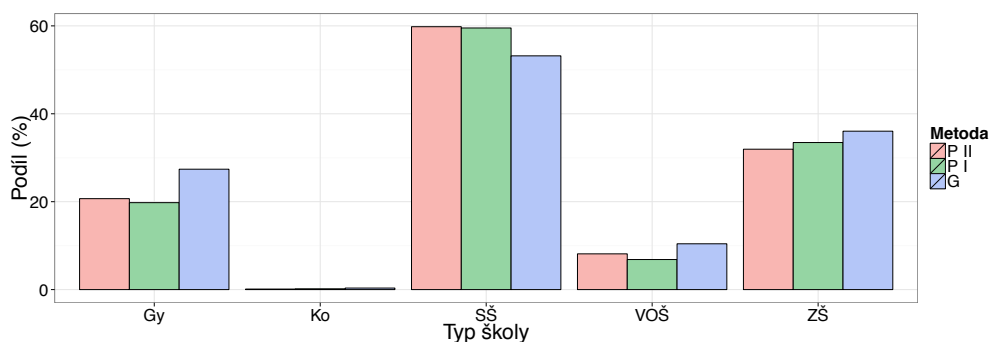
Hypotéza  $H_0$ : Rozdělení typů škol nalezených danou metodou odpovídá vztažnému rozdělení metody  $P_{II}$ .

Hypotéza  $H_1$ : Rozdělení typů škol nalezených danou metodou se liší od vztažného rozdělení metody  $P_{II}$ .

Tab. B.9: Porovnání metod – podíly typů škol

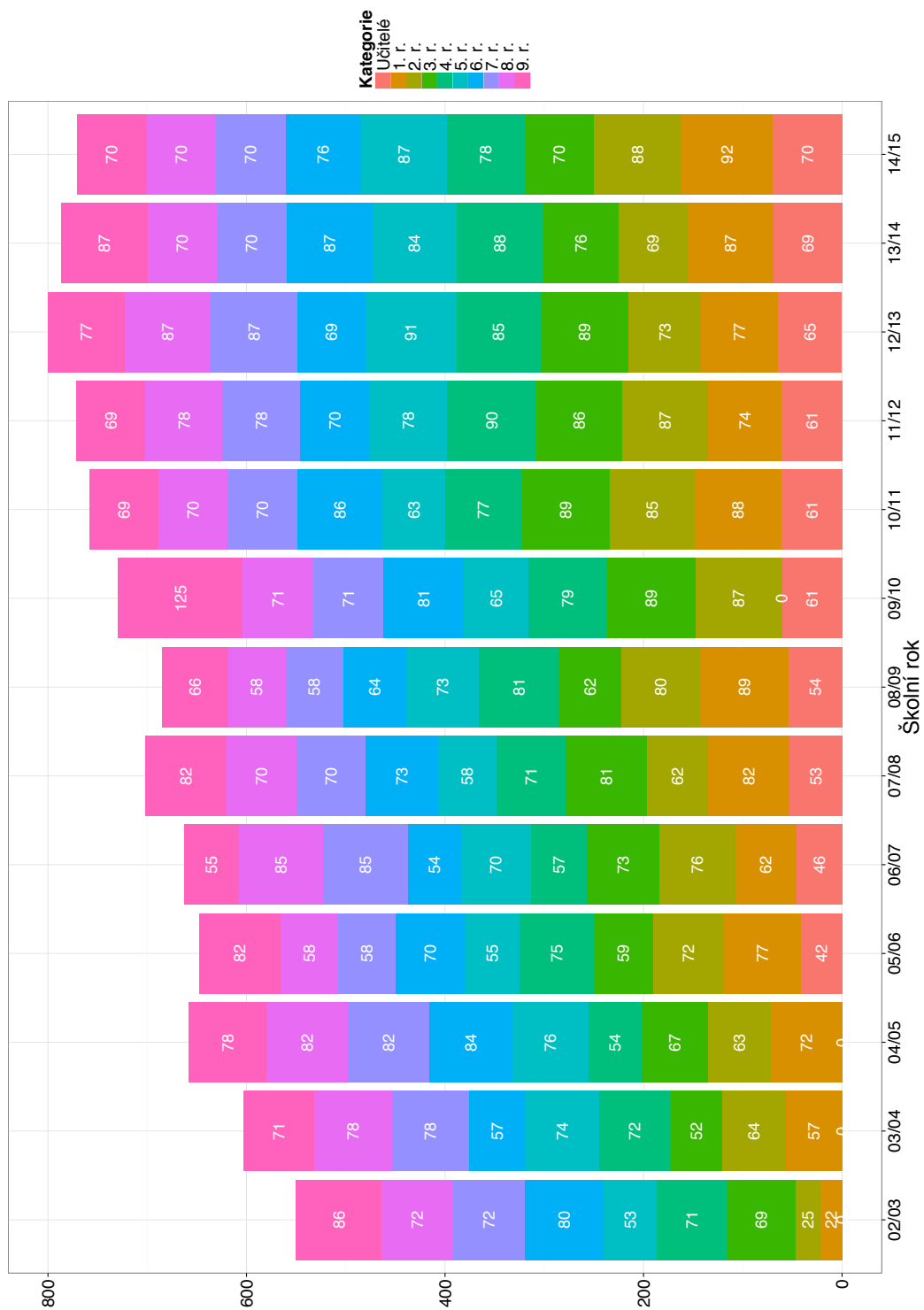
Metoda/typ	ZŠ	SŠ	Gy	Ko	VOŠ	$\Sigma$
$P_{II}$ ref.	267	500	173	1	68	836
$P_{II}$ %	31,94 %	59,81 %	20,69 %	0,12 %	8,13 %	100 %
$P_I$	181	322	107	1	37	541
$G$	204	301	155	2	59	566
Oč. $P_I$	172,78	323,56	111,95	0,65	44,00	–
Oč. $G$	180,77	338,52	117,13	0,68	46,04	–
$\tilde{\chi}_{P_I}^2$	0,39	0,01	0,22	0,19	1,12	1,92
$\tilde{\chi}_G^2$	2,99	4,16	12,25	2,59	3,65	25,62

Stupňů volnosti je  $N - 1 = 5$ , pro hladinu významnosti  $\alpha = 5 \%$  je kritická hodnota  $\chi^2 = 9,49$ . U metody  $P_I$  nelze hypotézu  $H_0$  zamítnout, získaná rozdělení typů škol jsou podobná referenčnímu vzorku. U metody  $G$  je hypotéza  $H_0$  zamítnuta, rozdělení typů škol neodpovídá referenčnímu vzorku. Příčinou je nedostatečná přesnost rozpoznávání typu školy pouze z jejího názvu, bez použití kódů z resortního číselníku AKDT (Druhy a typy škol a školských zařízení).



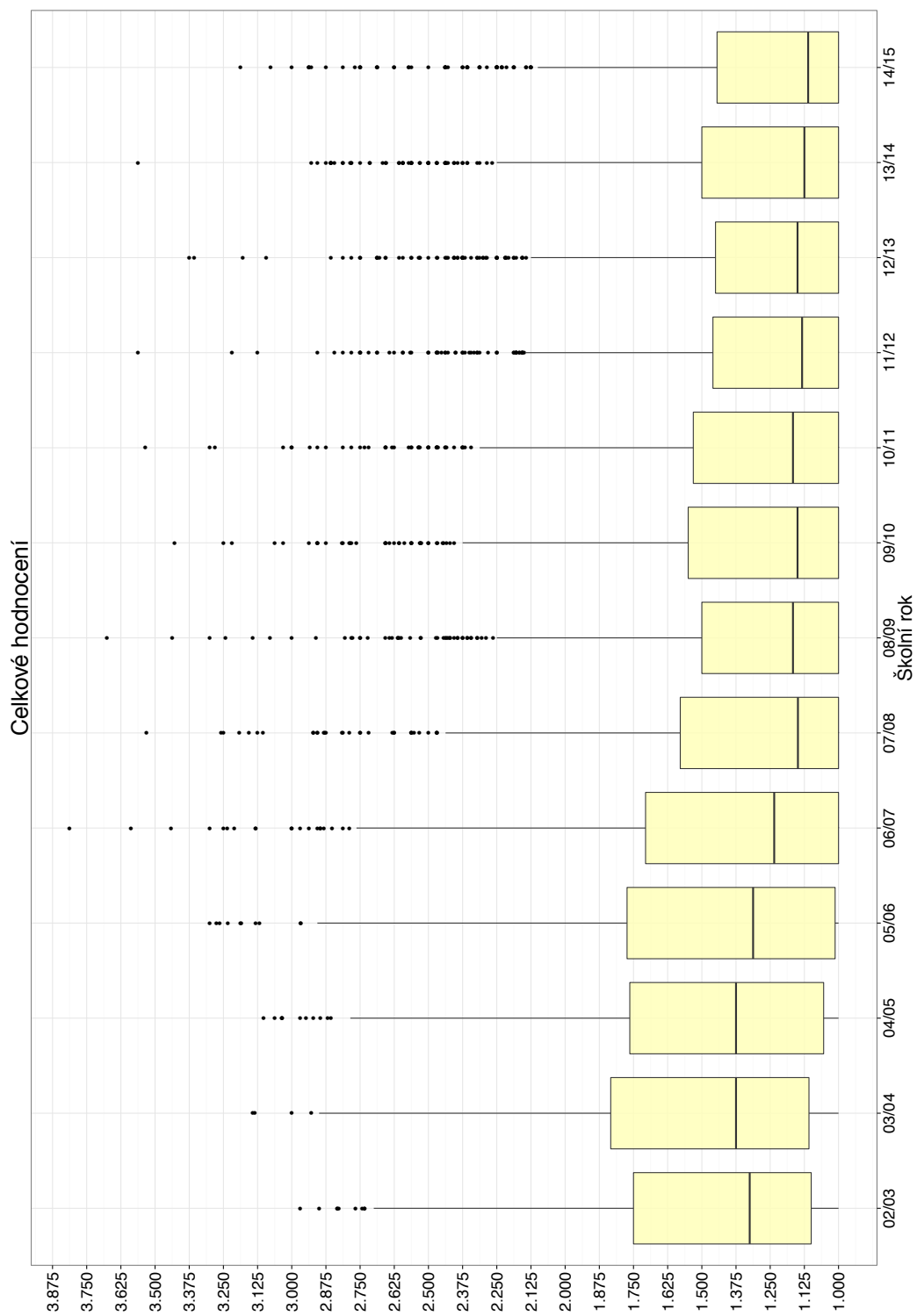
Obr. B.7: Graf – rozdělení typů škol získané různými metodami

## C VYUŽITÍ SYSTÉMU BAKALÁŘI

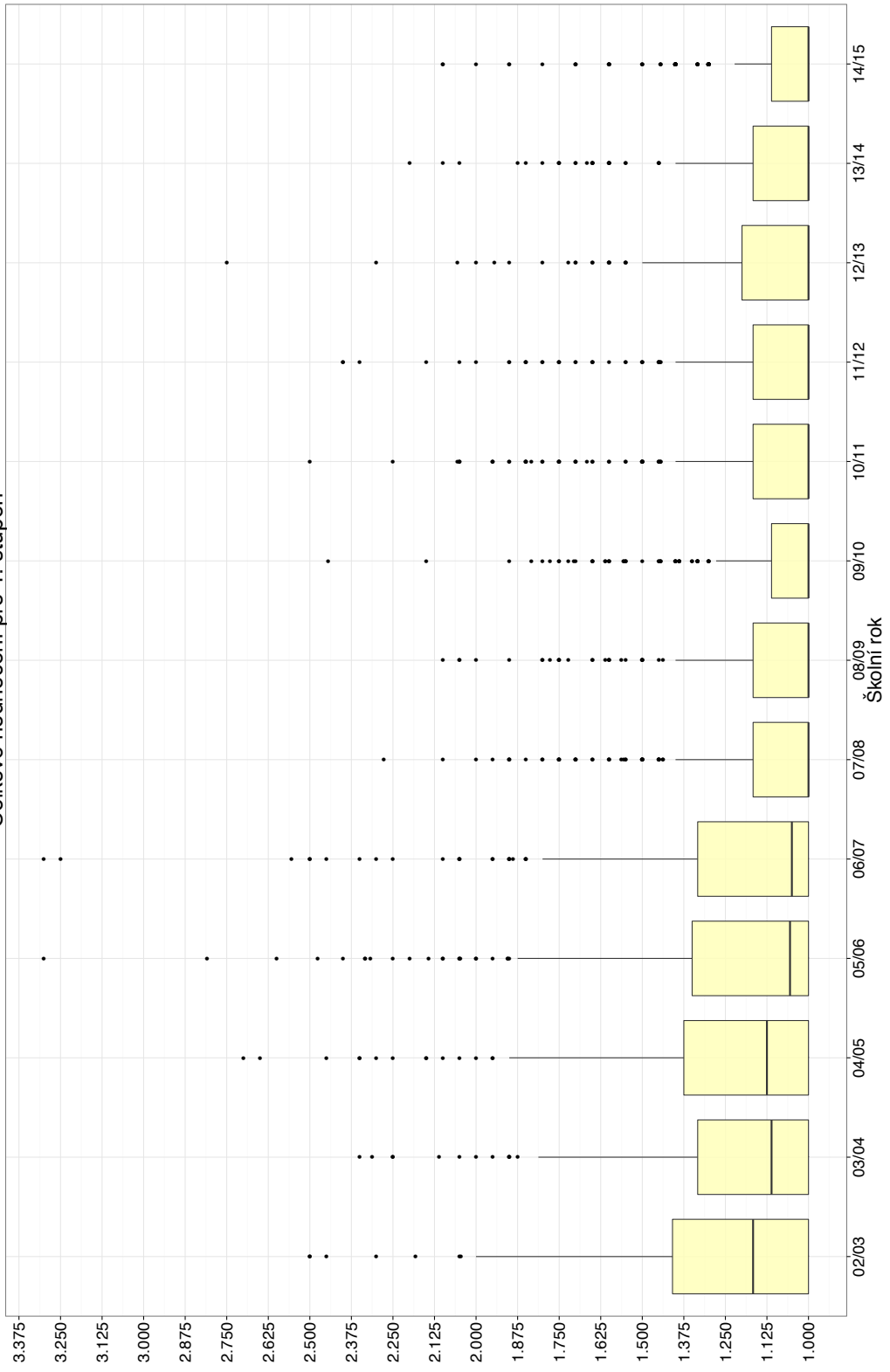


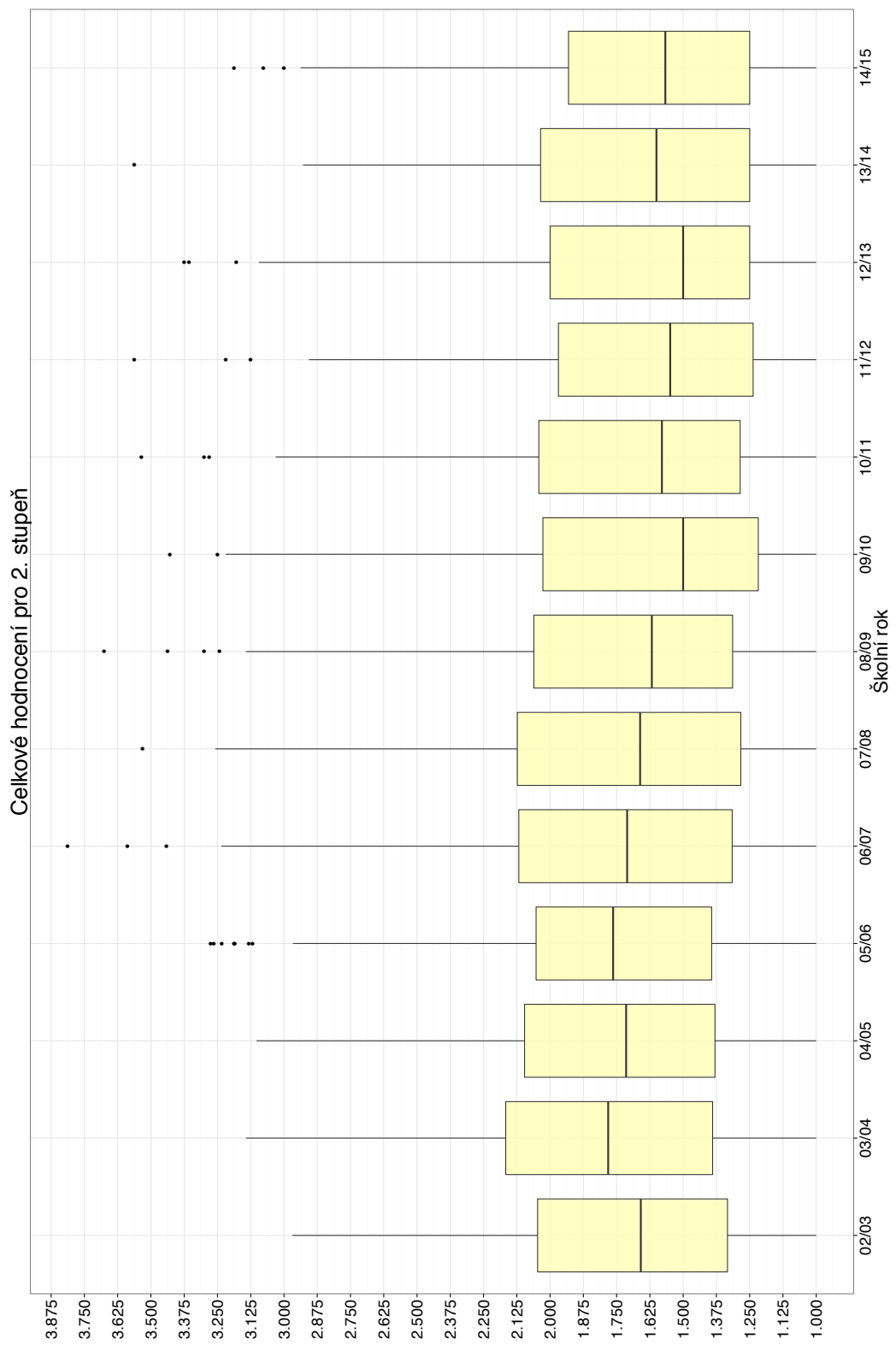


## D VÝVOJ ZMĚN V HODNOCENÍ ŽÁKŮ

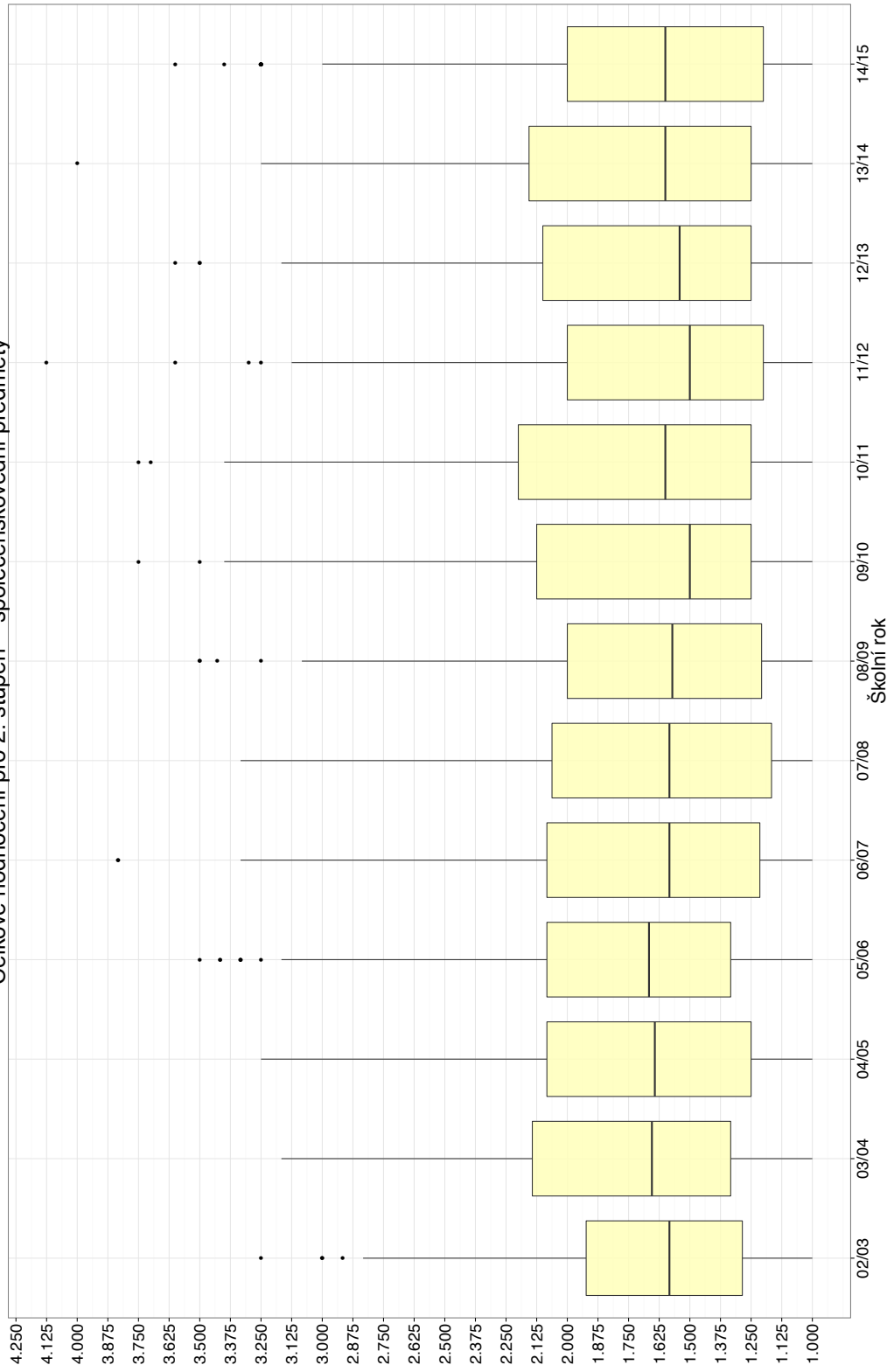


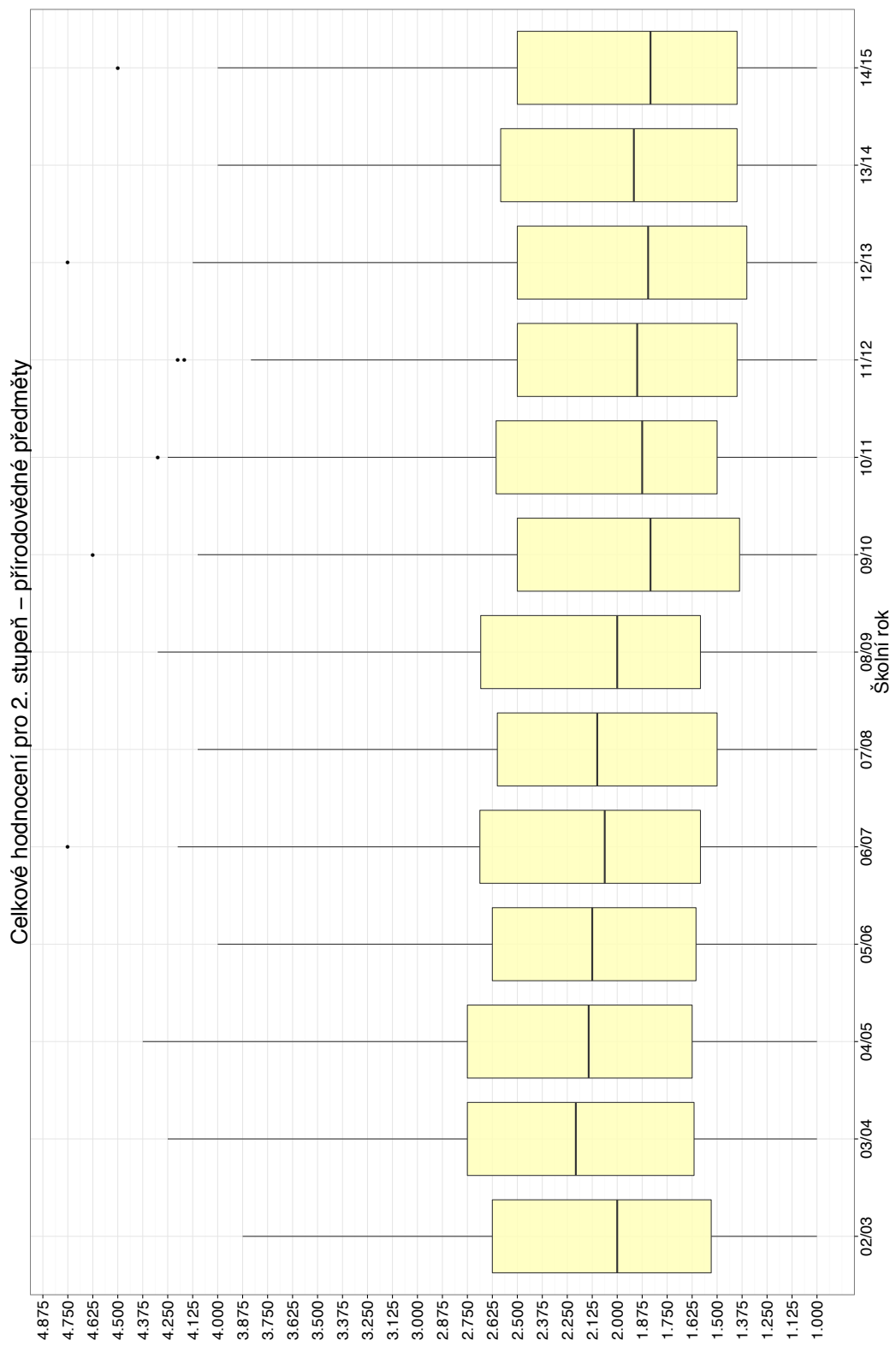
Celkové hodnocení pro 1. stupeň



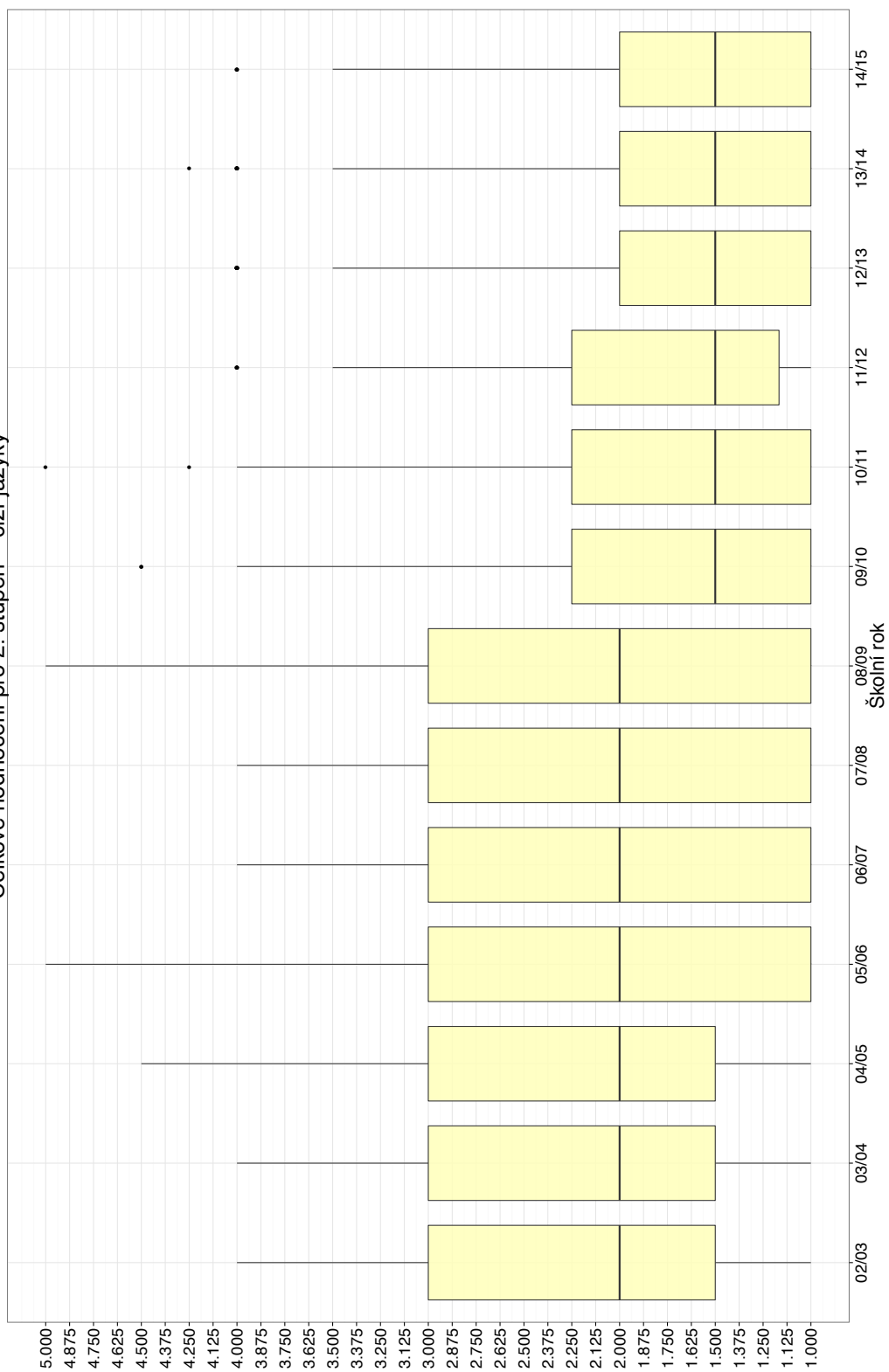


Celkové hodnocení pro 2. stupeň – společenskovední předměty

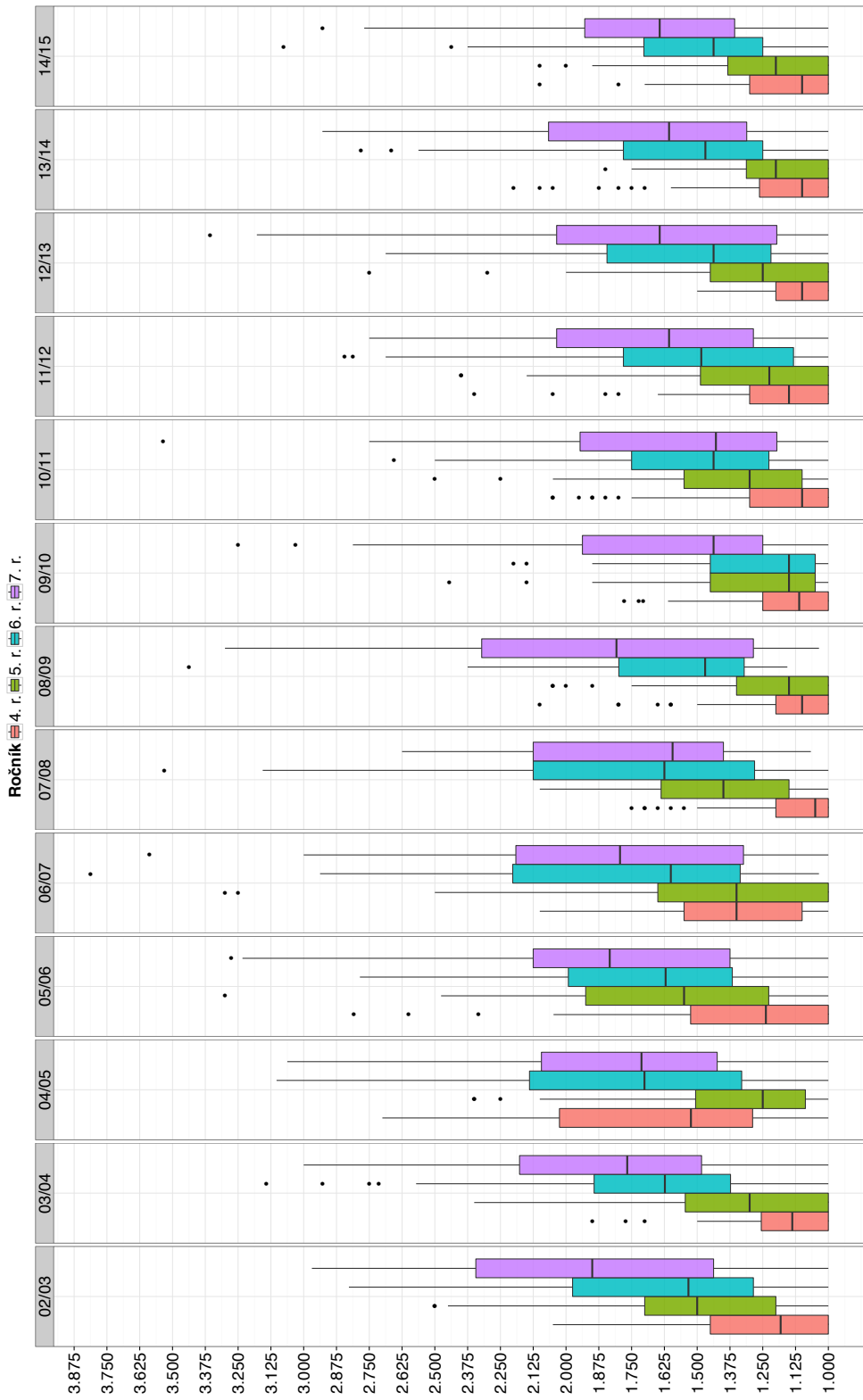




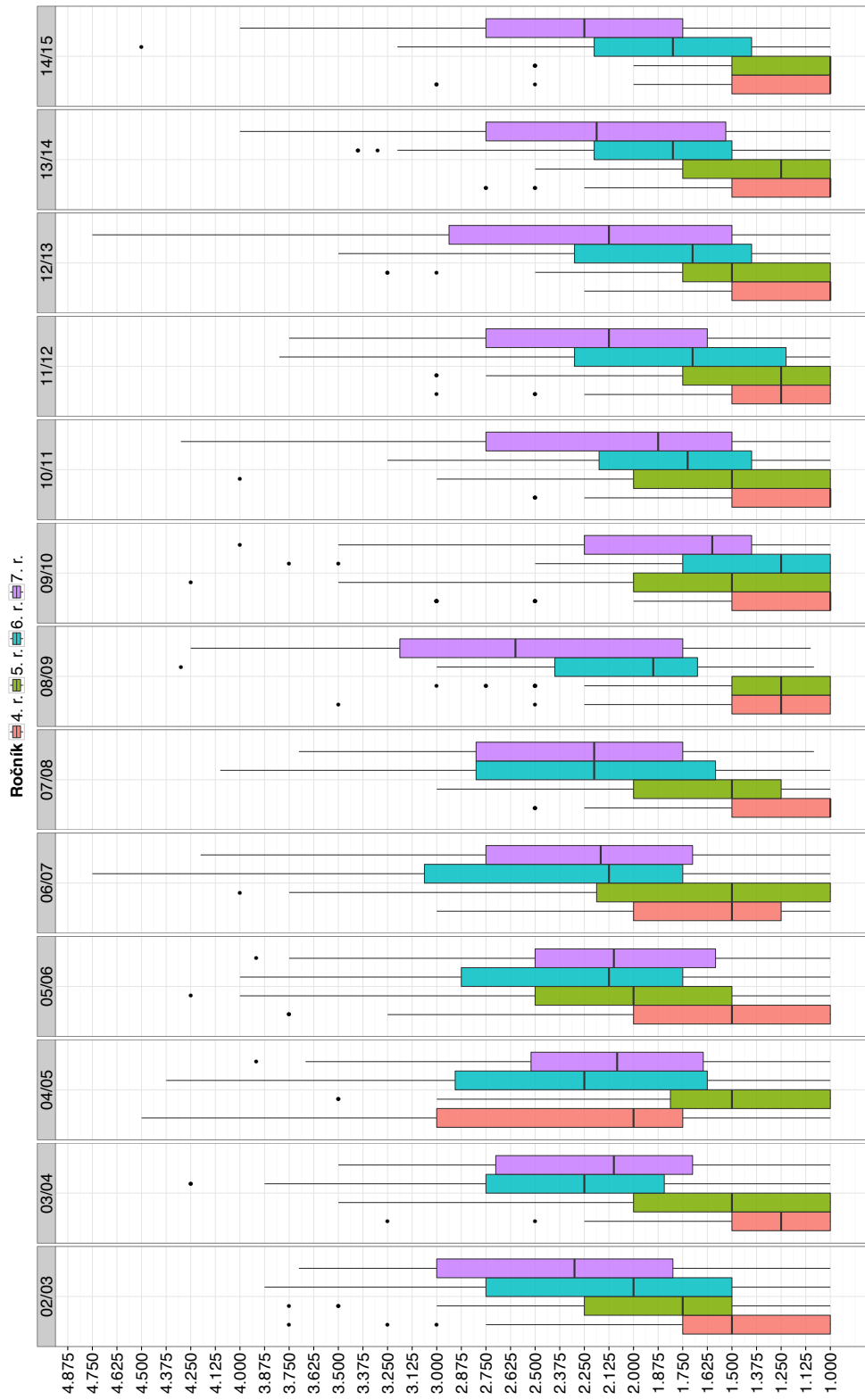
Celkové hodnocení pro 2. stupeň – cizí jazyky



### Adaptace 4./5./6./7. ročník – všechny předměty



# Adaptace 4./5./6./7. ročník – přírodovědné předměty





## E VLIV ASISTENTA PEDAGOGA NA TŘÍDU

Testovány byly jednotlivé pololetní známky žáků uvedených ročníků ve třídách s asistentem pedagoga (*A*) a ve třídách bez asistenta pedagoga (*bez*). Průměry známek jsou označeny  $\bar{x}_A$ , resp.  $\bar{x}_{bez}$ , jejich celkové počty pak  $n_A$ , resp.  $n_{bez}$ . Byl použit dvouvýběrový Welchův *t-test*. Tabulky zahrnují kritické hodnoty *t* a jim odpovídající *p*-hodnoty, porovnávané na různých hladinách významnosti  $\alpha$ , postupně 1 %, 5 %, 10 % a 15 %. Výběry dat a samotná testování byla realizována v prostředí jazyka R.

Hypotéza  $H_0$ : Přítomnost asistenta nemá žádný vliv na výkon třídy ( $\bar{x}_A \cong \bar{x}_{bez}$ ).

Hypotéza  $H_1$ : Přítomnost asistenta má pozitivní vliv na výkon třídy ( $\bar{x}_A < \bar{x}_{bez}$ ).

Tab. E.1: Vliv přítomnosti asistenta na celkový výkon třídy

r.	$\bar{x}_{bez}$	$n_{bez}$	$\bar{x}_A$	$n_A$	<i>t</i>	<i>p</i>	1 %	5 %	10 %	15 %
2.	1,024217	1404	1,025641	624	0,1888	0,5748	$H_0$	$H_0$	$H_0$	$H_0$
3.	1,092215	1323	1,088509	644	-0,2504	0,4011	$H_0$	$H_0$	$H_0$	$H_0$
4.	1,192	1875	1,182527	641	-0,4235	0,336	$H_0$	$H_0$	$H_0$	$H_0$
5.	1,265049	1509	1,198807	1030	-3,1718	7,67e-04	$H_1$	$H_1$	$H_1$	$H_1$
6.	1,525152	3121	1,493201	809	-1,1186	0,1318	$H_0$	$H_0$	$H_0$	$H_1$
9.	1,812639	3149	1,616809	702	-5,534	1,95e-08	$H_1$	$H_1$	$H_1$	$H_1$

Tab. E.2: Vliv přítomnosti asistenta na výkon třídy – cizí jazyky

r.	$\bar{x}_{bez}$	$n_{bez}$	$\bar{x}_A$	$n_A$	<i>t</i>	<i>p</i>	1 %	5 %	10 %	15 %
2.	1,00	156	1,00	70	–	–	–	–	–	–
3.	1,034014	147	1,098592	71	-1,4827	0,9292	$H_0$	$H_0$	$H_0$	$H_0$
4.	1,287234	188	1,265625	64	0,2706	0,3936	$H_0$	$H_0$	$H_0$	$H_0$
5.	1,524272	103	1,456954	151	0,893	0,1865	$H_0$	$H_0$	$H_0$	$H_0$
6.	1,703422	263	1,546875	64	1,6737	0,04847	$H_0$	$H_1$	$H_1$	$H_1$
9.	1,730594	438	1,483871	93	2,9832	1,64e-03	$H_1$	$H_1$	$H_1$	$H_1$

Tab. E.3: Vliv přítomnosti asistenta na výkon třídy – společenskovední předměty

r.	$\bar{x}_{bez}$	$n_{bez}$	$\bar{x}_A$	$n_A$	$t$	$p$	1 %	5 %	10 %	15 %
2.	1,041667	624	1,043478	276	-0,1235	0,5491	$H_0$	$H_0$	$H_0$	$H_0$
3.	1,127551	588	1,125874	286	0,0646	0,4743	$H_0$	$H_0$	$H_0$	$H_0$
4.	1,264	750	1,257812	256	0,1252	0,4394	$H_0$	$H_0$	$H_0$	$H_0$
5.	1,305825	412	1,226821	604	2,1801	0,01478	$H_0$	$H_1$	$H_1$	$H_1$
6.	1,501427	1051	1,533333	270	-0,643	0,7397	$H_0$	$H_0$	$H_0$	$H_0$
9.	1,982	1000	1,686364	220	4,6179	2,75e-06	$H_1$	$H_1$	$H_1$	$H_1$

Tab. E.4: Vliv přítomnosti asistenta na výkon třídy – přírodovědné předměty

r.	$\bar{x}_{bez}$	$n_{bez}$	$\bar{x}_A$	$n_A$	$t$	$p$	1 %	5 %	10 %	15 %
2.	1,051282	156	1,058824	68	-0,2233	0,5882	$H_0$	$H_0$	$H_0$	$H_0$
3.	1,285714	147	1,197183	71	1,3735	0,08584	$H_0$	$H_0$	$H_1$	$H_1$
4.	1,28877	374	1,265625	128	0,384	0,3507	$H_0$	$H_0$	$H_0$	$H_0$
5.	1,451456	206	1,311258	302	2,5427	0,0057	$H_1$	$H_1$	$H_1$	$H_1$
6.	1,847936	993	1,771318	258	1,3267	0,09269	$H_0$	$H_0$	$H_1$	$H_1$
9.	2,018257	1205	1,839552	268	2,8784	2,10e-03	$H_1$	$H_1$	$H_1$	$H_1$

## F OBSAH PŘILOŽENÉHO CD

CD, které je přiloženo, nese digitální kopii této práce ve formátu PDF a její odpovídající zdrojové kódy systému  $\text{\LaTeX}$ , v němž byla vysázena.

Původní, anonymizovaná data, jsou k dispozici v adresáři `orig-data`. Jeho součástí je mimo samotná data i anonymizační PHP skript `anonymize.php` a jím generovaný protokol `anonymize.log`. Adresář dále obsahuje textový soubor s neformálními poznámkami o vnitřním uspořádání a významu jednotlivých polí DBF souborů.

Konsolidované datové soubory formátu SQLite3 a výstupy analýzy obsahuje adresář `dat-analyza`. Přiložený konsolidační PHP skript `konsolidace.php` realizuje vytvoření chybějících tabulek s klasifikací v databázích starších verzí, jak je popsáno v kap. 4.2.2. Samotná analýza, provedená v prostředí R, je specifikována ve zdrojovém souboru `analyza.r`. Vektorové formáty výstupních grafů jsou přiloženy v podadresáři `grafy`.

Poslední součástí je adresář `web-app`, nesoucí všechny podklady, ze kterých vychází příloha B. Obsahuje v době práce aktuální seznam škol a školských zařízení v textovém souboru s hodnotami oddělenými čárkou. Jeho aktualizované podoby, zahrnující geolokační údaje, nalezené URL adresy webových aplikací Bakaláři a technické metriky serverového software, jsou umístěny v podadresáři výsledků `results`. Zdrojové kódy nástrojů pro obě automatizované metody, které příloha B popisuje, jsou k nalezení v podadresářích `google`, resp. `spider`. Podpůrné nástroje pro celkovou automatizaci, doplnění geolokačních informací, porovnávání textových řetězců a analýzu informací o školách a jimi používaných technických prostředcích, jsou shromážděny ve formě skriptů pro unixový shell Bash v podadresáři `bin`.



**UNIVERZITA HRADEC KRÁLOVÉ**

**Fakulta informatiky a managementu**

Rokitanského 62, 500 03 Hradec Králové, tel: 493 331 111, fax: 493 332 235

## **Zadání k závěrečné práci**

Jméno a příjmení studenta:

**Jan Hladěna**

Obor studia:

Aplikovaná informatika

Jméno a příjmení vedoucího práce:

**Hana Skalská**

Název práce:

**Analýza dat informačního systému**

Název práce v AJ:

Information system data analysis

Podtitul práce:

Podtitul práce v AJ:

Cíl práce: Návrh konsolidace a využití dat informačního systému základní školy. Převod dat do zdokumentované podoby. Shrnutí vývoje změn v hodnocení žáků za dobu používání IS.

Osnova práce:

1. Obecná charakteristika prostředí a analyzovaného IS
2. Rozbor interních datových struktur v IS a znázornění vztahů mezi nimi
3. Vytvoření relačního datového modelu, sjednocení dat
4. Příprava dat k analýze (transformace, čištění, ošetření chybějících dat)
5. Specifikace cíle analýzy, statistická analýza získaných dat, charakteristiky a korelace
6. Interpretace získaných výsledků v kontextu prostředí základní školy
7. Shrnutí, závěr, možnosti dalšího využití výsledků práce

Projednáno dne:

Podpis studenta

Podpis vedoucího práce