



TECHNICKÁ UNIVERZITA V LIBERCI
Fakulta mechatroniky, informatiky
a mezioborových studií ■

Pokročilý třídní jazykový slovníček

Diplomová práce

Studijní program: N2612 – Elektrotechnika a informatika

Studijní obor: 1802T007 – Informační technologie

Autor práce: **Daniel Maděra**

Vedoucí práce: Ing. Jana Vítvarová, Ph.D.





TECHNICAL UNIVERSITY OF LIBEREC
Faculty of Mechatronics, Informatics
and Interdisciplinary Studies ■

Advanced Class Language Dictionary

Master thesis

Study programme: N2612 – Engineering and Informatics

Study branch: 1802T007 – Information Technology

Author: **Daniel Maděra**

Supervisor: Ing. Jana Vitvarová, Ph.D.



ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Daniel Maděra**
Osobní číslo: **M14000170**
Studijní program: **N2612 Elektrotechnika a informatika**
Studijní obor: **Informační technologie**
Název tématu: **Pokročilý třídní jazykový slovníček**
Zadávající katedra: **Ústav mechatroniky a technické informatiky**

Z á s a d y p r o v y p r a c o v á n í :

1. Proveďte rešerši existujících řešení pro učení se a testování slovíček z cizích jazyků.
2. Navrhněte aplikaci jazykového slovníčku, která umožní studentům třídy motivující a personalizované procvičování slovíček z konkrétní učebnice.
3. Prozkoumejte a navrhněte způsoby, jak zjednodušit a automatizovat import požadovaných slovíček ve zvukové, obrazové a textové podobě.
4. Prozkoumejte metody vhodné na generování testovacích slovíček v závislosti na výsledcích předchozích testů a případně dalších parametrech.
5. Jazykový slovníček implementujte.
6. Slovníček naplňte slovíčky k vybrané cizojazyčné učebnici a otestujte.

Rozsah grafických prací: dle potřeby dokumentace

Rozsah pracovní zprávy: 40–50 stran

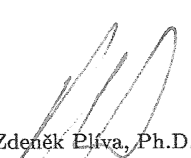
Forma zpracování diplomové práce: tištěná/elektronická

Seznam odborné literatury:


- [1] Russell S., Norvig P.: Artificial Intelligence: A Modern Approach, Pearson Education Limited, 2013, ISBN-13: 978-1292024202
- [2] Richardson L., Ruby S.: RESTful Web Services: Web services for the real world, O'Reilly Media, 2007, ISBN-13: 978-0596529260

Vedoucí diplomové práce: **Ing. Jana Vitvarová, Ph.D.**
Ústav mechatroniky a technické informatiky

Datum zadání diplomové práce: **10. října 2016**
Termín odevzdání diplomové práce: **15. května 2017**


prof. Ing. Zdeněk Pliva, Ph.D.
děkan




doc. Ing. Milan Kolář, CSc.
vedoucí ústavu

V Liberci dne 10. října 2016

Prohlášení

Byl jsem seznámen s tím, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé diplomové práce pro vnitřní potřebu TUL.

Užiji-li diplomovou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Diplomovou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím mé diplomové práce a konzultantem.

Současně čestně prohlašuji, že tištěná verze práce se shoduje s elektronickou verzí, vloženou do IS STAG.

Datum: 3. 1. 2017

Podpis:



Abstrakt

Klíčová slova

Abstract

Keywords

Obsah

1	Úvod	8
2	Analýza	9
2.1	Hlavní cíle	9
2.2	Existující řešení	10
2.3	Učení slovíček	12
2.4	Testování slovíček	14
2.5	Typy testů	14
2.6	Metoda rozloženého opakování	15
2.7	Specifikace požadavků	16
3	Návrh aplikace	18
3.1	Uživatelské role	18
3.2	Správa učebnic	19
3.3	Správa slovíček	19
3.4	Obtížnost slovíček	20
3.5	Generování slov pro procvičování	22
3.6	Procvičování slovíček	24
3.7	Připomínání slov	25
3.8	Motivace	27
3.9	Architektura aplikace	28
4	Klientská část	30
4.1	Návrh aplikace	30
4.2	Vývojové prostředí	33
4.3	Knihovna React	34
4.4	Testování	36
5	Serverová část	39
5.1	Technologie	39
5.2	API	40
5.3	Zabezpečení	42
5.4	Databáze	44
5.5	Testování	45
6	Závěr	46
	Použitá literatura	47

Seznam obrázků

1	Naznačení rozloženého opakování na základě křivky zapomínání	16
2	USE-CASE diagram aplikace	18
3	Ukázka struktury URL	31
4	Adresářové struktury klientské aplikace	37
5	Ukázka části databázového modelu zodpovědné za správu slovíček	45

Seznam tabulek

1	Používání slov v britském anglickém jazyce	12
2	Matice koeficientů pro výpočet optimálních intervalů	27

Seznam ukázek zdrojových kódů

1	Příklady použití Levenshteinovy vzdálenosti	32
2	Příklad komponenty pro zobrazování nápovědy při procvičování slovíček . .	35
3	Hlavní renderovací funkce aplikace	36
4	Serializační třída k transformaci dat o učebnici	41
5	Uložení zvukové nahrávky	42
6	Obsah metody <i>jwt_response_payload_handler</i>	43
7	Ukázka třídy s přizpůsobeným oprávněním	44

Seznam zkratek

AJAX Asynchronous JavaScript and XML.

ODBC Open Database Connectivity.

XML Extensible Markup Language.

1 Úvod

2 Analýza

Výuka cizích jazyků je pro aktuální společnost jedna z nejzásadnějších otázek, ať se jedná o pracovní příležitosti v zahraničí, tak sociální problematika světa. Žáci a studenti se často účastní různých kroužků nebo později studenti využívají Erasmus programů, kde vyhledávají právě zlepšení komunikace v cizím jazyce.

2.1 Hlavní cíle

Hlavním cílem aplikace je připravit žáky na hodinu cizího jazyka a zároveň přirozeně rozvíjet slovní zásobu, tak aby studenti neztratili motivaci a chuť k poznávání nových výrazů. Dále také umožnit žákům se protestovat a ověřit, zda naučenou sadu slov ovládají. Aplikace by se měla adaptovat na zdatnost a úroveň každého ze žáků. Vyučujícím by aplikace měla usnadnit správu a import slovíček, které třída má umět v rámci dané učebnice a následně předkládat žákům vhodnou slovní zásobu například pro následující lekci.

2.1.1 Personalizace

Důležitým cílem aplikace by měla být personalizace podle potřeby jednotlivých žáků a celých tříd. Personalizace na úrovni žáka znamená přistupovat individuálně na základě jeho úrovně znalostí a dovedností. V případě konkrétní třídy jde o individuální přístup a to zejména v sadě slov, které je třeba v přípravě procvičovat.

2.1.2 Motivace

Důležitou součástí vzdělávání obecně je motivace. Tedy přimět žáky, aby z vlastní iniciativy chtěli rozvíjet své vědomosti. Problém ale je, že se děti přirozeně neučí z vlastní iniciativy, ale aby uspokojili okolí. Motivace se rozděluje do skupin - vnitřní a vnější nebo pozitivní a negativní [3]. V analýze bude zaměřeno pouze na motivační prostředky, které lze zařadit do aplikace pro procvičování slovíček.

Motivace založená na základě vlastních úspěchů je důležitá pro utvrzení sebevědomí žáků. Ocenění v případě zvládnutí sady slov nebo gramatického bloku lze v případě aplikace implementovat například hláškami s projevem pochvaly nebo jiným upozorněním na dosažený výsledek. Zajímavým prvkem v aplikaci by mohlo být i herní prostředí. Žáci si rádi hrají a již J. A. Komenský poukázal na důležitost her ve vzdělávacím procesu.

Jedním z dalších stěžejních faktorů motivace je kolektiv. Právě díky kolektivu, ve kterém funguje přirozená rivalita, jsou schopni žáci dosáhnout mnohem vyšších výsledků než kdyby se vzdělávali odděleně a samostatně. Rivalita a soutěživost může projevovat i velmi

negativním způsobem. Místo kamarádských vztahů mezi dětmi mohou vznikat nepřátelské, kde může docházet i k posměchu těch, kteří nemusí mít pro výuku až takové nadání. Zajímavější motivací tedy pro kolektiv je například vidina společně dosažených výsledků. V případě učení slovíček by to byl počet slovíček naučených za rok jako celá třída. Dochází zde k utužování kolektivu a děti by mohlo těšit to, že nějakým způsobem přispívají k úspěchům celé třídy.

2.1.3 Využití IT pro výuku

Posledních několik let se společnost ubírá trendem informačních technologií. Každý ze žáků má už od útlého věku přístup k počítači nebo k chytrému telefonu. Orientace a schopnost tyto zařízení používat není pro ně žádný problém. Přirozeně se tedy tyto zařízení postupně stávají součástí každodenní přípravy žáka na následující školní den. V některých případech tyto zařízení plně nahrazují klasické učebnice a jsou přímo začleněny do výuky. Použití informačních technologií má za následek i zlepšení motivace při výuce. Obecně je známo, že žáci raději studují slovíčka interaktivní formou hádanek, křížovek nebo her, než nekonečných seznamů slov.

Důležitost cizích jazyků se projevuje na míře používání například chytrých tabulí nebo tabletů při výuce. Tyto zařízení umožňují interaktivní výuku, kde lze využít nejen textových, ale také obrázkových a zvukových prostředků pro lepší zasazení nově nabytých vědomostí do kontextu.

2.2 Existující řešení

Na trhu lze nalézt nepřeberné množství aplikací pro výuku cizích jazyků. Řada z nich jsou téměř komplexní systémy, které provází studenta od základních frází a slovíček až po gramatické standardy cizího jazyka. Analýza existujících řešení byla zaměřena na aplikace, které se zabývají především testováním slovíček a frází.

Do analýzy existujících řešení byly zahrnuty tři desktopové, jedna webová a jedna mobilní aplikace.

2.2.1 TS Angličtina

Z analyzovaných řešení se jevila desktopová aplikace TS Angličtina od firmy Terasoft ta nejlépe funkčně propracovaná. Tato firma se zabývá širokou škálou výukových nástrojů se zaměřením na základní školy. V případě cizích jazyků se zabývají výukou anglického a německého jazyka. Hlavní předností aplikace je podpora nejvíce používaných učebnic cizího jazyka. Aplikace umožňuje testování různými způsoby - psaný překlad slova, porozumění mluvenému slovu, vybírání správných významů nebo doplňování vynechaných slov [1].

Analýza byla tvořena pouze z informací vydavatele. Bohužel firma Terasoft neposkytuje DEMO nebo TRIAL verzi aplikace, která by hlubší analýzu umožňovala.

2.2.2 Langsoft Teacher

Dalším analyzovaným řešením byla aplikace Langsoft Teacher, která je dostupná v podání desktopové aplikace, ale také i mobilní aplikace pro platformy iOS a Android. Aplikace je velmi komplexní, obsahuje různé moduly pro testování například v obrazové formě pro předškolní děti. Zajímavá vlastnost, kterou aplikace disponuje, je pamatování problematických slov a nabízení těchto slov častěji než těch bezproblémových. Dále program umožňuje automaticky rozšiřovat slovní zásobu, která je v testování zahrnuta [2].

2.2.3 Duolingo

Dalším v pořadí byla aplikace Duolingo. Jedná se o mobilní aplikaci pro Android. Zahrnuje učivo cizího jazyka od základních komunikačních frází až po tvorbu gramaticky složitějších vět. V aplikaci je připravena dlouhá řada cizích jazyků - němčina, angličtina, španělština (kastilština), italština a další. Chybí ale více referenčních jazyků. Aktuálně lze využít pouze angličtinu. Aplikace kromě standardních funkcí zahrnuje i rozpoznávání mluvených odpovědí. Zajímavým poznatkem byl systém motivace uživatelů, kde si každý mohl pozvat své přátele, mezi kterými docházelo k sdílení dosažených výsledků. Dalším motivačním základem bylo nutkání udržení plánu pravidelného testování, jelikož v opačném případě docházelo k automatickému zvyšování objemu testovacích dat. Nevýhodou aplikace byla nutnost připojení k internetu. V případě požití mobilních dat, docházelo ke zpoždění zejména při rozpoznání slov.

2.2.4 Vocabulary Trainer

Vocabulary Trainer je webová aplikace zdarma napsaná v jazyce PHP, která naučí 5000 nejvíce používaných slov daného jazyka. Aplikace umožňuje hodně možného nastavení. K dispozici je i řada jazyků včetně češtiny a to jako referenční i jako učený jazyk. Testování spočívá nejdříve ve čtení slov a následně uživatel vybírá možnosti odpovědi. Aktuálně překládané slovo si lze kdykoliv přehrát v různé rychlosti. Jako motivační základ aplikace využívá jednoduchý bodový systém. Součástí je i kalendář s email upomínkou k dalšímu testování. Program je propracovaný, ale poměrně pomalý a dlouho trvá zejména úvodní načítání dat. Její výrobce LanguageCourse S.L. poskytuje i mobilní aplikaci pro Android k učení anglických slov a frází.

2.2.5 Supermemo aplikace

Za zmínku ještě stojí aplikace Supermemo. Není to aplikace s připravenými daty k testování slov cizího jazyka, ale slouží čistě jako šablona pro testování jakéhokoliv druhu otázek. Aplikace implementuje algoritmus Supermemo, který je založen na metodě postupného zvyšování intervalu dotazování na dané otázky. Při každé odpovědi program spočítá, kdy by si uživatel měl danou otázku zopakovat tak, aby odpověď byla správná a zároveň se co nejvíce zvyšoval interval mezi aktuální a předchozí odpovědí. Aplikace se adaptuje na schopnosti uživatele, v přiměřeném měřítku buď zvyšuje nebo snižuje interval dalšího připomenutí.

Z výše uvedených aplikací až na Supermemo žádná neumožňuje personalizovaný výběr učiva. Tedy nelze vložit vlastní slovíčko nebo si určit sadu slov pro testování. Proto jsou tyto aplikace především cílené pro uživatele, kteří vnímají výuku jazyka jako samostatné vzdělávání sami sebe. Pro studenty, kteří absolvují lekce z cizího jazyka ve škole je tento typ vzdělávání nevyhovující, jelikož se musí učit dvě nezávislé skupiny slov. Sice dochází k rozvinutí slovní zásoby studenta i do jiných okruhů než je jeho učebnice a málokterý student má ještě energii, časovou dotaci a vlastní iniciativu na to, aby se připravoval na školní test ze slovíček a ještě rozvíjel samostatně svoji cizojazyčnou slovní zásobu.

2.3 Učení slovíček

Standardní učení slovní zásoby cizího jazyka je založeno na častém opakování slov. Dle výzkumu Biemillera a Boote se lze ročně zvládnout až 400 slov u studentů 3.—6. tříd [4]. Což v případě cizího jazyka poměrně velké číslo, ale problém je, do jaké míry je slovo správně ukotveno v dlouhodobé paměti. Porovnání, kolik je potřeba slov pro ovládnutí anglického jazyka, usnadní následující tabulka 1, která zahrnuje procentuální využití nejvíce používaných slov v každém z odvětví [6].

	Konverzace	Noviny	Akademický text
1. 1000 slov	84,3%	75,6%	73,5%
2. 1000 slov	6%	4,7%	4,6%
Akademické výrazy	1,9%	3,9%	8,5%
Ostatní	7,8%	15,7%	13,3%

Tabulka 1: Používání slov v britském anglickém jazyce

2.3.1 Zastaralý způsob učení

Dle vlastního průzkumu žáci základních škol nevyužívají k učení slovíček nikterak pokročilé technologie. Většinou si udržují vlastní slovníček, do kterého nově nabytá slova. Při učení zakrývají část s cizími slovy a na základě českého ekvivalentu se snaží vybavit

překlad slova. Tato metoda neposkytuje prakticky žádnou zpětnou vazbu. Žáci většinou pouze do krátkodobé paměti uloží slova, později při testu rychle zodpoví a následně během pár hodin si na slovo už ani nevzpomenou. Nevýhod a námětů na zlepšení metody má tato metoda celou řadu, ale jedním z klíčových vad je, že žáci si nevytvoří dostatečně souvislostí, aby řádně ukotvilo slovo v paměti.

2.3.2 Zvuková interpretace

Zejména v učení cizího jazyka je zvuková podoba a interpretace slov velice důležitá. Jelikož každý jazyk může hlásky různě zvukově interpretovat a pro nováčka v cizím jazyku nemusí textová výslovnost plně vyhovovat. Žák dále díky znění slova získá podvědomí o dialektu daného jazyka a zároveň dochází k lepšímu zapamatování slova. Právě díky zvukům dochází k propojení při učení i pravé mozkové hemisféry. A napomáhá tedy k vytvoření pevnější ukotvení v paměti. Dle studie bulharského vědce George Lozanova, který se zabýval studiem mozku a učebních metod, byl zjištěn obrovský přínos zvukových vjemů [5].

2.3.3 Problematika obtížnosti

Velkým problémem v učení slov je přizpůsobit obtížnost každému ze žáků individuálně. Jelikož ne všichni mají stejnou úroveň znalostí cizího jazyka a každý potřebuje jiné tempo pro zapamatování sady slov. Jsou žáci s výbornou pamětí, kterým stačí si slova pouze jednou projít a dokáží je používat, ale jsou žáci, kde nestačí je pětikrát zopakovat. Dalším problémem obtížnosti je z pohledu jednotlivých slov. Každé slovo má odlišnou obtížnost, které lze soudit například podle míry používanosti v jazyce, délky slova, zdali obsahuje přehlasování, dvojité písmena nebo podobnost s referenčním slovem.

2.3.4 Učení slov v kontextu

Pro učení slov cizího jazyka je rovněž důležité správné zasazení významu slova do kontextu. Dle průzkumu Biemillera a Bootea z roku 2006 bylo zjištěno, že u žáků od 10—13 let docházelo k nárůstu zapamatovaných slov o 4%, pokud byla slova předložena v kontextu vět. Důležitým poznatkem z toho průzkumu je, že žáci si nejen déle naučené slovo pamatovali, ale správně ho i interpretovali, když měli za úkol ho svými slovy vysvětlit [4]. Ve stejném průzkumu rovněž docházelo ještě k vyššímu zlepšení v učení, kdy si žáci zapisovali vlastními slovy definici a použití slova.

2.4 Testování slovíček

2.4.1 Aktivní a pasivní slovní zásoba

Slovní zásobu, kterou využíváme k tvorbě vět ať už v cizím nebo mateřském jazyce rozdělujeme na dvě skupiny - aktivní a pasivní. Pasivní zásoba je sada slov, která jsou pevně a spolehlivě uložena v naší paměti. Průměrný žák zná cca 50 000 výrazů. Její velikost je ovlivněna věkem, vzděláním a četbou. Slova ze této sady využíváme zejména při písemné formě, ať už se jedná o čtení nebo psaní. Aktivní zásoba je sada slov, ve které lze najít žádané slovo během desítek milisekund. U většiny lidí dosahuje velikosti 4 000 až 8 000 výrazů [7]. Slovo se díky používání dostává z pasivní do aktivní slovní zásoby.

2.4.2 Metody testování

Metod testování slovní zásoby je mnoho. Základním rozdělením je na pasivní a aktivní. V případě pasivního se jedná například o výběr z nabídnutých možností. Jde o případ, kdy student nemusí znát přesnou odpověď a dokáže otázku vyhodnotit správně vylučovací metodou. Při aktivním testování žáci musí odpověď znát, aby otázka byla vyhodnocena správně. Pasivní metoda má pozitivní vliv například na motivaci žáka, který u ní tolik netápe a za pomoci zdravého rozumu může test vyhodnotit správně. Problém ale vzniká při používání získaných a procvičených informací. V případě cizího jazyka lze pasivní slovní zásobu využít pro čtení a náslech, ale pro ovládnutí cizího jazyka je nedostatečná a tudíž nevhodná.

Dalším rozdělením pasivního a aktivního testování slovíček je rozpoznávání (*recognition*) a vzpomnutí (*recall*). Rozpoznávání spočívá v předložení cizího slova žákovi a pro správné zodpovězení musí najít český ekvivalent. Vzpomnutí je způsob testování opačným způsobem než rozpoznání. Tedy uživateli je předložen výraz v jeho mateřském jazyce a on musí nalézt správný výraz v cizím. Rozpoznání je z principu jednodušší pro uživatele než vzpomnutí. Lze ho tedy využít rovněž pro zvýšení motivace při procvičování slov. Ale pro ověření, zda je slovo ovládnuté či nikoliv je metoda rozpoznání také nedostatečná.

2.5 Typy testů

V analýze existujících řešení bylo procvičování a testování slov interpretováno různými způsoby. Obecně by se typ testů dal rozdělit na tři kategorie - textové testy, multimediální a herní testy.

2.5.1 Textové testy

Textové testy se vyskytovaly například jako výběr z více možností nebo spojováním spolu souvisejících významů, jak už ale bylo uvedeno v kapitole 2.4.2, jedná se o rozvíjení pasivní slovní zásoby. Zajímavějším typem testů je doplňování slov do vět a klasický překlad slova. Tyto typy rozvíjejí žádanou aktivní slovní zásobu. Standardně textové testy byly doplněny zvukovou interpretací cizího slova.

2.5.2 Multimediální a herní testy

Zajímavým využitím multimédií by mohlo být zahrnutí otestování výslovnosti. Tedy možnost nahrání odpovědi a následně by došlo k rozpoznání. Ale analýza a rozpoznání řeči není nikterak triviální záležitost. Dalším zajímavým řešením byly herní testy. Díky nimž docházelo k zvýšení motivace uživatelů. Problém je, že tento typ procvičování většinou není zas tolik efektivní, jelikož si uživatel procvičí například v případě doplňování písmen křížovky nebo známé hry *Hangman* pouze pár slov za poměrně dlouhý čas.

2.6 Metoda rozloženého opakování

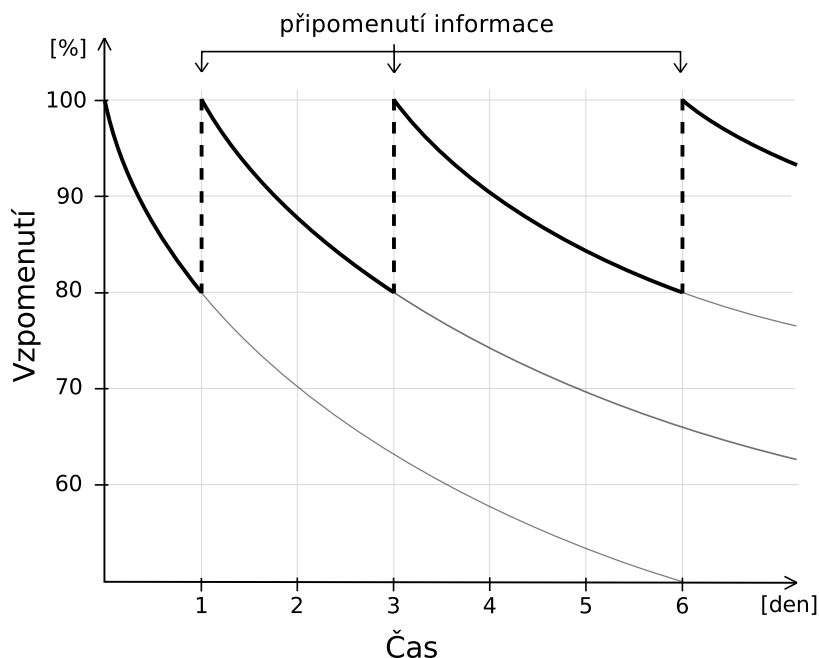
Až na výjimečné paměti je obecně známo, že pokud chceme informaci v paměti uchovat dlouhodobě, je nutné opakovaně připomínat. Pokud informace není vyžívána s největší pravděpodobností dojde ke ztrátě této informace. V případě, že uživatel s touto informací pracuje častěji, výrazně navyšuje šance pro její zapamatování. Metoda rozloženého opakování (*spaced repetition*) je učební technika založená na opakovaném připomínání informace s navyšujícím intervalem. Na začátku učebního procesu, jsou intervaly krátké například na hodinu, 4 hodiny nebo celý den. S postupem se tyto intervaly mohou zvyšují až na týdny a měsíce. Ideální systém rozloženého opakování nabídne zopakování informace před tím, než dojde k jejímu eventuálnímu nevzpomenutí.

2.6.1 Křivka zapomínání

Křivka zapomínání je graf exponenciálního charakteru, který znázorňuje za jak dlouho a s jakou pravděpodobností je možné si na nově nabytou informaci vzpomenout. První tuto křivku specifikoval už v roce 1885 Hermann Ebbinghaus. Na ose Y lze nalézt procentuální pravděpodobnost vzpomenutí na informaci, na ose X je čas. V čase 0 dojde k nabytí informace a s postupem času dochází k zapomínání. U průměrných pamětí dojde k potenciálnímu zapomenutí nové nesouvisející informace bez kontextu během pěti dní [10].

Následující obrázek 1 znázorňuje metodu rozloženého opakování na křivce zapomínání. V čase s 80% pravděpodobností vzpomenutí na danou informaci dojde k znovu připome-

nutí informace. Křivka zapomínání se posune na ose Y a znovu dochází k postupnému zapomínání. Díky exponenciálnímu charakteru interval mezi jednotlivými opakováními narůstá také exponenciálně. V ideálním případě se posune exponenciála tak, že se bude limitně blížit 80% hranici.



Obrázek 1: Naznačení rozloženého opakování na základě křivky zapomínání

2.6.2 Leitnerův algoritmus

Na Leitnerův algoritmus lze nahlížet jako na známý učební postup v podobě karet s otázkami na informace. Tyto karty jsou rozděleny do skupin, které jsou označeny například 1 až 3, kategorií ale obecně může být více. Nová karta je zařazena do první kategorie. Karty v první kategorii jsou opakovány častěji, v poslední kategorii jsou karty opakovány méně často. Karty jsou přemisťovány podle toho, jak uživatel zvládá odpovídat na jednotlivé otázky. V případě, že uživateli otázka na kartě nedělá problémy, přesune kartu do vyšší kategorie. V opačném případě, kdy uživatel nemůže opovědět správně na danou otázku, karta se přesune do první kategorie a cyklus pro danou kartu začíná znovu.

2.7 Specifikace požadavků

Na základě analýzy byla provedena specifikace požadavků, které budou implementovány v aplikaci pro testování slovíček z cizího jazyka.

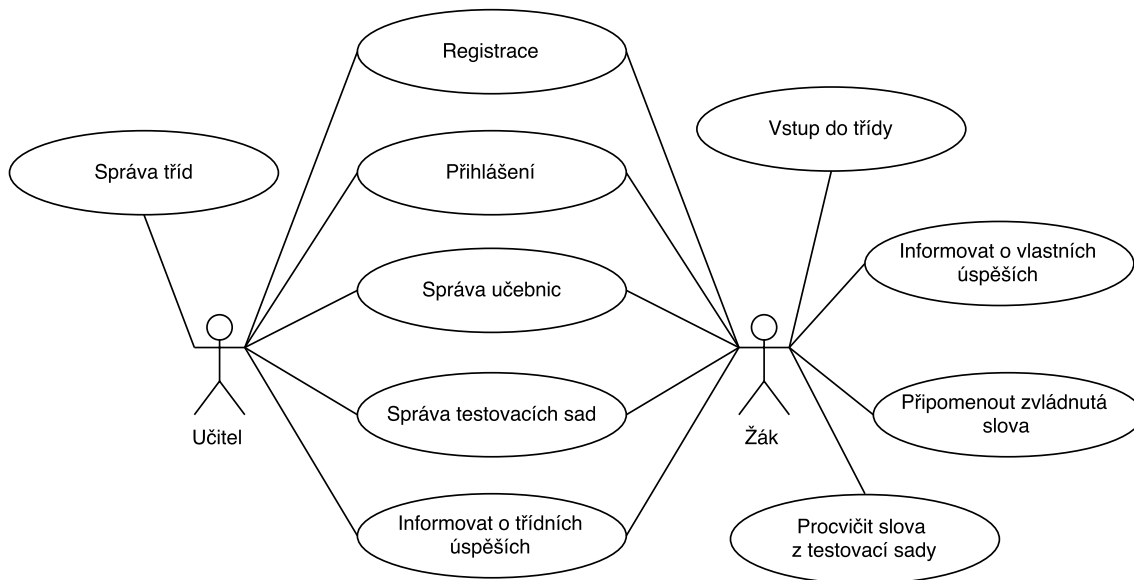
1. správa učebnic

- tvořit, editovat a mazat vlastní učebnice

- umožnit publikovat učebnici pro veřejnost
2. správa slovíček v učebnici
 - tvořit, editovat a mazat slovíčka
 - hromadně slovíčka do učebnice importovat
 - částečně automatizovat zvukovou a obrazovou interpretaci slovíčka
 3. správa modulů a tématických okruhů v učebnici
 - tvořit, editovat a mazat moduly a tématické okruhy učebnice
 - přiřazovat slovíčka do daných modulů a okruhů
 4. správa uživatelských skupin (tříd)
 - tvořit, editovat a mazat uživatelské skupiny
 - umožnit uživatelům se přihlásit do skupiny
 5. správa testovacích sad
 - tvořit, editovat a mazat testovací sady
 - umožnit vybírat slova pro testovací sadu z vlastních i veřejných učebnic
 - přiřazovat testovací sady ke skupinám uživatelů
 6. procvičování slovíček
 - vybrat testovací sadu slovíček
 - generovat slova na základě úrovně uživatele
 - zahrnout obrazovou a zvukovou interpretaci do procvičování
 - možnost uložit stav testování a umožnit pozdější navázání
 7. připomínat a procvičovat ovládnutou slovní zásobu
 8. motivace
 - motivovat v rámci uživatelské skupiny
 - motivovat vlastní iniciativu k procvičování

3 Návrh aplikace

Na základě specifikace požadavků byl vytvořen USE-CASE diagram aplikace. Diagram na obrázku 2 znázorňuje pouze obecné bloky funkčnosti aplikace. V následujících kapitolách bude každý z jednotlivých USE-CASE bloků popsán detailněji.



Obrázek 2: USE-CASE diagram aplikace

3.1 Uživatelské role

Do aplikace by měly vstupovat dva typy uživatelů - učitel a žák. Učitel z principu má za úkol spravovat testovací sady, učebnice a třídy. Žáci oproti nim mají za úkol vstupovat do svých tříd, procvičovat slova z testovacích dat, které jsou připravené od učitele a připomínat si zvládnutá slova.

Po konzultacích s vedoucím práce a vyučujícím cizího jazyka na základní škole došlo k několika změnám v návrhu právě v oblasti uživatelských rolí a i v principu aplikace. Vznikl požadavek, že aplikace by mohla fungovat spíše jako portál pro vzdělávání a ne jako výukový nástroj pro učitele. V návrhu tudíž vzniklo rovnocenné partnerství mezi učitelem a žákem.

3.1.1 Sjedení uživatelským skupin

Podnětem pro sjednocení uživatelských skupin vedlo umožnění žákům procvičovat si slova i v případě, že nejsou součástí žádné třídy. Tak aby mohl kdokoli se do aplikace přihlásit, vybrat si testovací sadu a vyzkoušet svoje znalosti slovíček. Tato funkce umožňuje žákům volnější přístup k výuce. Nebylo totiž v požadavcích to, aby učitel měl přehled o znalostech žáka - právě naopak. Účelem aplikace je poskytnout žákům možnost se samostatně

vzdělávat a procvičovat slovní zásobu a ne vyučujícím poskytnout nástroj pro otestování, zda žák disponuje požadovanými znalostmi.

3.2 Správa učebnic

Z důvodu personalizace slovíček bylo nutné v návrhu aplikace zařadit učebnice, díky nimž si žáci budou procvičovat pouze slovíčka, která jsou pro ně aktuální ve výuce cizího jazyka. Učebnice obsahují moduly a v modulech se nacházejí jednotlivá slova. Při tvorbě učebnice bude zvolen cizí jazyk, pro který je učebnice připravená. V rámci jedné učebnice jsou slova unikátní. Tzn. do konkrétní učebnice nemohou patřit dvě stejná slova. Majitel učebnice je může editovat, vytvářet, skrýt pro veřejnost a mazat i v případě, že je využívána některým z ostatních uživatelů.

V aplikaci se může vyskytovat i více stejných učebnic. Například dva různí vyučující vedou své hodiny cizího jazyka dle stejné učebnice, ale každý si chce přizpůsobit sadu slov po svém. Proto se v aplikaci musí učebnice identifikovat majitelem a názvem učebnice.

3.2.1 Sdílení učebnic

Jelikož aplikace slouží jako portál pro vzdělávání, ve výchozím stavu jsou učebnice a slovíčka v nich veřejně přístupné jakémukoliv uživateli. Příchozí uživatel si tedy bude moci buď vytvořit vlastní učebnici se svými slovy, které si chce procvičit nebo může vyhledat už z vytvořených učebnic. Díky sdílení nový uživatel může prakticky okamžitě po přihlášení začít procvičovat slova a nemusí složité žádná slova importovat. Tento přístup je cílený právě především pro žáky, kteří si chtějí samostatně procvičovat slova a nepatřit do žádné konkrétní třídy.

3.3 Správa slovíček

Jednotlivá slovíčka se budou ukládat do daným modulů učebnice. Slovíčko se bude charakterizovat - překladem v cizím jazyce, významem v mateřském jazyce, definice slova v cizím i mateřském jazyce a použití ve větách z učebnice. Dalšími atributy, kterými slovíčko bude disponovat, jsou obtížnost a slovní druh. Slovíčko bude možné doplnit o multimediální interpretaci a to v podobě zvuku a obrázku.

3.3.1 Textová forma

Textová forma slovíčka spočívá v překladu slova a jeho významu v mateřském jazyce. Pro zjištění významu slova se může využít již hotové aplikace. Například Google Translate poskytuje aplikační rozhraní pro překlady slov i celých vět. Výhodou by bylo usnadnění

zadávaní a import slovíček do aplikace. Jelikož se aplikace bude zabývat aktivním rozpoznáním a vzpomínáním, je vhodné neautomatizovat významy slov. Učebnice nebo učitelé se mohou s Google Translate lišit v definování významů slov. Dalším problémem je cena využívání překladů z Google Translate, která je účtovaná měsíčně a to \$20 za 1 milion znaků [9].

3.3.2 Zvuková forma

Do aplikace by měla být implementována i zvuková interpretace slovíčka v cizím jazyce. Zvukové nahrávky mohou být zařazeny do aplikace dvěma způsoby - manuálně vlastními nahrávkami nebo využít Google Text to Speech API, které poskytuje zdarma. Aplikace bude záznam lokálně ukládat, aby bylo se co nejvíce omezilo využívání API. Zvuková nahrávka bude získána při importu slovíček do učebnice. Zadávající uživatel si bude moci zvolit, k jakému slovu chce zvukovou interpretaci.

3.3.3 Obrazová forma

Obrazová interpretace slovíčka bude rovněž získána z Google Custom Search API. Toto aplikační rozhraní není omezené a lze v parametrech omezit vyhledávání pouze na obrázky. Při importu si uživatel zvolí pro jaká slovíčka chce obrazovou formu. Automatizace v tomto případě není vhodná, jelikož ne všechny slovíčka lze interpretovat jako obrázek. Uživatel si také bude moci zvolit obrázek z několika možností. Po vybrání obrázku bude vytvořena komprimovaná kopie souboru a uloží se rovněž zdroj, odkud je obrázek získán. Tato forma bude v aplikaci sloužit pouze jako nápověda pro žáky, díky níž si mohou ke slovíčku vytvořit hlubší asociaci.

Díky Google Custom Search API lze vyhledávání parametrizovat. Výrazem pro vyhledávání bude referenční slovo v cizím jazyce, jelikož cílový jazyk bude nejspíš anglický, německý nebo španělský a všechny tyto jazyky jsou více rozšířené než ten český. Dalším parametrem vyhledávání je omezení pouze na obrázky s licencí a „Volně užívat nebo sdílet“, která umožňuje obrázek zkopírovat a nekomerčně publikovat s uvedením zdroje. Bohužel společnost Google Inc. nezaručuje pod jakou licencí obrázky aktuálně se na stránkách prezentuje. Proto bude uživatel při výběru obrázku vyzván, aby zkontroloval licenci použití.

3.4 Obtížnost slovíček

Při zadávání slova do učebnice bude automaticky vyhodnocena jeho obtížnost, která je rozdělena do čtyř kategorií - snadné, střední, těžké a nemožné. Vyhodnocení obtížnosti je

založeno na kombinaci několika parametrů slova - jeho délky, podobnosti s českým ekvivalentem a další atributy jako počet přehlasovaných písmen nebo počet výskytů dvojitéch písmen. V případě nevhodně vygenerované obtížnosti slova, uživatel ji bude moci opravit dle vlastního uvážení. Délka slova je prahově rozdělena do kategorií obtížnosti, následně bude vyhodnocena podobnost s významem pomocí Levensteinovy vzdálenosti popsané v podkapitole 3.4.2. Poté následuje vyhodnocení přehlasování a dvojitéch hlásek, které ale ovlivňují obtížnost už minimálně.

3.4.1 Výběr algoritmu

Levenshteinova vzdálenost je velmi podobná známé Hammingově vzdálenosti s tím rozdílem, že Hammingova vzdálenost uvádí počet pozic se stejným symbolem v obou řetězcích. Kdežto Levenshteinova uvádí počet jednoznakových editací. Pro určení míry podobnosti mezi slovem v cizím jazyce a jeho překladem je vhodnější Levenshteinova vzdálenost, jelikož akceptuje řetězce s různou délkou a v případě, že písmeno chybí, je to vyhodnoceno pouze jako odstraněné písmeno a algoritmus pokračuje dál. V případě Hammingovy vzdálenosti, vynechané písmeno naruší porovnávání zbytku řetězce.

3.4.2 Levenshteinova vzdálenost

Levenshteinova vzdálenost v informatice je míra rozdílu mezi dvěma řetězci. Pro výpočet vzdálenosti se využívá matice s rozměry velikostí délek řetězců. V podstatě vzdálenost je minimální počet jednoznakových úprav ve slově, aby vzniklo slovo referenční. Za jednoznakovou úpravu se považuje smazání, záměna nebo vložení písmene na dané místo.

Vzorec 1 je matematický zápis Levenshteinovy vzdálenosti, kde a a b jsou řetězce $a(i)$ a $b(j)$ jsou indexované znaky v řetězcích. Výsledná vzdálenost je rovna hodnotě v matici $dist_{a,b}$ na pozici $dist_{a,b}(|a|, |b|)$, kde $|a|$ a $|b|$ jsou délky řetězců [8].

$$dist_{a,b}(i, j) = \begin{cases} \max(i, j) & \min(i, j) = 0 \\ \min \begin{cases} dist_{a,b}(i-1, j) + 1 \\ dist_{a,b}(i, j-1) + 1 \\ dist_{a,b}(i-1, j-1) + \begin{cases} 0 & a(i) = b(j) \\ 1 & a(i) \neq b(j) \end{cases} \end{cases} & \min(i, j) \neq 0 \end{cases} \quad (1)$$

V případě, že se pohybujeme v matici v prvním sloupci a prvním řádku přiřazujeme hodnoty 1 až délka řetězce a do sloupce a 1 až délka řetězce b do řádku. Tedy první sloupec a první řádek v matici slouží jako referenční inicializace. V generování matice se následně pokračuje a porovnávají se jednotlivé znaky. První položka v minimu je případ,

kdy písmeno bylo odstraněno z a . Druhá položka je případ, kdy došlo k vložení znaku do a na základě znaku v b řetězci. A poslední položka je případ porovnání znaků. V okamžiku, kdy znaky se rovnají, hodnota vzdálenosti na diagonále zůstává, v opačném případě se přičítá jednička.

Výpočet Levenshteinovy vzdálenosti dvou slov znázorňuje matice 2, kde dochází k porovnání českého slova *cyklus* a anglického výrazu *cycle*. Tučně je znázorněna cesta výpočtu. Prováděné jednoznakové operace jsou žádná, žádná, žádná, záměna, záměna a vložení.

		<i>c</i>	<i>y</i>	<i>c</i>	<i>l</i>	<i>e</i>
	0	1	2	3	4	5
<i>c</i>	1	0	1	2	3	4
<i>y</i>	2	1	0	1	2	3
<i>k</i>	3	2	1	1	2	3
<i>l</i>	4	3	2	2	1	2
<i>u</i>	5	4	3	3	2	2
<i>s</i>	2	5	4	4	3	3

(2)

3.4.3 Určení globální obtížnosti

Základním parametrem globální obtížnosti slova je jeho délka, která má váhu 4 v celkovém průměru jednotlivých hodnot obtížnosti. Na základě prahových hodnot je délka rozdělena do čtyř kategorií - $\langle 1, 6 \rangle$ je snadné, $\langle 6, 11 \rangle$ je střední a $\langle 11, \infty \rangle$ je slovo těžké. Dalším parametrem k určení obtížnosti je Levenshteinova vzdálenost mezi referenčním slovem a jeho překladem, jejíž váha je 2. Znovu zde figurují nastavitelné prahové hodnoty - v případě, že se procentuální podobnost (tj. $pp = (1 - a/l) * 100$; kde l je délka slova a a je počet jednoznakových editací) nachází v intervalu $\langle 70, 100 \rangle$, je obtížnost nejlehčí; v intervalu $\langle 30, 70 \rangle$ je střední a v intervalu $\langle 0, 30 \rangle$ je slovo těžké. Přehlasování znaku a dvojitý výskyt stejného znaku mají oba stejnou váhu 1. Bez výskytu přehlasování a dvojitého znaku je obtížnost kvalifikována jako snadná, jeden výskyt je střední obtížnost a více výskytů je obtížnost vyhodnocena jako těžká.

3.5 Generování slov pro procvičování

Způsob generování slov je důležité pro správné ukotvení slova v dlouhodobé a ne krátkodobé paměti. Při testování bude docházet i k několika násobnému opakování slova. Základem generování je častěji a vícekrát předkládat slova, která jsou pro žáka problematická a slova, která jsou zvládnuta žákem s přehledem, jsou méně častěji testována. Pro generování byla zvolena metoda rozloženého opakování v podobě přizpůsobeného Leitnerova algoritmu, který je popsán v kapitole 2.6.2.

Algoritmus pro procvičování slov se dělí na dvě části. V první části se zjišťuje, jaká slova jsou pro zkušného uživatele jednoduchá a naopak s jakými slovy má problém - tato fáze se bude dále nazývat inicializační. Po této fázi následuje fáze s použitím Leitnerova algoritmu - tato fáze bude nazývána testovací.

3.5.1 Inicializační fáze

Na začátku testování jsou k dispozici slova s přiřazenou obtížností. Tato obtížnost je buď globální v případě, že se uživatel nesetkal ještě s daným slovem v aplikaci nebo uživatelská v případě, že uživatel dané slovo již procvičoval. Dle této obtížnosti se slovíčka seřadí tak, aby se střídala nejnižší a nejvyšší obtížnost. Na prvním místě tedy je slovo s nejnižší obtížností, následuje slovo s nejvyšší, dále slovo s druhou nejnižší atd. Důvodem tohoto seřazení je zvýšení motivace dětí. V případě, že by žák dostal na začátku nejtěžší slova, mohlo by si vytvořit odpor k procvičování v domněnku, že slova jsou příliš složitá pro něj.

Uživatel tedy postupně odpovídá na slovíčka. Jeho odpověď je vyhodnocena buď správně, špatně nebo neúplně. Na základě vyhodnocené odpovědi a globální nebo uživatelské obtížnosti je ke slovu přiřazen počet správných opakování pro dokončení slova.

3.5.2 Testovací fáze

Po inicializační fázi následuje fáze testovací. V této fázi již je známo, jaká slova dělají uživateli problém. Oproti standardnímu Leitnerovu systému popsaném v kapitole 2.6.2, kde na začátku jsou všechny otázky zařazeny do první kategorie, zde se otázky rozdělují do třech kategorií na základě inicializační fáze. Následuje první fáze, kde dochází postupnému procházení otázek v první Leitnerově kategorii od nejtěžší pro uživatele po nejlehčí. V případě správně nebo neúplně zodpovězené otázky, slovo je přesunuto do vyšší kategorie. Následuje druhá fáze, kde se prochází otázky z první a druhé Leitnerovy kategorie. I zde platí, na základě odpovědi dochází k přesouvání otázek mezi kategoriemi, jak už tomu je ve standardním provedení Leitnerova systému. Ve třetí fázi jsou nabízeny otázky z první, druhé a třetí Leitnerovy kategorie. Po průchodu třetí fáze následuje znovu fáze první. Takto algoritmus cykluje do doby, dokud uživatel nepřeruší test nebo nezůstanou žádná slova, které by uživatel nedokončil.

Za dokončené slovo se považuje takové, kde uživatel odpověděl správně nebo neúplně tolikrát, kolik má slovo přiřazených nutných správných opakování, a zároveň poslední odpověď nesmí být neúplná, ale pouze správná. Je to z důvodu, aby bylo ověřeno, že uživatel má správně zapamatovanou informaci.

3.5.3 Adaptivní a globální obtížnost

V případě přerušení testu dochází k ukládání aktuální Leitnerovy kategorie k danému uživateli. Díky tomu lze znovu navázat v případě, že uživatel se bude chtít k testování dané skupiny slov vrátit. Leitnerova kategorie vypovídá, jak problematické je slovo pro uživatele. V případě, že se slovo nachází v první kategorii je problematické a v případě, že se nachází ve třetí, uživatel ho zvládá bez obtíží. Díky této informaci dochází k adaptivnímu generování obtížnosti. V případě, že není u daného slova a uživatele uložena Leitnerova kategorie, je v inicializační části brána globální obtížnost.

Například se může stát, že globální obtížnost slova je v kategorii těžkých, ale uživatel se s tímto slovem už několikrát setkal a dobře se mu pamatuje. Je proto zbytečné slovo mu opakovat vícekrát, jelikož na základě obtížnosti a správnosti odpovědi se generuje počet opakování.

3.6 Procvičování slovíček

Základní myšlenkou procvičování a testování slovíček je pomoci se žákům efektivně naučit slovíčka, nikoliv poskytnou vyučujícím možnost posuzovat, jak na tom žáci jsou. Následující podkapitoly popisují, jakým způsobem probíhá testování, napovídání, vyhodnocování odpovědi a kontrola podvádění.

3.6.1 Metody testování

Pro procvičování slovíček je zvolen pouze aktivní přístup a to v metodách aktivního rozpoznání a aktivního vzpomnutí. Tyto metody se automaticky střídají v závislosti na odpovídání uživatele. V případě inicializační fáze procvičování, popsané v kapitole 3.5.1, se metoda testování nemění, aby uživatel měl stejné podmínky pro vyhodnocení úvodní obtížnosti slova. V následující fázi, kde probíhá samotné procvičování se metody střídají. Když uživatel odpoví špatně na vzpomnutí a u slova má více než jednu nutnou správnou odpověď pro dokončení slova, následující procvičení stejného slova je metodou rozpoznání. Je to z důvodu zvýšení motivace, aby procvičování uživatele neodradilo, když si nevzpomene na cizojazyčný výraz.

Jelikož vzpomnutí je paměťově náročnější úkol pro uživatele, je nutné, aby metoda potenčně poslední odpovědi byla vzpomnutí. Jinak není možné ověřit, zda uživatel opravdu se slovo naučil i v cizojazyčném výrazu. Tzn. V případě, že počet nutných správných odpovědí byl vyhodnocen na číslo 2, první metoda testování bude rozpoznání - uživatel napíše význam slova v mateřském jazyce. Pokud odpoví správně následuje metoda vzpomnutí. Pokud ale odpoví špatně, následuje znovu metoda rozpoznání.

3.6.2 Nápořědy

V procvičování slovíček uživatel má k dispozici i nápořědy, v případě, že vlastník učebnice vyplní ke slořům i doplňující informace jako definici a použití ve větě z fyzické učebnice. Tyto nápořědy jsou použity pouze v případě, že se testuje metodou vzpomenutí. Další nápořědou je zobrazení prvního písmene odpovědi. Definice je nápořěda, kterou uživatelé vidí jako doplňující atribut otázky. V případě, že uživatel klikne na tlačítko "Použít nápořědu", zobrazí se mu první písmeno odpovědi. Pokud klikne na tlačítko znovu, zobrazí se mu použití slova ve větě.

3.6.3 Vyhodnocování odpovědi

K vyhodnocování odpovědi se využívá znovu Levenshteinovy vzdálenosti popsané v kapitole 1. Odpověď je vyhodnocena do tří kategorií - správně, špatně a neúplně. Před porovnáváním slova, je z odpovědi odstraněné mezery, tabulátory a další speciální znaky. Správně je vyhodnocena pouze odpověď, která se plně shoduje s referenčním slovem. Jako neúplná je vyhodnocena v případě, že na 5 písmen je pouze jedna jednoznaková editace - tedy chybovost 20%. V ostatních případech je odpověď vyhodnocená jako špatná. Pokud uživatel využije alespoň jednu z nápořěd, automaticky se vyhodnocuje odpověď jako neúplná i v případě zcela správné odpovědi.

3.6.4 Kontrola podvádění

Aplikace nemá za úkol testovat žáky, zda zodpovědně a bez pomoci procvičují slova. Filozofie aplikace má být motivovat žáky a poskytnout jim příjemný nástroj na efektivnější učení slovíček z cizího jazyka. Proto v aplikaci nejsou zakomponovány žádné mechanismy, které mají za úkol kontrolovat uživatele, zda si například nevyhledávají slovíčka ve slovníku nebo nepracují na procvičení společně s další osobou.

Jsou způsoby, jak by se ale tato funkčnost dala do aplikace implementovat a to například v podobě měření času odpovědi. Vyhodnocení by bylo na základě délky odpovědi a časem odpovědi. Další, více striktnější, omezení podvádění by mohla být kontrola, zda uživatel opustil aktuální okno aplikace.

3.7 Připomínání slov

Připomínání slov je důležitou součástí aplikace. Funkce slouží k tomu, aby studenti nezapomněli na již naučená a zvládnutá slova z procvičování. Jelikož se jedná o webovou aplikaci, není úplně možné připomínat slova v přesných intervalech přímo v aplikaci. Způsob implementace připomínání slov je závislý na tom, aby se uživatel sám přihlásil a spustil funkci

připomínání slov. Je možné na základě softwarového démonu Cron v určitých intervalech ze serveru odesílat emailová upozornění se slovy, které by si měl uživatel připomenout.

3.7.1 Supermemo algoritmus

Supermemo algoritmus je rovněž další implementace metody rozloženého opakování popsané v kapitole 2.6. První implementace Supermemo algoritmu v papírové podobě byla vytvořena již v roce 1985. V aplikaci je využívání aktuální verze 11, poprvé implementována v roce 2005.

Algoritmus má dvě možnosti implementace - optimální interval a pokročilé opakování [11]. Verze s optimálním interval je pro účely aplikace dostačující, jelikož pro implementaci s pokročilým opakováním by bylo nutné ukládat ke každému uživateli nepřehledné množství dat. Pro správnou funkci je nutné ukládat každou odpověď ke každému ze slovíček, na základě nichž se aplikace adaptuje na schopnosti uživatele. Oba algoritmy je možné si vyzkoušet v podobě desktopové aplikace popsané v kapitole 2.2.5. V aplikaci lze pozorovat změny na uživateli v podobě křivky zapomínání a dalších statistických informací.

Definice spočtení optimálního interval [11]:

$$I(1) = OF[1, L + 1] \tag{3}$$

$$I(n) = I(n - 1) * OF[n, AF] \tag{4}$$

kde:

- n počet již provedených připomenutí
- $I(n)$ je n -tý interval v řadě intervalů pro opakování
- AF je absolutní obtížnost
- L počet, kolikrát si uživatel nemohl vzpomenout na slovo
- OF matice optimálních koeficientů pro zvyšování intervalu
- $OF[1, L + 1]$ koeficient z matice na prvním řádku a $L+1$ sloupci
- $OF[n, AF]$ koeficient z matice, který koresponduje n -tému opakování a obtížnosti slova

Matice koeficientů k výpočtu optimálních intervalů je získána z freeware aplikace SuperMemo 15. Tato matice je statisticky nejvhodnější pro žáky základních tříd. Ukázku matice znázorňuje následující tabulka 2.

	1	2	...	8	9	10
1	2.48	2.22		1.12	1.00	0.89
2	1.20	1.80		5.40	6.00	6.60
3	1.20	1.50		3.30	3.60	3.90
4	1.20	1.40		2.60	2.80	3.00
...						
17	1.20	1.24		1.46	1.50	1.54
18	1.20	1.24		1.45	1.48	1.52
19	1.20	1.23		1.43	1.47	1.50
20	1.20	1.23		1.42	1.45	1.48

Tabulka 2: Matice koeficientů pro výpočet optimálních intervalů

3.7.2 Příklad řady optimálních intervalů

V případě, že si uživatel pokaždé správně vzpomene na naučenou informaci spočtené intervaly pro středně těžké slovo budou přibližně - 2 dny, 6 dní, 14 dní, 27 dní, 49 dní, 82 dní, 131 dní, 7 měsíců, 10 měsíců, 15 měsíců, 21 měsíců, 2,5 roku, 3,5 roku atd. V případě, že dojde k zapomenutí slova, tj. uživatel při připomenutí nebude moci slovo vybavit, interval se ze začátku zkrátí a začíná se v matici koeficientů na dalším sloupci - což znázorňuje parametr L v rovnici 3.

3.8 Motivace

Motivace je důležitou součástí, jelikož aplikace apeluje na vlastní iniciativu, popřípadě mohou být nuceni externě (mimo aplikaci) například vyučujícím. Aby se žáci rádi vraceli k aplikaci, je nutné zaujmout je designem nebo nějakou vnitřní psychologickou motivací. Další možností motivace je zařadit do aplikace nějakou formu hry. Například za milník naučených slov žáky odměnit například jednoduchou hrou. To by ale vyžadovalo složitější návrh aplikace.

3.8.1 Třídní počet zvládnutých slov

Jednou z vnitřních psychologických motivací je ukázání uživateli, že nějakým způsobem přispívá do kolektivu. Pokud je uživatel přihlášen do nějaké třídy nebo skupiny uživatelů, vidí v horním panelu počet zvládnutých slov v rámci třídy. Tento způsob má za následek to, že může dojít k předháněním mezi třídami, přičemž přispívání zvládnutými slovy na úrovni uživatelů je anonymní. V případě, že uživatel dokončí slovo v rámci dané třídy, aplikace ho upozorní v podobě zvukové signalizace a zvýrazní sekci, kde se nachází tyto statistické údaje.

3.8.2 Vlastní počet zvládnutých slov

Další vnitřní psychologickou motivací je zobrazení uživateli, kolik slov doposud v aplikaci zvládl procvičit. Tato informace je uživateli skryta ve výchozím stavu, pokud uživatel chce vidět jeho pokrok, musí kliknout na doplňující statistické informace. Důvodem skrytého výchozího stavu je to, aby uživatel neviděl, kolik úsilí je nutné vynaložit, aby slovo bylo korektně dokončeno. Mohlo by se totiž zdát, že slova do dokončených se načítají příliš pomalu, jelikož procvičení slova vyžaduje několika násobné opakování.

Dalším problémem je unikátnost slov, protože učebnice jsou editovatelné prostřednictvím různých vlastníků a do každé z nich lze zapsat identické slovo. Nelze proto počet zvládnutých slov vyhodnocovat pouze na základě úspěšně dokončených slov, ale na základě unikátních úspěšně dokončených slov. Unikátnost se provádí na referenčním slově v cizím jazyce. Problém ale vzniká, pokud nějaký uživatel při tvorbě slovíček zadává slovo včetně členu a jiný ne. Tato slova se posuzují jako dvě různá.

3.8.3 Design aplikace

Design aplikace má také velký vliv na motivaci žáků se do ní vracet. Pokud design disponuje dostatkem barev, různých designových prvků v podobě zvířat nebo jiných entit, ze kterých jsou děti odvázané. Aplikace je cílená pro děti základních škol zejména prvního stupně tedy ve věku 6 až 10 let, u nichž design hraje velkou roli. Další vlastností, kterou by měl design disponovat je responzivnost pro mobilní telefony, potažmo přímo dedikovanou aplikaci pro chytrý telefon. Aktuální trend je takový, že každý žák na první stupni ZŠ ovládá chytrý telefon.

Pro kvalifikovaný přístup k designu by se měla aplikace zadat profesionálnímu grafikovi, který by design přizpůsobil pro děti.

3.9 Architektura aplikace

Aplikace je implementovaná v podobě webové služby a skládá se z dvou částí - klientská a serverová. V návrhu aplikace dochází k vysokému sdílení dat mezi uživateli, proto je vhodné aplikaci implementovat jako webovou. Základní myšlenkou návrhu bylo oddělit klientskou a serverovou část, aby bylo možné do budoucna aplikaci rozšířit implementací například pro chytrý telefon.

Serverová část je zodpovědná za servírování klientské aplikace, rovněž aplikace poskytuje aplikační rozhraní, kde na základě dotazů lze získat data ve formátu JSON. V serverové části je implementovaný Supermemo algoritmus pro připomínání slov. Dále je server zodpovědný za získání zvukové nahrávky z Google Text to Speech API, ukládání její kopie a

její poskytnutí pro klienta. V neposlední řadě se server zabývá problematikou autorizace a autentizace uživatelů v podobě JavaScript Web Tokens (JWT).

Klientská aplikace je zodpovědná za zobrazení načtených dat ze serveru, poskytnutí přehledného grafického rozhraní pro tvorbu, editaci a mazání dat nutných pro procvičování a připomínání slov (slovíčka, učebnice, třídy apod.). V klientské části je implementovány algoritmy Leitnerova systému a spočtení Levenshteinovy vzdálenosti. Dále klient je zodpovědný za načítání obrázků z Google Search API.

4 Klientská část

Tato kapitola se zbývá klientskou částí aplikace od technologického návrhu aplikace, přes vývojové prostředí a vlastní implementaci, až po testování. Aplikace je napsána v jazyce JavaScript za podpory kódovacího jazyka HTML a kaskádovými styly CSS. Základní specifikace jazyků jsou doplněny o knihovny React s využitím MOBX rozšíření a o knihovnu pouze s CSS styly z Bootstrap frameworku.

4.1 Návrh aplikace

Využití webových technologií bylo jasnou volbou v návrhu aplikace ihned po specifikování požadavků, protože hlavním požadavkem bylo vytvořit portál pro žáky a učitele, kteří si budou moci sdílet mezi sebou učebnice se slovíčky, budou se moci přihlašovat do svých tříd a společně pracovat na zvýšení slovní zásoby.

Jelikož se mělo jednat o aplikaci, která do jisté míry měla nahrazovat desktopové aplikace, bylo nutné zvolit takovou technologii, která se nejvíce blíží statickým aplikacím a kde se nemusí čekat na obnovení celé stránky, ale kde dochází k okamžitým reakcím na uživatelské podněty. Na základě tohoto požadavku je využita technologie Single Page aplikace (SPA).

4.1.1 Single-page

Single page application (SPA) je jedna webová stránka, která poskytuje celou aplikaci. Při prvním načtení je přenesen veškerý kód pro aplikační běh, tj. HTML šablony, JavaScript i kaskádové styly (CSS). Na základě akcí prováděných uživatelem jsou poté dynamicky načítány ostatní zdroje jako data, obrázky apod. Celá stránka se nikdy neobnovuje jako tomu je v klasické webové aplikaci. Veškerá aplikační logika je přenesena ze serveru na klienta za účelem server odlehčit. V případě SPA server většinou po načtení aplikace poskytuje pouze data a to nejčastěji ve formátech JavaScript Object Notation (JSON) nebo Extensible Markup Language (XML).

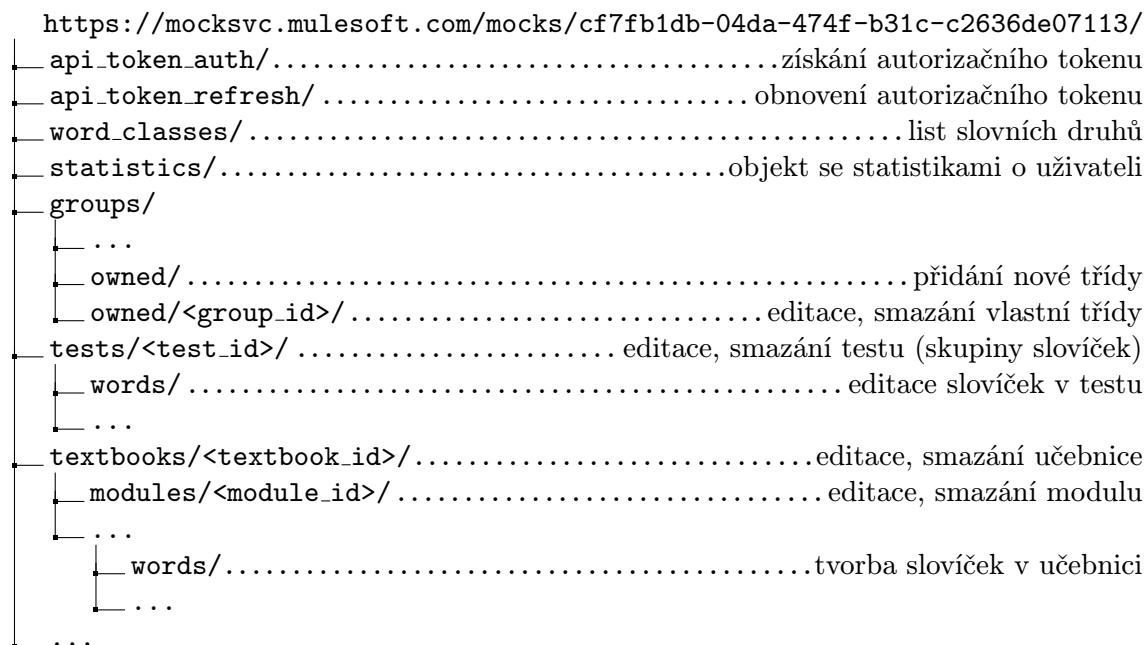
Hlavní výhoda SPA je kromě úvodního načítání dat rychlá odezva. Jelikož se ze serveru nestahuje celý HTML kód, nýbrž jen čistá data. Dochází tedy hlavně k datové i výkonové úspoře, protože prohlížeč nemusí renderovat celou stránku, ale pouze tu část, kde dochází ke změně. Výhodou je také možnost implementace tzv. lazy-loadingu, kdy dojde k rychlému vykreslení obsahu a náročnější operace se mohou být načteny nezávisle později. SPA má i řadu nevýhod. Například v prohlížeči pro chod aplikace je nutné mít zapnutý JavaScript, další nevýhodou je obtížné zpracování obsahu pro roboty vyhledávačů. Ale zařazení do vyhledávání je nepodstatné pro aktuální potřeby aplikace.

4.1.2 Komunikace se serverem

Klient komunikuje se serverem obvykle dvěma způsoby - web sokety nebo AJAX (Asynchronous JavaScript and XML) dotazy. Další technologií pro komunikaci, která by mohla být využita je Server-sent events (SSEs), která umožňuje serveru zasílat data klientovi přes HTTP protokol. Pro účely aplikace bude omezenou pouze na AJAX komunikaci, jelikož pro web sokety a tudíž obousměrnou komunikaci v aplikaci využití není.

4.1.3 Model aplikace

Před začátkem vývoje klientské aplikace bylo vytvořeno prozatímní aplikační rozhraní pomocí platformy Anypoint od firmy Mulesoft. Je to volně dostupná služba, kde lze velmi jednoduše, staticky a poměrně variabilně popsat serverové API v podobě souborů se syntaxí RAML (RESTful API Modeling Language). Díky této službě lze namodelovat klientskou aplikaci pomocí URL mapy a následně toto API zveřejnit a umožnit testování klienta. Na zdrojový kód s RAML syntaxí lze nahlédnout v příloze.



Obrázek 3: Ukázka struktury URL

4.1.4 Editovatelné seznamy

Pro jednotnost aplikace všechny data lze editovat přímo prostřednictvím seznamů. Vazba M:1 je v řádcích editovatelná prostřednictvím výběrových listů, vazby M:N a 1:M jsou editovatelné skrz vyskakovací okna s možnými návrhy. K získání potřebných dat dochází před načtením editovatelného seznamu. Obsah seznamu je vkládán přímo do editovatelného pole, v případě, že uživatel změní jeho obsah a odejde z buňky, obsah se automaticky

uloží. Ukládaný obsah na server je omezen pouze na změněnou položku a identifikátor pro optimalizaci přenášeného obsahu.

4.1.5 Design

Pro design aplikace byla využita čistá CSS knihovna z frameworku Bootstrap. Původní verze aplikace byla vyvíjena pod tématickou šablonou Yeti Bootswatch. Na základě konzultace bylo od této šablony opuštěno, jelikož působila příliš hranatě. Téma bylo vyměněno za Cerulean, které disponuje oblejšími prvky a výraznějšími barvami, aby žáky více zaujalo. Design aplikace je responsivní a připraven pro testování na tabletech a chytrých telefonech.

4.1.6 Adaptovaná Levenshteinova vzdálenost

Pro vyhodnocování odpovědi a určení podobnosti mezi referenčním slovem a jeho překladem se využívá Levenshteinovy vzdálenosti popsané v kapitole 1. Tato vzdálenost pro určení podobnosti slov musela být v jisté míře upravena, aby lépe vyhodnocovala podobnost slov mezi jazyky. Standardní Levenshtein, který měl parametry dva řetězce byl rozšířen o další nepovinné atributy - možnost odebrat diakritiku u znaků a možnost předat funkci slovník, kde klíč ve slovníku je písmeno v mateřském jazyce a hodnota ve slovníku je písmeno v cizím jazyce.

Pro příklad v anglickém jazyce písmeno *c* odpovídá písmenu *k* v českém. Po předání slovníku do funkce pro výpočet Levenshteinovy vzdálenosti dojde k porovnání znaků na daných pozicích a v případě, že nesouhlasí, funkce se pokusí porovnat s odpovídajícím písmenem ve slovníku.

```
1 let dictionary = [{'k':'c'}, {'ch':'š'}, ...]
2 "šance".levenshtein("chance", dictionary) // === 0
3 "šance".levenshtein("chance") // === 2
4 "korektní".levenshtein("correct", dictionary) // === 3
5 "korektní".levenshtein("correct") // === 5
```

Ukázka kódu 1: Příklady použití Levenshteinovy vzdálenosti

4.1.7 Google Search API

Pro získání obrázků se využívá Google Search API. Obrázky jsou získány v podobě odkazů. Uživatel si může vybrat z deseti obrázků, pokud žádný z nich nevyhovuje, vytvoří se nový dotaz s další stránkou. Pro možnost tvorby dotazů je nutné v Google API Console vytvořit vlastní API klíč, který se přidává jako atribut k URL. K API klíči je také nutné vytvořit

vlastní Google vyhledávač se svým identifikátorem. Tento atribut je rovněž přidáván jako value-key parametr dotazu. URL vlastního dotazu vypadá takto:

```
https://www.googleapis.com/customsearch/v1?q=<slovo>&searchType=image&cx=<id_vyhledavace>&key=<api_klic>
```

Odpověď na dotaz je JSON soubor s odkazy na jednotlivé obrázky. K dispozici jsou jak náhledové, malé obrázky (řádově několik desítek kB), tak obrázky v plné velikosti. Jelikož nelze zaručit permanentní přístup k souboru, po vybrání obrázku se na serveru vytvoří komprimovaná kopie obrázku s jednotnou šířkou.

4.1.8 Zabezpečení

Pro všechny funkce v aplikaci je nutné se přihlásit popřípadě registrovat. Uživatel vytvoří na základě svého jména a hesla účet, pod kterým bude v aplikaci vystupovat. Aplikace komunikuje pouze přes zabezpečenou komunikaci HTTPS. Aby nebylo možné heslo a jméno odposlechnout technikou man-in-the-middle. Jakmile je uživatel na základě hesla a jména ověřen, je mu zaslán JWT token, se kterým se identifikuje v hlavičce s každým dalším dotazem. Generování a ověřování JWT tokenu je popsáno v kapitole 5.3.1.

4.2 Vývojové prostředí

Pro vývoj SPA je důležitá správná konfigurace vývojového prostředí, jelikož dokáže hodně ovlivnit produktivnost programátora. Základními požadavky na vývojové prostředí je předzpracování kódu, automatické obnovení stránky po uložení editovaného souboru, vytvoření jednotného JS bundle souboru, který obsahuje všechny potřebné zdrojové kódy rozdělené do logických celků a poskytnutí HTTP serveru. Dále je vhodná pro produkční verzi aplikace například komprimace zdrojových kódů.

4.2.1 Webpack

Webpack je nástroj napsaný pro NodeJS, který umožňuje vytvořit požadované vývojové prostředí. Pro vývoj klientské aplikace jsou vytvořeny dvě webpack konfigurace - vývojová a produkční. Pro kompilaci JS souborů je využíván Babel kompilátor, který je popsán v následující kapitole. CSS soubory jsou v vývojové konfiguraci vkládány přímo jako inline styly do index.html souboru, aby bylo možné v reálném čase sledovat změny designu a nečekat na obnovení stránky, popřípadě emulovat určitý stav aplikace. V produkční verzi je vytvořen zvláštní CSS soubor, protože při načítání aplikace dochází k paralelnímu načítání CSS a JS souborů a dochází díky tomu k urychlení úvodního načtení aplikace. Pro vytvoření produkční verze je využívána pro kompilaci řada webpackových modulů jako

například *favicons-webpack-plugin* k tvorbě ikon pro řadu různých prohlížečů nebo *html-webpack-plugin* pro minifikaci (kompresi) a znemožnění přečtení vytvořeného JS souboru, aby nebylo úplně jednoduché pročítat si přehledný JS kód [12].

4.2.2 Babel

Babel je kompilátor, který transformuje rozšířený JavaScript do standardu, který většina prohlížečů podporuje - aktuálně ES5. Důvod využití tohoto kompilátoru je vysoká podpora EcmaScript2015 (ES6) a integrace JSX popsaných v následujících kapitolách. Babel je mimo jiné také doporučován komunitou zabývající se knihovnou React a používají ho společnosti jako Facebook nebo Evernote. Výhodou je, že Babel pouze rozšiřuje současnou verzi JavaScriptu a snaží se zachovat jeho dopřednou kompatibilitu [13].

4.2.3 Podpora ES6

Díky kompilátoru Babel lze využít novinky ze specifikace ES6, i když ještě není implementována do prohlížečů. Při vývoji nedílnou součástí bylo klíčové slovo *export*, které umožňovalo rozdělení kódů do mnoha modulů a souboru. V aplikaci je také hojně využívána nová syntaxe tříd, která umožňuje přehlednější dědičnost, tvorbu konstruktorů a statických metod. Novinek, které jsou využívány v aplikaci, je celá řada, za zmínku stojí zkrácená šipková syntaxe \Rightarrow , která umožňuje sdílet klíčové slovo *this* uvnitř logického bloku [14].

4.2.4 JSX

JSX je specifikace, která umožňuje kompilovat stromovou strukturu podobnou XML do objektů standardu ECMAScript [15]. Jelikož má stejnou strukturu jako XML, je možné předávat značkám atributy klíč-hodnota. Díky kompilaci této struktury lze tedy uložit do JS proměnných celé šablony HTML kódu. Na základě těchto šablon je poté renderováno grafické rozhraní.

4.3 Knihovna React

React je komunitně rozšířená JavaScriptová knihovna pro vytváření webových komponent. Její předností je efektivní aktualizace a renderování pouze těch komponent, ve kterých dochází ke změnám dat. Uživatel při jejím použití neprovádí změny přímo na DOM (Document Object Model) dokumentu, ale na jeho abstraktní vrstvě. Výhodou je, že uzly v abstraktní vrstvě mohou uchovávat jejich vnitřní stav. V případě jeho změny React zajistí automaticky a efektivně promítnutí změny v DOM dokumentu.

4.3.1 Komponenty

Prezentační vrstva aplikace lze rozdělit na jednotlivé logické celky v podobě jednoduchých a složitých komponent. Jednotlivé komponenty lze jednoduše vnořovat do sebe a tím vytvořit složitější interaktivní uživatelské rozhraní. Každá komponenta dědí od třídy *Component* v knihovně *React* a obsahuje funkci *render*, která vrací obsah, který bude renderovaný v dokumentu. Dále komponenty obsahují funkci *setState*, která mění její vnitřní stav a atributy *props* a *state*. *Props* obsahuje konstantní atributy předané komponentě a *state* obsahuje stavové objekty komponenty. V případě, že dojde ke zavolání *setState* funkce a tedy změně objektu *state*, dojde k zavolání funkce *render* a tudíž k překreslení obsahu komponenty [16].

```

1 export default class Hint extends React.Component {
2   showHint = e => {
3     this.setState({showHint: true})
4   }
5   render() {
6     const { definition, value, examStore } = this.props
7     const { showHint } = this.state
8     let firstLetter = value.substring(0, 1)
9     return <div className="hints">
10       <Button onClick={this.showHint}>Zobrazit nápovědu</Button>
11       {definition && <p>Definice: {definition}</p>}
12       {showHint && <p>První písmeno odpovědi: {firstLetter}</p>}
13     </div>
14   }
15 }

```

Ukázka kódu 2: Příklad komponenty pro zobrazování nápovědy při procvičování slovíček

4.3.2 MobX

MobX poskytuje mechanismus k ukládání a aktualizaci dat a tudíž i stavů aplikace [17]. Lze si ho představit jako oddělené úložiště dat, které zároveň obsahuje veškeré možné manipulace s daty. V případě, že v úložišti jsou data označena jako *@observable* a uživatel jejich obsah změní, dojde k překreslení všech komponent, ve kterých se daná *@observable* proměnná vyskytuje. To přináší React aplikaci možnost zcela oddělit datovou vrstvu aplikace od zobrazovací vrstvy a přiblížit se tak MVC modelu. V aplikaci je využitý MobX společně s dekorátory a anotacemi z ES.Next.

Všechny komponenty, které mění data MobX úložiště, musí být anotovány *@observer*. Aby bylo možné v komponentě data získat z úložiště, je využívána anotace *@inject*. V aplikaci jsou využívány i doplňkové funkce MobX jako například *@computed* atribut nebo *@action*

funkce.

V úložištích se rovněž provádí veškeré AJAX dotazy prostřednictvím externí knihovny *Axios*, která umožňuje mimo jiné generické nastavení hlaviček u jednotlivých HTTP dotazů.

4.3.3 Směrování a historie

Pro směrování a historii v SPA jsou využity externí knihovny z NPM (Node Package Manager) - *react-router* a *mobx-react-router*. Používat navigační tlačítka a historii prohlížeče umožňuje modul *syncHistoryWithStore*, který synchronizuje historii v podobě JS objektu s historií prohlížeče. K manipulaci historie prohlížeče se využívá JS funkcí *pushState* a *replaceState*, definovaných ve standardu HTML5.

Směrování v aplikaci umožňuje modul *Router* z knihovny *react-router*. Tento modul lze do sebe nesčetněkrát vnořovat a umožňuje lépe oddělit logické části aplikace. *Router* mimo jiné disponuje nepovinným atributem *onEnter*, ve kterém lze zajistit oprávnění pro různé sekce aplikace. Dalším důležitým modulem je *Provider* z knihovny *react-router*, která poskytuje pro komponenty Reactu úložiště popsané v 4.3.2. Následující ukázka zdrojového kódu 3 představuje hlavní vykreslovací funkci aplikace, jejíž první atribut je obsah a druhý je element, kam se má obsah vložit.

```

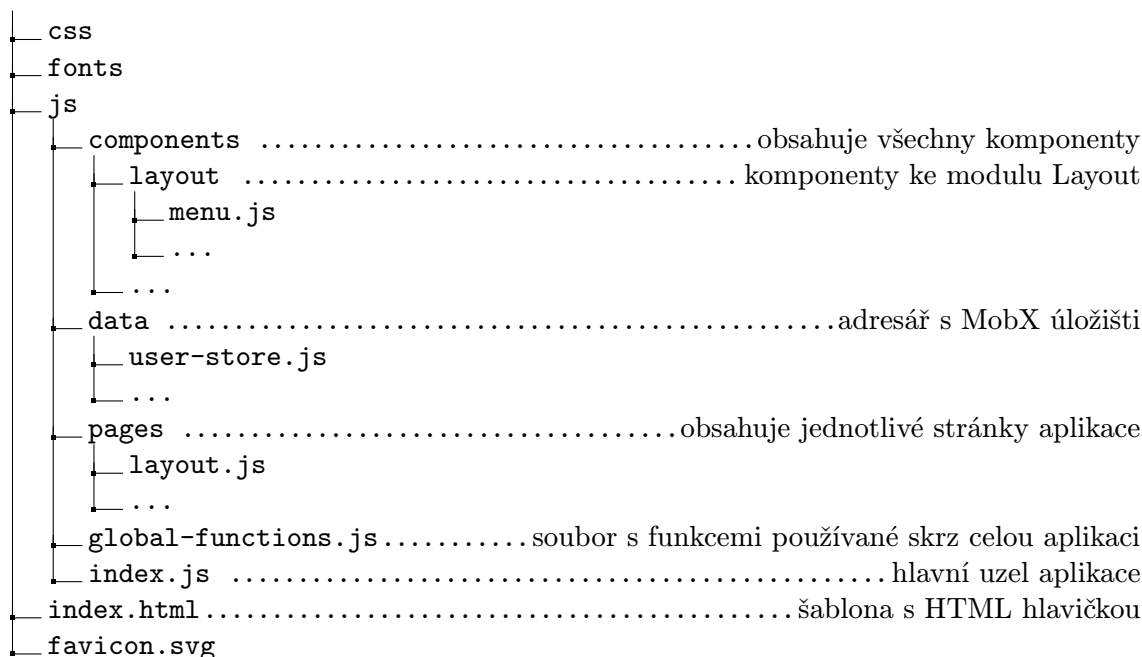
1 ReactDOM.render(
2   <Provider {...stores}>
3     <Router history={history}>
4       <Route path="/" component={Layout}>
5         <IndexRoute component={Welcome}/>
6         <Route path="/auth" component={Authentication}></Route>
7         ...
8         <Route path="/textbooks/:textbookId" component={
9           TextbookEditor} onEnter={requireAuth}></Route>
10       </Route>
11     </Router>
12   </Provider>,
13   document.getElementById('app'));
```

Ukázka kódu 3: Hlavní renderovací funkce aplikace

4.3.4 Adresářová struktura

4.4 Testování

Pro testování klientské části byl využity nástroje Enzyme a Jest. Enzyme je testovací nástroj, který umožní testovat React komponent. Jeho syntaxe je velmi podobná známé



Obrázek 4: Adresářové struktury klientské aplikace

knihovně JQuery. Jest je nástroj, který tyto testy spouští a vyhodnocuje. Pro kompilaci ES6 kódů byl využit balíček *babel-jest*.

4.4.1 Unit testování

Díky MobX úložišti je velmi jednoduché otestovat jednotlivé funkce, které manipulují s daty. MobX úložiště je JS objekt, který lze v jakémkoliv stavu aplikace uložit a pro potřeby testů znovu načíst. Díky této vlastnosti jsou jednotkové testy zcela odděleny od aplikace a lze je spouštět nezávisle. Enzyme umožňuje simulovat i JS události zavoláním funkce *simulate* na daný React element. První atribut je typ události a druhý atribut je objekt, který simuluje objekt události.

Ke každému úložišti byl vytvořen soubor s příponou *.test.js*, tyto testy jsou umístěny v adresáři v podadresáři *./data/__tests__*/. Pro testování jsou k dispozici funkce *describe*, *it* a *expect*. Enzyme je doplněn i o další standardní funkce jednotkového testování jako například *beforeEach*, která umožňuje načíst stejná data do úložiště pro každý test zvlášť.

4.4.2 Integrační testování

Pro testování, zda komponenty společně správně pracují, slouží integrační testování. Článek, který komponenty spojuje, je právě již zmiňované MobX úložiště. Každý z integračních testů obsahuje úložiště, do kterého jsou vložena data představující určitý stav v aplikaci. Toto úložiště je poté předloženo daným komponentám. Test následně provede

sadu operací nad komponentami a na závěr zkontroluje, zda výsledná komponenta vrací požadovaný obsah.

Důležitou součástí v kontrole obsahu je funkce *shallow*, která obalí JS objekt představující uzel v dokumentu funkcemi, které umožní jednoduché prohledávání jednotlivých elementů, jejich obsahu a atributů.

5 Serverová část

Serverová část aplikace je v podstatě aplikační rozhraní, které poskytuje data ve formátu JSON pro klientskou aplikaci. API je napsané v programovacím jazyce Python s využitím Django webového frameworku a Django REST frameworku.

5.1 Technologie

Důvod, proč bylo zvoleno Django a Django REST framework jako platforma pro aplikační rozhraní, je převážně z dřívějších zkušeností s variabilitou a rychlostí vývoje. Pro potřeby aplikace se jedná možná o zbytečně komplexní framework, jelikož aplikace jako celek stojí převážně na zpracování klientské části.

5.1.1 Django

Django je open-source webový framework, který využívá vlastní implementaci MVC modelu a označuje ho jako MTV (Model Template View) [18]. Z databázového modelu se získávají data, které jsou interpretovány v případě aplikace skrz Serializéry a následně jsou data podstrčeny podhledům, které v našem případě odešle uživateli JSON soubor s daty. Django je velmi variabilní a lze ho použít pro tvorbu standardní webové aplikace, tak za pomoci REST frameworku i tvorbu API.

Z čitého Django projektu je v aplikaci využívána hlavně modelová vrstva, která je implementována přes ORM (Object-Relational Mapping) přístup. U přístupu do databáze se nepíšou SQL dotazy, nýbrž se provádí operace nad objekty za pomoci definovaných funkcí v Pythonu. ORM implementované v Django transformuje objekty na dotazy, snaží se o jejich optimalizaci, a poskytuje nám vyšší abstrakci [18]. Využití Django ORM v aplikaci se detailněji zabývá kapitola 5.4.

5.1.2 Django REST framework

Webový framework Django je v aplikaci rozšířen o nadstavbu jménem Django REST framework. Jeho přínosy jsou rychlá tvorba vlastního HTTP nebo i RESTful API, připravené bezpečnostní principy s možností využití OAuth2. Dále obsahuje připravené serializační, parsovací a pohledové komponenty s možností generické a automatické validace dat a dalších užitečných funkcí [19].

5.1.3 HTTP a RESTful API

Klientská aplikace se dotazuje serveru pomocí požadavků metod GET, POST, PUT a DELETE. Jelikož je klient úzce svázaný se serverem, API nemusí nést principy HATE-OAS (Hypermedia as the Engine of Application State) ani RESTful rozhraní. Tedy takové, které mimo jiné poskytuje dynamicky vytvářené hypertextové odkazy pro manipulaci stavů aplikace. Toto rozhraní je vhodnější pro aplikace, kde veškeré logické části aplikace jsou prováděny na serveru a klient slouží pouze jako prohlížeč dat. Ve slovníkové aplikaci je naopak veškerá logika a stavovost přesunuta na klienta a API slouží pouze jako vrstva pro ukládání dat. Rozhraní neposkytuje ani odkazy pro navigaci mezi daty a to z důvodu optimalizace velikosti přenášených dat a jak bylo zmíněno klient je úzce svázaný se serverovým rozhraním. Z těchto důvodů by se dalo rozhraní klasifikovat spíše jako HTTP API.

5.2 API

Tato sekce poskytuje náhled na vnitřní strukturu a principy serverové části aplikace.

5.2.1 Struktura dat

Data odesílané klientovi jsou číté reprezentace atributů v modelové vrstvě. Každý z objektů je identifikován svým číslem datového typu *integer*. Co se týče relačních vztahů mezi modely, ty jsou reprezentovány pomocí zaslaného identifikátoru, popřípadě listu identifikátorů, pokud se jedná o násobný vztah. Pro získání doplňujících informací k zaslanému identifikátoru, je nutné vytvořit nový dotaz typu GET a v uvést identifikátor v parametru URL.

Z důvodu optimalizace počtu dotazů na server z určité části klientské aplikace bylo nutné nahradit relačně vztahové identifikátory vlastními daty, aby nedocházelo k několika násobným dotazům s velikostně velmi nízkou odpovědí. Například v případě získání listu učebnic, je klientovi zaslán list objektů, které obsahují atributy jako název učebnice, datum vytvoření, jazyk, identifikátor vlastníka atd. Aby klient nemusel dotazovat server několika násobně další informace o vlastníkově, jsou základní data o vlastníkově vnořena přímo do listu učebnice.

V Django REST frameworku tuto funkci zajišťují serializátory. Jsou to třídy, které transformují modelové objekty a dotazové sety do Pythonových data-typů tak, aby je bylo možné zapsat do formátu JSON nebo XML [19]. Následující ukázka 4 představuje serializer pro transformaci informací o učebnici. Výstupní objekt bude nést atributy, které jsou uvedeny v tuplu *fields*. Na základě definice *PrimaryKeyRelatedField* bude hodnota *language* atributu pouze identifikátor. Funkce *PrimaryKeyRelatedField* akceptuje atribut

queryset, který definuje list možných identifikátorů v podobě dotazu. Naproti tomu atribut *owner* bude ve výstupu objekt, který obsahuje právě dodatečné informace o vlastníkovi - id, uživatelské jméno, jméno a příjmení.

```

1 class OwnerSerializer(serializers.ModelSerializer):
2     class Meta:
3         model = User
4         fields = ('id', 'username', 'first_name', 'last_name')
5
6 class TextbookSerializer(serializers.ModelSerializer):
7     owner = OwnerSerializer()
8     language = serializers.PrimaryKeyRelatedField(
9         queryset=models.Language.objects.all())
10    class Meta:
11        model = models.Textbook
12        fields = ('id', 'title', 'language', 'owner')

```

Ukázka kódu 4: Serializační třída k transformaci dat o učebnici

5.2.2 Implementace Supermemo algoritmu

Algoritmus Supermemo popsaný v kapitole 3.7.1. pro výpočet optimálních intervalů je implementovaný na serverové části aplikace. V okamžiku, kdy uživatel dokončí neboli ovládne slovo v procvičování, se ukládá na server informace o tom, že slovo bylo pro daného uživatele dokončeno. V tomto okamžiku se spočítá absolutní obtížnost slova na základě adaptované obtížnosti a zjistí se první interval v podobě data, kdy má slovo být připomenuto. Tento datum společně s hodnotou 0 představující počet, kolikrát si uživatel nemohl vzpomenout na slovo, uloží do databáze. Pro uložení matice optimálních faktorů je v Pythonu vyžít standardní list listů.

Následně v okamžiku znovu přihlášení uživatele do aplikace, server společně s informacemi o statistikách uživatele odešle list slov, kde aktuální datum je vyšší než datum připomenutí. V aplikaci si uživatel může spustit mód připomenutí již naučených slov. V případě, že si uživatel nevzpomene na slovo, při ukládání informací o připomínání, se přičte jednička do hodnoty s počtem, kolikrát uživatel nemohl vzpomenout na slovo a pro každé slovo se přepočítá další datum připomenutí.

5.2.3 Získání zvukové nahrávky

Pro získání zvukové nahrávky slova je využitý Python modul s názvem gTTS. Stažení nahrávky je velmi jednoduché. Nahrávky se identifikují jazykem a názvem omezeným na ASCII znaky, což v ukázce 5 představuje funkce *slugify*. K uložení nahrávky dochází pouze v případě, že stejné slovo již stejné slovo není v adresáři uloženo. Klientovi se následně zasílají odkazy na nahrávku, kde si uživatel může slovo přehrát.

```

1 from gtts import gTTS
2 def save_sound(text, lang):
3     fname = slugify(text)
4     filename = 'static/sounds/%s/%s.mp3' % (lang, fname)
5     if not os.path.exists(filename):
6         tts = gTTS(text=text, lang=lang)
7         tts.save(filename)

```

Ukázka kódu 5: Uložení zvukové nahrávky

5.3 Zabezpečení

Serverová a klientská aplikace spolu komunikují skrz zabezpečený protokol HTTPS. Pro autorizaci uživatelů v rozhraní slouží standard JWT popsáný v následující kapitole. Kapitola dále rozvíjí, jak jsou uživatelé autorizováni k jednotlivým operacím a jaké bezpečnostní mechanismy jsou použity proti internetovým útokům.

5.3.1 JavaScript Web Tokens

JSON Web Token (JWT) je otevřený standard (RFC 7519) definující způsob posílání zabezpečené informace ve formě zakódovaného JSON objektu. Tyto informace jsou zabezpečeny buď s využitím tajného klíče a algoritmu HMAC a nebo pomocí páru privátní/-veřejný klíč a algoritmu RSA. Zašifrovaný obsah JWT tokenu obsahuje důležité informace o uživateli a klient tedy nemusí dotazovat server pro tyto informace, což vede k snížení režie serveru [22].

Způsob přihlášení uživatele probíhá následovně. Uživatel zadá do klientské aplikace jméno a heslo, které je dále interpretováno pouze jako MD5 hash. Tyto informace jsou zaslány na server. Server na základě ověření hesla a jména vygeneruje nový JWT token, který zašle zpět klientské části. Ta si token uloží do lokálního úložiště a přidává token ke každému HTTP dotazu v hlavičce se strukturou **Authorization: JWT <token>**.

Na serveru generování JWT obstarává knihovna *django-rest-framework-jwt*. Expirace tokenu je 3 hodiny, knihovna ale poskytuje možnost obnovení tokenu s novou expirační dobou. Na klientské části je obnovení řešeno Timeout callback funkcí, která obnoví token ještě než vyprší jeho expirace. Společně s tokenem se klientské části zasílají i statistické informace o uživateli. To zajišťuje vlastní implementace funkce *jwt_response_payload_handler* na následující ukázce 6.

```

1 def jwt_response_payload_handler(token, user=None, request=None):
2     user_info = models.UserInfo.objects.get(user=user)
3     return {
4         'token': token,
5         'statistics': UserInfoSerializer(user_info, context={'request':
6             request}).data
7     }

```

Ukázka kódu 6: Obsah metody *jwt_response_payload_handler*

5.3.2 CSRF a XSS

Cross Site Request Forgery (CSRF) je útok, kdy útočník donutí oběť odeslat upravený požadavek na server. V podstatě se jedná o nějaký podvržený odkaz na jiné doméně a jelikož uživatel je do aplikace přihlášen, má práva provést například změnu emailové adresy [20]. Proti tomuto typu útoku je aplikace chráněná díky JWT, jelikož s každým požadavkem na server je v hlavičce dotazu zaslán token. JWT nelze z jiné domény zjistit, jelikož prohlížeč pro každé okno vytváří samostatnou instanci pro provoz JavaScriptu, které mezi sebou data nesdílí.

Cross-site scripting (XSS) je případ, kdy útočník pomocí vstupního formulářového prvku zapíše JS kód do stránky, který se následně vykoná [21]. XSS je mnohem závažnější útok, jelikož útočník může získat například Cookie dokumentu, které může obsahovat například JWT. Jelikož klientská část funguje jako SPA, nedochází k obnovení stránky a tudíž nedochází k resetování JavaScriptové instance. JWT token je uložen pouze v JS objektu. Proto v případě, že uživatel z aplikace odejde, musí znovu projít autentizačním cyklem.

5.3.3 CORS

Pro testovací účely bylo nutné v serverové aplikaci a Django REST frameworku povolit CORS, jelikož server a klient nesdílely stejnou doménu. CORS (Cross-Origin Resource Sharing) je mechanismus, který spočívá v povolení klientům dotazovat rozhraní, které je poskytováno na jiné doméně. Pro povolení CORS v serverové aplikaci byl využit další middleware s názvem *corsheaders.middleware.CorsMiddleware*, ve kterém lze nastavit parametrem *CORS_ORIGIN_WHITELIST* soupis klientů, kteří mohou dotazovat server. V produkční verzi aplikace tento problém nenastává, jelikož serverová a klientská část sdílí servírovanou doménu.

5.3.4 Autorizace

Autorizace ke zdrojům je v aplikaci jednoduchá. Jelikož došlo ke sloučení rolí učitele a žáka, není nutné role v aplikaci rozdělovat. Proto všichni uživatelé mají přístup do všech sekcí

aplikace. Aby ale uživatel měl možnost vůbec dotazovat API, musí být identifikován a přihlášen. K nastavení a zprovoznění tohoto základního oprávnění posloužila třída, která je součástí REST frameworku *rest_framework.permissions.IsAuthenticated*, kterou je nutné uvést v nastavení pod atributem *DEFAULT_PERMISSION_CLASSES*.

Tato autorizace nestačí, jelikož je nutné také omezit editační a mazací metody pouze na vlastníky dat. Následující ukázka 7 představuje třídu oprávnění, která kontroluje, zda vlastník záznamu je přihlášený uživatel v případě, že se snaží provést nějakou úpravu. Pro správnou funkčnost je samozřejmě nutné, aby daný záznam přesněji objekt modelové třídy měl definovaný atribut *owner*.

```

1 class IsOwnerOrReadOnly(permissions.BasePermission):
2     def has_object_permission(self, request, view, obj):
3         if request.method in permissions.SAFE_METHODS:
4             return True
5
6         return obj.owner == request.user

```

Ukázka kódu 7: Ukázka třídy s přizpůsobeným oprávněním

5.4 Databáze

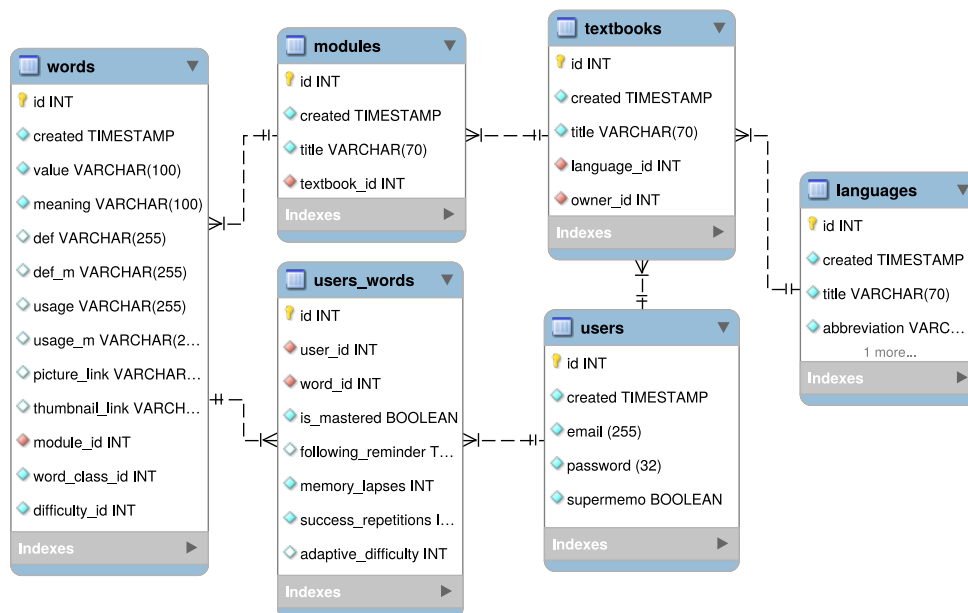
K ukládání dat v aplikaci je využita open-sourcová databáze PostgreSQL. Django podporuje celou řadu databázových backendů jako například MySQL, SQLite nebo lze nalézt i obecný konektor ODBC (Open Database Connectivity). Ale největší podporu ve vývoji frameworku dostává právě zmíněná databáze PostgreSQL [18].

Django jak už bylo zmíněno nepovoluje přímý přístup k databázi. Ve frameworku existuje abstraktní vrstva ORM, která mapuje objekty v Pythonu na databázové dotazy. To přináší spoustu výhod, jedna z nich je vysoká optimalizace dotazů prováděné na databázi, další může být například podpora databázových migrací. Django si sám spravuje strukturu tabulek v databázi. Potom co se implementují modelové třídy, je ve frameworku připraven nástroj, který vytvoří v databázi tabulky s atributy, indexy a se vším, co se v modelových třídách definuje. Django rovněž podporuje propagaci změn na databázi v případě, že dojde k editaci modelových tříd [18].

5.4.1 Model

Následující obrázek 5 znázorňuje část modelu databáze, která je zodpovědná za správu učebnice, slovíček a uživatelů. Za zmínku stojí tabulka *words*, která ukládá informace od slovu - překlad, význam, definice, odkazy na náhledový obrázek i ten v plné velikosti, globální obtížnost atd. Tabulka *users_words* slouží k ukládání informací o slovu ke každému uživateli. V případě, že uživatel se v procvičování potká s daným slovem, vytvoří se záznam

v této tabulce. S každou další interakcí se změní obtížnost slova interpretovaná atributem *adaptive_difficulty*. Tato tabulka rovněž ukládá informace o připomenutí slov. Atribut *memory_lapses* představuje počet, kolikrát uživatel zapomněl slovo, *success_repetitions* značí, kolikrát uživatel byl schopný si úspěšně vzpomenout na slovo a *textitfollowing_reminder* představuje čas, kdy by mělo dojít k dalšímu připomenutí.



Obrázek 5: Ukázka části databázového modelu zodpovědné za správu slovíček

5.5 Testování

Testování rozhraní bylo zaměřené pouze na správnou funkčnost API, tzn. kontrola, zda jednotlivé dotazy vracejí správná data. Jednotkové testování krátkých logických funkcí nemělo příliš význam, jelikož v REST frameworku jde spíše o nastavení součástí a spojení je do jednotného celku než implementace nízkourovňových logických bloků.

Pro testování správné funkčnosti API byla využita třída `APITestCase` poskytovaná REST frameworkem. Testy jsou napsány na každou metodu každé URL. Kontrola jedné metody URL spočívala v ověření stavového kódu HTTP, dále v počtu vrácených dat, pokud se jednalo o listy objektů a náhodná kontrola hodnot objektu. K provádění dotazů posloužil objekt třídy `APIClient`, který umožňoval jednoduše editovat hlavičky dotazu pro celou testovací sadu.

Užitečným nástrojem pro testování jsou tzv. *fixtures*. Jsou to soubory ve formátu JSON, které obsahují databázová data. Na začátku každé testovací sady je vytvořena nové dočasné schéma v databázi, které má identickou strukturu s produkčním schématem v databázi. Dále jsou na začátku testu předány jednotlivé *fixtures*, které pro účely testování naplní databázi požadovanými daty.

6 Závěr

Použitá literatura

- [1] Terasoft a.s. *Terasoft - Výchovné programy* [online] 2002-10-07. [cit. 2016-12-15]. Dostupné z: http://www.terasoft.cz/czpages/cd_aj15.htm
- [2] LangSoft s.r.o. *Language Teacher* [online]. [cit. 2016-12-07]. Dostupné z: <http://www.langsoft.cz/teacher.htm>
- [3] KREJČOVÁ Lenka. *Psychologické aspekty vzdělávání dospívajících*. 1. vyd. Praha: Grada Publishing, 2011 [cit 2016-12-18]. ISBN 978-80-247-3474-3.
- [4] BIEMILLER Andrew, BOOTE Catherine. *An effective method for building meaning vocabulary in primary grades*. Vol 98(1), Journal of Educational Psychology, 2006 [cit 2016-12-18].
- [5] LOZANOV Georgi. *Suggestology and Outlines of Suggestopedy*. 1. vyd. Gordon and Breach, 1978 [cit 2016-12-18]. ISBN 0-203-39282-5.
- [6] I. S. P. Nation *Learning Vocabulary in Another Language*. Cambridge University Press, 2001 [cit 2016-12-18]. ISBN 0-521-800927.
- [7] SVOBODOVÁ Jana, SVOBODOVÁ Diana, KULDANOVÁ Pavlína, GEJGUŠOVÁ Ivana, ROSOVÁ Milena, NOVÁK Radomil. *Lexikologie* [online] 2003. [cit. 2016-12-19]. Dostupné z: <http://www.osu.cz/fpd/kcd/dokumenty/cestinapositi/lexikologie.htm>
- [8] Wikimedia Foundation, Inc. *Levenshtein distance* [online] 2016-12-20. [cit. 2016-12-21]. Dostupné z: https://en.wikipedia.org/wiki/Levenshtein_distance
- [9] Google, Inc. *Google Cloud Platform* [online]. [cit 2016-12-21]. Dostupné z: <https://cloud.google.com/>
- [10] Hermann Ebbinghaus *Memory: A Contribution to Experimental Psychology*. New York, 1913 [cit 25-12-2016]. ISBN 978-16-142-7166-6.
- [11] WOZNIAK Piotr A. *Repetition spacing algorithm used in SuperMemo 2002 through SuperMemo 2006* [online] 2006-04-02. [cit. 2016-12-26]. Dostupné z: <https://www.supermemo.com/english/algs11.htm>
- [12] KOPPERS Tobias a spol. *Webpack 2 Documentation* [online] 2016. [cit. 2016-12-26]. Dostupné z: <https://webpack.js.org/configuration/>
- [13] SUSCAK Marek a spol. *Learn ES2015* [online] 2016. [cit. 2016-12-27]. Dostupné z: <https://babeljs.io/learn-es2015/>
- [14] Ecma International. *ECMAScript® 2015 Language Specification* [online] 2015 [cit. 2016-12-27]. Dostupné z: <http://www.ecma-international.org/ecma-262/6.0/index.html>.

- [15] Facebook Inc. *JSX In Depth* [online] 2016 [cit. 2016-12-28]. Dostupné z: <https://facebook.github.io/react/docs/jsx-in-depth.html>.
- [16] Facebook Inc. *React Components and Props* [online] 2016 [cit. 2016-12-28]. Dostupné z: <https://facebook.github.io/react/docs/components-and-props.html>.
- [17] WESTSTRATE Michel a spol. *MobX Core Concepts* [online] 2016 [cit. 2016-12-28]. Dostupné z: <https://mobxjs.github.io/mobx/index.html>.
- [18] Django Software Foundation. *Django documentation* [online] 2015 [cit. 2016-12-28]. Dostupné z: <http://www.docs.djangoproject.com/en/>.
- [19] CHRISTIE Tom. *Django REST framework* [online] 2015 [cit. 2016-12-28]. Dostupné z: <http://www.django-rest-framework.org/>.
- [20] Open Web Application Security Project. *Cross-Site Request Forgery* [online] 2016-11-02 [cit. 2016-12-29]. Dostupné z: https://www.owasp.org/index.php/Cross-Site_Request_Forgery_%28CSRF%29
- [21] Open Web Application Security Project. *Cross-site Scripting (XSS)* [online] 2016-11-02 [cit. 2016-12-29]. Dostupné z: <https://www.owasp.org/index.php/XSS>
- [22] Auth0 Inc. *Introduction to JSON Web Tokens* [online] 2016-11-02 [cit. 2016-12-29]. Dostupné z: <https://jwt.io/introduction/>