

Univerzita Pardubice
Fakulta elektrotechniky a informatiky

Webová aplikace na stahování obsahu ze sociální sítě Reddit

Bakalářská práce

2023

Josef Knotek

Univerzita Pardubice
Fakulta elektrotechniky a informatiky
Akademický rok: 2022/2023

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Josef Knotek**
Osobní číslo: **I20113**
Studijní program: **B0688A140009 Informační technologie**
Téma práce: **Webová aplikace na stahování obsahu ze sociální sítě Reddit**
Zadávající katedra: **Katedra informačních technologií**

Zásady pro vypracování

Cílem práce je vytvoření webové aplikace, která umožní uživatelům hromadně stahovat příspěvky ze sociální sítě Reddit. Uživatel bude moci specifikovat preference stahování (umístění, adresářová struktura, pojmenování souborů). Aplikace bude využívat databázi na uložení uživatelských preferencí. Aplikace získá příspěvky ze sítě Reddit pomocí Reddit API. Webová stránka bude vytvořena pomocí knihovny ASP.NET a jazyku C#. V teoretické části se bude pojednávat o využitých technologiích a jejich alternativách. V závěru teoretické části bude uvedena analýza návrhu a realizace aplikace.

Rozsah pracovní zprávy: **min. 30 stran**
Rozsah grafických prací:
Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam doporučené literatury:

BAPTISTA, Gabriel a Francesco ABBRUZZES. Software Architecture with C# 10 and .NET 6: Develop software solutions using microservices, DevOps, EF Core, and design patterns for Azure, 3rd Edition: Develop software solutions using microservices, DevOps, EF Core, and design patterns for Azure. 3rd Edition. Birmingham: Packt Publishing, 2022. ISBN 180323525X.

FREEMAN, Adam. Pro ASP.NET Core 3: Develop Cloud-Ready Web Applications Using MVC 3, Blazor, and Razor Pages. 8th edition. New York: Apress, 2020. ISBN 9781484254400.

WONG, Clinton. HTTP Pocket Reference: Hypertext Transfer Protocol. Sebastopol, California: O'Reilly Media, 2000. ISBN 9781565928626.

Vedoucí bakalářské práce: **Ing. Jan Merta, Ph.D.**
Katedra softwarových technologií

Datum zadání bakalářské práce: **16. prosince 2022**

Termín odevzdání bakalářské práce: **12. května 2023**

L.S.

Ing. Zdeněk Němec, Ph.D. v.r.
děkan

Ing. Jan Panuš, Ph.D. v.r.
vedoucí katedry

Prohlašuji:

Práci s názvem *Webová aplikace na stahování obsahu ze sociální sítě Reddit* jsem vypracoval samostatně. Veškeré literární prameny a informace, které jsem v práci využil, jsou uvedeny v seznamu použité literatury.

Byl jsem seznámen s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze 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, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Beru na vědomí, že v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů, a směrnicí Univerzity Pardubice č. 7/2019 Pravidla pro odevzdávání, zveřejňování a formální úpravu závěrečných prací, ve znění pozdějších dodatků, bude práce zveřejněna prostřednictvím Digitální knihovny Univerzity Pardubice.

V Pardubicích dne 04. 05. 2023

Josef Knotek

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu mé bakalářské práce panu Ing. Janu Mertovi, Ph.D. za vedení práce a za cené rady. Také srdečně děkuji své rodině, která mi poskytovala podporu, nejen při realizaci mé bakalářské práce, ale i během celého mého studia.

ANOTACE

Cílem práce je vytvoření webové aplikace, která umožní uživatelům hromadně stahovat příspěvky ze sociální sítě Reddit. Uživatel bude moci specifikovat preference stahování (umístění, adresářová struktura, pojmenování souborů). Aplikace bude využívat databázi na uložení uživatelských preferencí. Aplikace získá příspěvky ze sítě Reddit pomocí Reddit API. Webová stránka bude vytvořena pomocí knihovny *ASP.NET* a jazyku *C#*.

KLÍČOVÁ SLOVA

webové aplikace, sociální sítě, API, stahování, *ASP.NET*, *C#*, Reddit

TITLE

Web application for bulk downloading of content from social site Reddit

ANNOTATION

The goal of this work is creation of web application, which will allow users to download posts in bulk from social site called Reddit. Users will be able to specify preferences of the download (location, folder structure, naming of files). Application will use a database for saving user preferences. Application will fetch the posts from Reddit with Reddit API. The web page will be created with the help of a library called *ASP.NET* and language *C#*.

KEYWORDS

web applications, social sites, API, downloading, *ASP.NET*, *C#*, Reddit

OBSAH

SEZNAM ILUSTRACÍ A TABULEK.....	9
SEZNAM ZKRATEK A ZNAČEK	10
SEZNAM ZDROJOVÝCH KÓDŮ.....	11
ÚVOD.....	12
1 TEORETICKÁ ČÁST	13
1.1 Existující řešení.....	13
1.2 API	14
1.3 Reddit a Reddit API	14
1.3.1 Sociální síť Reddit	14
1.3.2 Reddit API	15
1.4 Jazyk C#.....	16
1.5 Webové frameworky	17
1.5.1 ASP.NET Core.....	18
1.5.2 Razor Pages.....	18
1.6 Microsoft SQL Server	20
1.7 Teoretická analýza	20
2 PRAKTICKÁ ČÁST	22
2.1 Praktická analýza	22
2.1.1 Použití API.....	22
2.1.2 Vývojové prostředí	25
2.2 Databáze.....	25
2.3 Stránky	26
2.3.1 Index	26
2.3.2 LoggedIn.....	27
2.3.3 Download/Loading	27

2.3.4	Download/Select.....	28
2.3.5	Download/Structure	29
2.3.6	Download/Progress	29
2.3.7	Statistics	30
2.4	Flow aplikace	30
2.5	Přihlášení.....	31
2.6	Získání příspěvků.....	33
2.6.1	Metoda GetAllSavedPosts	33
2.7	Filtrování a zobrazení příspěvků.....	35
2.7.1	Filtrování na stránce „Download/Select“	35
2.7.2	Zobrazení na stránce „Download/Structure“	37
2.8	Stahování příspěvků.....	38
2.8.1	Historie stahování	41
2.9	Šablony.....	42
2.10	Statistika	42
ZÁVĚR		44
POUŽITÁ LITERATURA		45
SEZNAM PŘÍLOH.....		47
PŘÍLOHA A: Uživatelský návod		48
PŘÍLOHA B: Aplikace Reddit Online Downloader.....		54
PŘÍLOHA C: DDL skript na vytvoření databáze		55

SEZNAM ILUSTRACÍ A TABULEK

Obrázek 1 - logo aplikace Reddit	14
Obrázek 2 - domovská stránka Redditu	15
Obrázek 3 - endpointy Reddit API týkající se uživatelů	16
Obrázek 4 - složka Pages v novém projektu ASP.NET Core s Razor Pages	19
Obrázek 5 - příklad webové aplikace	23
Obrázek 6 - záhlaví stránek bez přihlášeného uživatele	26
Obrázek 7 - záhlaví stránek s přihlášeným uživatelem	26
Obrázek 8 - domovská stránka aplikace s přihlášeným uživatelem	27
Obrázek 9 - stránka Download/Loading	27
Obrázek 10 - stránka Download/Select	28
Obrázek 11 - stránka Download/Structure	29
Obrázek 12 - stránka Download/Progress	30
Obrázek 13 – Reddit formulář na schválení práv	32
Obrázek 14 – příklad dat v tabulce Users	33
Obrázek 15 - formulář na filtrování příspěvků	35
Obrázek 16 - formulář na pojmenování a strukturu složek	37
Obrázek 17 – ukázka šablon v tabulce „Templates“	42

SEZNAM ZKRATEK A ZNAČEK

API	Application Programming Interface
HTTP	Hypertext Transfer Protocol
SQL	Structured query language
T-SQL	Transact-SQL
MSSQL	Microsoft Server SQL
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
JSON	JavaScript Object Notation
CSS	Cascading Style Sheets
HTML	Hypertext Markup Language
MEF	Microsoft Entity Framework
SFW	Safe For Work
NSFW	Not Safe For Work
CLI	Common Language Infrastructure

SEZNAM ZDROJOVÝCH KÓDŮ

Zdrojový kód 1 - ukázkový Razor Page	19
Zdrojový kód 2 - část razor page dynamicky generující obsah	20
Zdrojový kód 3 - část metody GetAllSavedPosts odpovědná za získání příspěvků z uživateleova Reddit účtu	34
Zdrojový kód 4 - filtrování příspěvků v Javascriptu.....	36
Zdrojový kód 5 - funkce redraw ze souboru „Structure.cshtml“	38
Zdrojový kód 6 - metoda pro ukládání domén i.redd.it a i.imgur.com.....	40
Zdrojový kód 7 - metoda pro ukládání textových příspěvků	41

ÚVOD

V několika posledních letech existuje trend, ve kterém uživatelé chytrých mobilních telefonů tráví více času ve stále menším počtu mobilních aplikací [1]. Tohoto trendu tudíž mnozí vývojáři využívají a vytváří pomocné aplikace k těmto nejvíce používaným aplikacím. Jako uživatel Redditu jsem si všiml možné příležitosti, která zatím není komerčně využita.

Reddit je sociální síť, která umožňuje uživatelům zveřejňovat a sdílet příspěvky o různých tématech. Tyto příspěvky si mohou uživatelé ukládat, nicméně je toto ukládání pouze dočasné, jelikož Reddit povoluje uložení maximálně 1000 příspěvků. Pokud si uživatel uloží více, pak je nejstarší příspěvek odstraněn. Pokud si uživatel chce tyto příspěvky trvale uložit, tak mu Reddit nenabízí skoro žádný vhodný nástroj. Obrázky se třeba dají jednoduše stahovat, ale videa stahovat nelze. Textové příspěvky také není možné stahovat, leda by si je uživatel manuálně kopíroval.

Průzkum existujících řešení našel některé možnosti, jak hromadně stahovat obsah, nicméně zde nebylo žádné řešení, které bylo jednoduché na použití (většina kvalitních řešení vyžaduje složitou instalaci či náročné prvotní nastavování).

Cílem této práce je tedy nabídnout možnost hromadného stahování příspěvků z pohodlí webové stránky, bez potřeby zbytečného nastavování či instalace ze strany uživatele. Zároveň je cílem této práce také nabídnout vylepšení k obyčejnému stahování (zaznamenávání již stažených příspěvků, filtrování, struktura stažených příspěvků apod.).

Tato práce bude vytvořena pomocí *ASP.NET Core s Razor pages*, což je framework pro vývoj webových stránek. Tento framework bude používán v kombinaci s jazykem C#. K získání dat z Redditu bude použito Reddit API. Pro uchování dat aplikace bude využito *Microsoft SQL Serveru*, tedy relační databáze založené na jazyku *T-SQL* (Transactional-SQL), což je rozšíření jazyka SQL.

1 TEORETICKÁ ČÁST

1.1 Existující řešení

Na internetu zatím není dostupné webové řešení tohoto problému. Na internetu jsou alespoň některá řešení. Mezi ty manuálnější například patří Reddit bot „u/savevideo“ nebo třeba jednoduchý online nástroj na stažení konkrétních videí, kterým je webová stránka „<https://redditsave.com>“. Tyto a podobné nástroje zatím slouží jenom ke stahování konkrétních příspěvků, ale neslouží k hromadnému stahování.

Existují však i programy, které se dají použít pro hromadné stahování. Jedním z nich je „Reddsaver“. Provedením je Reddsaver kvalitní, ale problém je v jeho složité instalaci a v používání. Instalace je dostupná z githubu. Poté je ještě potřeba doinstalovat „ffmpeg“, následně musíte založit vlastní reddit aplikaci a nakonec se program spouští z konzole. Tento přístup není dostupný široké veřejnosti. [2]

Trochu dostupnějším řešením je kombinace webové stránky „Reddit Manager“ a programu „JDownloader 2“. „Reddit Manager“ dokáže získat odkazy na média uložených příspěvků a jelikož se jedná o webovou stránku, tak je jeho použití relativně jednoduché. Ale postup se stává komplikovanější ve chvíli, když dojde řada na program „JDownloader 2“, který vyžaduje stažení a instalaci. Program jako takový není příliš přehledný. V tomto programu bude muset uživatel naimportovat stažený HTML soubor z „Reddit Downloader“ a následně zahájit stahování. [3]

Za zmínku také stojí služba „eternity“. Tato služba umožní uložené příspěvky zachovávat i přes limit jednoho tisíce. Problémem je, že prvotní nastavení této služby obsahuje vytvoření vlastní databáze na Google Firebase. Také je třeba zmínit, že tato služba vám žádné média stahovat nebude, ukládá do databáze jenom informace o příspěvcích, které je možné si prohlížet na jejich webové stránce <https://eternity.portals.sh/>. [4]

Na internetu se dají najít i jiná řešení, ale žádné není jednoduše použitelné širokou veřejností. Většina technicky kvalitních řešení vyžaduje jistou úroveň IT znalostí. Proto vznikl nápad této aplikace. U této aplikace by uživatel neměl mít problém dosáhnout hromadného stažení bez větších znalostí ze světa IT a zároveň bez potřeby stahovat a instalovat programy.

1.2 API

API je zkratka pro Application Programming Interface. Česky to znamená Rozhraní pro programování aplikací. Je to sada definovaných protokolů a pravidel, které dovolují různým programům a aplikacím vzájemnou komunikaci. API slouží jako prostředník mezi systémy, zpracovává přenos dat a umožňuje společnostem zpřístupnit svá data a funkce aplikacím externích vývojářů, obchodních partnerů a interních oddělení.

API není určeno pro přímé používání koncovými uživateli. API využívají programátoři, kteří je mohou integrovat do svých aplikací. Metody, ze kterých se API skládá, se nazývají koncové body (endpoint). API skrývá své vnitřní fungování. Odkrývá jen takové části, které programátor může shledat užitečnými a tyto části zachovává konzistentní i přes budoucí vnitřní změny. [5]

Důvod, proč je API tolik používaná, je její jednoduchost pro implementaci. Například pokud by vývojář chtěl implementovat do své webové stránky mapu, která zobrazí zákazníkům lokaci jeho podniku, tak vývojář nemusí vůbec znát fungování map. Vše co potřebuje vědět, je forma požadavku, který musí na API poslat.

1.3 Reddit a Reddit API

1.3.1 Sociální síť Reddit

Reddit je sociální síť, která umožňuje uživatelům zveřejňovat a sdílet obsah o různých tématech. Každé téma má svůj vlastní subreddit. Subreddit je uživatelsky spravovaná skupina pro uživatele, které zajímá právě dané téma subredditu. Tato témata mohou být od aut přes roztomilá zvířata až po tipy, jak ušetřit peníze.

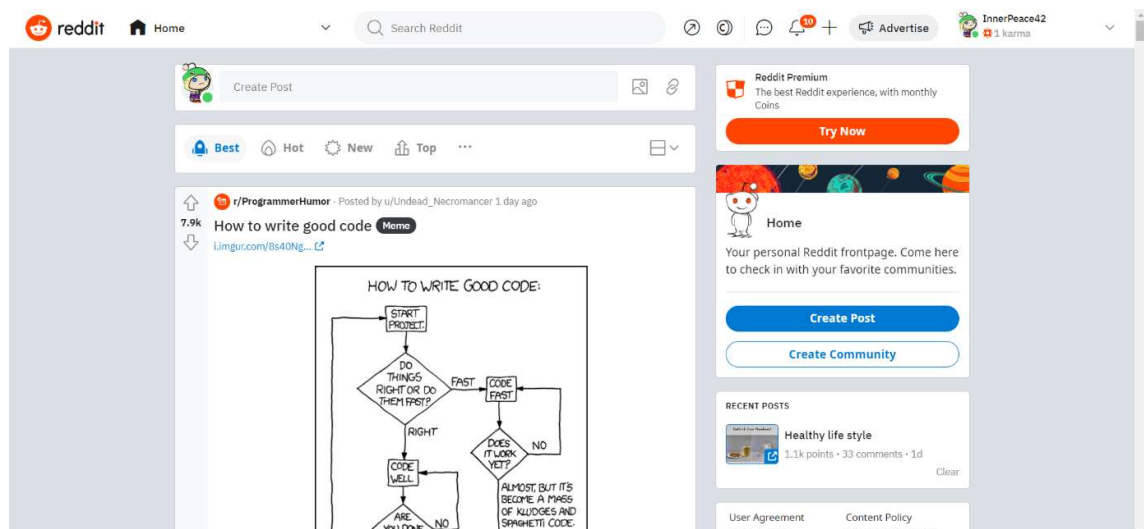


Obrázek 1 - logo aplikace Reddit

Registrovaní uživatelé mohou zveřejňovat příspěvky, odpovídat u příspěvků, ukládat si obsah nebo třeba hlasovat. Reddit také podporuje sponzorované příspěvky, tedy reklamu ve formě příspěvku označeného nápisem „Promoted“. Dalším prvkem monetizace platformy je Reddit

měna, za kterou můžete jiným uživatelům darovat ocenění, které nemá žádný funkcionální význam, až na to, že obdarovaný uživatel může obdržet nějakou odměnu (reddit měna, pár dní procházení bez reklam). Posledním prvkem monetizace je Reddit premium.

Příspěvek na Redditu je základním obsahem sociální stránky. Jakmile Reddit otevřete, první co vidíte, je seznam příspěvků, které vám algoritmus doporučil. Ty se berou jak ze subredditů, kterých je uživatel členem, ale i z jiných subredditů, které by vás mohly zajímat. Příspěvek má vždy nějaký nadpis, neboli titulek, a následně určitý obsah. Tento obsah může být různorodý. Obsahem může být například text, odkaz, obrázek, obrázky, video a další. Nejčastějším obsahem na síti jsou redditem hostované obrázky, videa a textové příspěvky. Další populární druhy obsahu jsou obrázkové galerie, odkazy na články a videa či obrázky hostované na různých jiných webech.



Obrázek 2 - domovská stránka Redditu

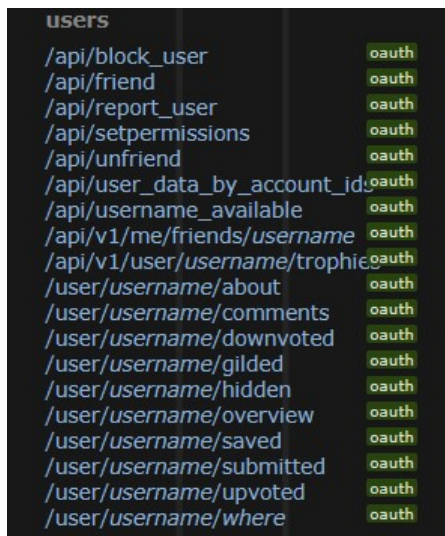
Reddit je poměrně velkou sociální sítí a to hlavně ve Spojených státech amerických. Na konci roku 2022 měl Reddit 57 miliónů aktivních denních uživatelů. Zároveň měl ve stejném období Reddit přes 100 tisíc aktivních subredditů. [6]

1.3.2 Reddit API

Reddit API slouží vývojářům k interakci s obsahem a funkcionalitou Redditu. API umožňuje vývojářům vytvářet aplikace a jiné nástroje, které dokáží automatizovat některé postupy na platformě. Nejvíce viditelným příkladem využití API na platformě jsou boti. Ti podle svého programu dokáží provádět různé činnosti, nejčastěji odpovídají na příspěvky a komentáře.

Reddit API používá HTTP dotazy a odpovědi pro výměnu informací mezi Reddit serverem a klientskou aplikací. Pro přístup k API si vývojáři musí nechat ověřit aplikaci pomocí OAuth2. Toto zajistí, že pouze ověřené aplikace mají přístup k datům z Redditu.

K API také existuje řada komunitně vytvořených wrapperů. Nejznámějším wrapperem mezi vývojáři pro Reddit je PRAW (Python Reddit API Wrapper). Ale existují i další jako jsou: RedditSharp (wrapper pro C#), raw.js (wrapper pro Node.js), graw (knihovna na vytváření botů pomocí jazyku Go). [7,8]



users	
/api/block_user	oauth
/api/friend	oauth
/api/report_user	oauth
/api/setpermissions	oauth
/api/unfriend	oauth
/api/user_data_by_account_id	oauth
/api/username_available	oauth
/api/v1/me/friends/username	oauth
/api/v1/user/username/trophies	oauth
/user/username/about	oauth
/user/username/comments	oauth
/user/username/downvoted	oauth
/user/username/gilded	oauth
/user/username/hidden	oauth
/user/username/overview	oauth
/user/username/saved	oauth
/user/username/submitted	oauth
/user/username/upvoted	oauth
/user/username/where	oauth

Obrázek 3 - endpointy Reddit API týkající se uživatelů

Koncové body API jsou rozdělené do rámců (scopes). O tyto rámce musí externí aplikace, vyvinutá pro Reddit, žádat, aby jim je daný uživatel povolil. Ve zdroji [9] můžete najít jednotlivé rámce koncových bodů. Jako příklad je zde uveden rámec „Identity“, ten udělí aplikaci přístup k uživatelskému jménu a datu registrace. Další z důležitých rámců je „History“, ta udělí aplikaci přístup k historii uložených příspěvků, zveřejněných komentářů i příspěvků nebo třeba i skrytého obsahu. Reddit API má 29 rámců.

1.4 Jazyk C#

C# je moderní vysokoúrovňový programovací jazyk, který podporuje několik paradigmat včetně imperativního, deklarativního, funkcionálního, generického, objektově orientovaného a komponentově orientovaného programování. Byl vytvořen společností Microsoft v roce 2000, aby splnil rostoucí poptávku po webových aplikacích, které Visual Basic a C++ nedokázaly uspokojit. Je součástí vývojářského ekosystému od Microsoftu jménem .NET. Vychází z jazyků C a C++.

Jeho hlavní použití je vytváření webových stránek, nativních Windows aplikací a pro vývoj her, ve spolupráci s Unity Engine. Použit se dá i na jiné úkoly, například na vývoj mobilních aplikací pomocí technologie Xamarin.

C# podporuje silně a implicitně typované proměnné klíčovým slovem `var`. Samozřejmě je také možné explicitně deklarovat proměnné. C# také podporuje silně typované ukazatele pomocí klíčového slova `delegate`. C# má i sémantiku specificky pro události `publish-subscribe`, na které jsou využity právě delegáty. C# dost silně odpovídá `Common Language Infrastructure (CLI)`. CLI je otevřená specifikace a technický standard, původně vyvinutý Microsoftem a později standardizován ISO/IEC. Tento standard popisuje vykonatelný kód a prostředí za běhu u vysokoúrovňových jazyků. [10,11]

1.5 Webové frameworky

Vývoj webových stránek je časově náročná činnost, proto se k usnadnění vývoje webových stránek používají různé frameworky. Framework je předvytvořená sada komponentů pro vytváření aplikací či systémů. Těchto frameworků je hned několik. A jejich výběr často může značně ovlivnit rychlost i kvalitu vývoje projektu. Vliv frameworku je více zřejmý na straně vývojáře, nežli uživatele. V následujícím textu bude uvedeno, podle čeho si mezi nimi vybrat, a také budou zmíněni známější zástupci těchto frameworků.

Kritéria:

- **Požadavky projektu** – Nejprve je potřeba zvážit potřeby projektu, zda specifické požadavky projektu odpovídají lépe některé z technologií, nebo jestli projekt již má nějakou část vytvořenou a nezvolit takovou technologii, aby integrace s předchozím kódem byla co nejjednodušší. Někdy se také může stát, že projekt má předem určený požadavek, který říká, jaký framework máme použít.
- **Křivka učení** – Některé frameworky mohou mít značně strmější křivku učení než jiné frameworky. Je důležité vybrat takový framework, se kterým mají vývojáři zkušenost, nebo se ho jsou ochotni naučit. Práce s novým frameworkem může často způsobit zpoždění a zvýšení chybovosti během vývoje projektu.
- **Podpora komunity** – Při nesnázích je velkou pomocí podpora komunity. Na hromadu otázek, které může mít vývojář, je při dobré komunitě již na internetu odpovězeno. Velké množství více obecných řešení je již vyřešené a často i dostupné. To potom vede k dalšímu zrychlení a zefektivnění vývoje.

- **Výkon** – V závislosti na charakteru aplikace může být vhodné vybrat takovou knihovnu, která je co nejvýkonnější. U jiných aplikací může být tento faktor nepříliš důležitý.
- **Podpora vývojářů** – V neposlední řadě je důležité zvážit, zda vybraný framework bude dlouhodobě podporovaný. Výběr frameworku, který už je široce používaný, a je i nadále podporovaný, může poskytnout větší stabilitu a bezpečí, než některé novější frameworky.

Následují některé více známé webové frameworky: ASP.NET Core, React, Angular, Vue.js. Po zvážení zmíněných frameworků byl zvolen ASP.NET Core, jelikož s touto technologií má autor práce předchozí zkušenosti. Faktorem byl také plán autora se profesně zaměřit na .NET ekosystém.

1.5.1 ASP.NET Core

ASP.NET Core je open-source, multiplatformní framework pro vytváření aplikací, jako jsou webové aplikace, IoT aplikace nebo třeba mobilní backend. Byl navržen tak, aby poskytoval optimalizované vývojové prostředí pro aplikace nasaditelné buď na cloud nebo běžící lokálně. Aplikace vyvinuté pomocí *ASP.NET Core* fungují na různých operačních systémech včetně Windows, Mac a Linux. První verze byla vydána v roce 2016. Aktuální verze je 7.0.5. Verze 8.0 je momentálně dostupná, ale v preview.

Jako programovací jazyk lze použít C#, F# a Visual Basic. Z těchto jazyků je nejpoužívanější C#, kvůli své modernosti, objektově orientovanému přístupu, který umožňuje dobře strukturovat kód aplikace a je relativně snadný na naučení.

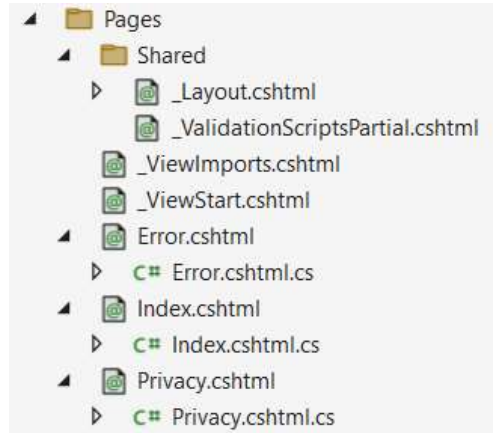
ASP.NET Core aplikace je v podstatě konzolová aplikace, která ve svém těle vytvoří webový server. Ten vytváří pomocí „WebApplicationBuilder“. Builder má metody a vlastnosti, pomocí kterých lze nastavit, jak se má server chovat.[3]

1.5.2 Razor Pages

Je webový vývojářský framework pro budování dynamických webových stránek pomocí *ASP.NET Core* frameworku. *Razor Pages* je stránkově orientovaným (page-centric) modelem. Jinými zástupci jsou *PHP*, *Classic ASP*, *Java Server Pages*, *ASP.NET Web Pages* a další.

Každá *Razor Pages* stránka ve složce „Pages“ představuje jeden endpoint. Každá stránka má přidružený soubor s C# modelem, který obsahuje funkcionalitu stránky.

Konvencí *Razor Pages* je pojmenování modelů shodně se stránkou. Jediným rozdílem je přípona. Místo „cshtml“ je přípona „cshtml.cs“. Zároveň všechny podpůrné soubory mají název, který začíná na podtržítka jako třeba „_Layout.cshtml“, což je soubor obsahující záhlaví a zápatí stránky.[4,13]



Obrázek 4 - složka Pages v novém projektu ASP.NET Core s Razor Pages

```
@page
@model IndexModel
@{
    ViewData["Title"] = "Home page";
}

<div class="text-center">
    <h1 class="display-4">Welcome @Model.Name</h1>
</div>

<form method="post" asp-page="Index">
    <div class="form-group">
        <label asp-for="Name"></label>
        <input asp-for="Name" class="form-control" placeholder="Enter Name" autocomplete="off" />
    </div>
    <button type="submit" class="btn btn-primary">Submit</button>
</form>
```

Zdrojový kód 1 - ukázkový Razor Page

Na zdrojovém kódu 1 lze vidět, jak může vypadat stránka v *Razor Pages*. `@page` označuje soubor jako razor stránku. `@model` definuje jaký soubor bude modelem této stránky, tedy již dříve zmíněným přidruženým souborem. `@{ }` v těchto složených závorkách jde psát čistý C# kód. V tagu `label` a `input` si také můžete všimnout atributu „asp-for“ nastaveného na „Name“, toto jméno souhlasí se jménem vlastnosti v přidruženém modelu. V praktičtějším příkladě potom můžeme použít bloku `@{ }` a pouze `@` pro dynamické generování stránek. Znak `@` znamená, že řádek obsahuje kód C#. Níže najdete příklad, jak může vypadat část razor stránky, dynamicky generující obsah.

```

<div class="col">
  <ul>
    <!--Downloads-->
    @foreach (Download download in Model.MyDownloads)
    {
      <div>
        Download @download.Id - @download.DownloadStart - @download.ProgressAbs / @download.ProgressAbsMax - Done
        <a asp-page-handler="Download" asp-route-id="@download.Id" class="btn btn-primary">Download</a>
        <a asp-page-handler="Delete" asp-route-id="@download.Id" class="btn btn-danger">X</a>
      </div>
    }
  </ul>
</div>

```

Zdrojový kód 2 - část razor page dynamicky generující obsah

1.6 Microsoft SQL Server

Microsoft SQL Server (MSSQL) je relační databázový systém, který podporuje širokou škálu zpracování transakcí, business logiky a dalších. Podobně jako i jeho konkurenti, je MSSQL vybudován nad variantou SQL, standardizovaným programovacím jazykem, který je využíván databázovými administrátory a jinými IT profesionály pro správu databáze a dotazování dat. MSSQL využívá jazyk Transact-SQL (T-SQL), což je implementace SQL od Microsoftu, která přidává řadu proprietárních programovacích rozšíření ke standardu jazyka.

Pokročilé bezpečnostní funkce podporované ve všech edicích Microsoft SQL Server od verze SQL Server 2016 SP1 zahrnují tři hlavní technologie: **Always Encrypted**, ta umožňuje uživateli aktualizovat šifrovaná data bez nutnosti je nejprve dešifrovat. **Row-level** zabezpečení, které umožňuje řízení přístupu k datům na úrovni řádků v tabulkách databáze. **Dynamické maskování dat**, které automaticky skrývá prvky citlivých dat před uživateli bez plných přístupových oprávnění.[14]

1.7 Teoretická analýza

Nakonec byla zvolena technologie *ASP.NET Core* s *Razor pages*. Tato možnost byla zvolena z několika důvodů, nejdůležitější z nich je autorova předchozí zkušenost s jazykem C# a celková přívětivost technologické platformy od společnosti Microsoft. Předchozí zkušenost s C# má autor primárně díky hobby vývoji videoher pomocí jazyku C# v kombinaci s *Unity Engine*. Toto rozhodnutí také plyne ze zajímavých moderních řešení co *ASP.NET Core* nabízí.

Jak bylo dříve zmíněno, cíl aplikace je poskytnout jednoduchou možnost, jak rychle a hromadně stáhnout uložené příspěvky ze sítě Reddit, proto byl navržen co nejjednodušší vzhled stránek i jejich malý počet.

Stránky aplikace:

- Home – základní informace o aplikaci, tipy o tom, jak aplikaci použít
- About – více detailní informace o stránce, autor, detail, kontakt
- Download – tato stránka bude technicky rozdělena na dvě stránky. Jedna pro zvolení, které příspěvky chce uživatel stáhnout, druhá pro zvolení pojmenování a strukturování souborů
- Progress – místo, kde uživatel bude moci vidět a stáhnout dokončené stahování. A zde také půjde vidět postup na ještě probíhajících stahováních
- Statistics – stránka, na které se budou zobrazovat statistická data aplikace

Jedním z logických kroků pro aplikaci tohoto stylu by bylo přidat obdobné funkce stahování i třeba pro subreddity. Nicméně po důkladném zvážení bylo rozhodnuto aplikaci zachovat co nejjednodušší pro snadné použití uživatelem. A to i proto, že ti uživatelé, kteří by se rozhodli stahovat obsah z celých subredditů, by mohli negativně ovlivnit výkonnost webové aplikace pro ostatní uživatele. Zároveň druh lidí, kteří chtějí stahovat tak velká data, jsou většinou takoví lidé, co jsou ochotni věnovat trochu více času nastavení tohoto stahovacího procesu. A jakmile je takový proces nastaven lokálně na počítači, tak je pochopitelně výkonnější než z webové stránky.

Aplikace v rámci stránky „Download“ nabídne filtrování, výběr příspěvků a poté také pojmenování a možnost nastavení struktury složek, ve kterých se příspěvky stáhnou. Pro tyto nastavení bude existovat systém, který dovolí uživatelům si nastavení uložit proto, aby nemuseli pokaždé vyplňovat podobné informace.

2 PRAKTICKÁ ČÁST

2.1 Praktická analýza

K architektuře aplikace bylo přistoupeno tak, že nejdůležitější funkcionality byly vytvořeny do vlastních služeb, které využívají zabudované Dependency Injection (DI je design model, ve kterém je možné, aby objekty dostaly svoje závislosti z externího zdroje, například z konstruktoru. Toto způsobuje snížení provázanosti kódu, zvýšení flexibility a udržitelnosti kódu). Zbytek funkcionality bude obsažen v rámci „cshtml.cs“ souborů, což jsou přidružené soubory s kódem C# u jednotlivých stránek. Důvodem k nezvolení komplikovanější architektury je relativní nenáročnost aplikace.

Funkcionality aplikace, které si zaslouží dedikovanou službu, jsou následující:

- `ApplicationDbContext` – služba pro ovládání databáze. Tato služba bude implementována pomocí balíčku `Microsoft.EntityFrameworkCore`. Práce s tímto balíčkem spočívá v deklaraci entit, inicializaci a spravování migrací.
- `RedditAPI` – služba pro získání dat z reddit API. Primárně pro získání všech uložených příspěvků a informací o profilu.
- `DownloadManager` – spravování stahování, což je jejich vytvoření, průběh i odstraňování.

Stránek bude ve výsledné aplikaci více než v teoretické analýze, což bylo očekáváno. Konkrétně jde o stránky `LoggedIn.shtml`, `Loading.cshmtl`. Jak bylo avizováno v teoretické analýze, tak stránka `download` bude rozdělena na `Download/Select` a `Download/Structure`. `Loading.cshtml` bude přidána pro zobrazení načítací stránky, jelikož po uvodním testování Reddit API bylo zjištěno, že získávání příspěvků z API je poněkud pomalé, přičemž při počtech příspěvků dosahujících k tisíci bude získávání trvat i několik sekund.

2.1.1 Použití API

Rozhodl jsem se nepoužívat žádný API wrapper a to i přesto že existuje wrapper „`RedditSharp`“ což je wrapper pro C#. Tento wrapper již není vývojářem udržován. Zároveň se domnívám, že využít originální řešení se pro charakter bakalářské práce hodí více.

Z Reddit API nakonec používám 4 endpointy (koncové body API). První endpoint slouží pro ověření reddit účtu. Druhý je pro získání přístupového tokenu. Třetí slouží pro získání

informací o uživatelově Reddit účtu. Poslední, čtvrtý dotaz, je pro získání uložených příspěvků od uživatele.

Aby bylo možné manipulovat s reddit API tak si vývojář musí založit Reddit aplikaci na svém Reddit účtu na webové stránce <https://www.reddit.com/prefs/apps>. Zde je možné si zvolit různé druhy aplikace, mezi které patří script, instalovaná aplikace a webová aplikace. Každá aplikace má svoje specifika a svoje podmínky použití.

Pro účely této aplikace je třeba založit webovou aplikaci. V nastavení aplikace si je třeba dát pozor na URI v položce „redirect uri“. Tato položka značí, kam bude dotaz přeměřován po přihlášení uživatele, a zároveň se musí schodovat s redirect URI, která je součástí dotazu.[15]



The screenshot shows the configuration interface for a Reddit application named "Online Downloader for - old". It includes a "secret" field with the value "N_yUDCCT3I_FrTbXVkf_Jgj8Y3_aLg", a "name" field with "Online Downloader for", and a "redirect uri" field with "https://localhost:44335/LoggedIn". There are also buttons for "update app" and "delete app".

Obrázek 5 - příklad webové Reddit aplikace

1) Endpoint - www.reddit.com/api/v1/authorize

Tento endpoint slouží k autentizaci uživatele. Při vstupu na stránku tohoto endpointu se uživatel ocitne na stránce redditu, kde je vyzván k přihlášení (pokud již tak neučinil). Po přihlášení má uživatel možnost udělit či zamítnout vyžádaná práva aplikaci.

Vyžádaná práva, stejně jako další informace, se vkládají jako GET požadavek za endpoint. Tedy v následujícím formátu:

„https://www.reddit.com/api/v1/authorize?client_id=xyz&response_type=code“

Mezi některé parametry patří následující:

- Client_id – id Reddit aplikace, zaregistrované přes Reddit.

- `Response_type` – druh odpovědi. Musí být „code“ pro Reddit aplikace typu „web“ a „script“, jelikož jsou považovány za důvěrné.
- `State` – textový řetězec dle vlastního uvážení. Lze využít pro silnější zabezpečení nebo/a zároveň může být použit pro vrácení zpět na určenou stránku jako argument.
- `Redirect_uri` – adresa, na kterou se uživatel vrátí. A to po přijetí či odmítnutí práv. Musí odpovídat `redirect uri`, které je specifikované při vytvoření Reddit aplikace.
- `Duration` – buď „temporary“ nebo „permanent“. Indikuje, jestli má být přístup časově omezený nebo neomezený. V obou případech vyprší access token během jedné hodiny, ale při zvolení „permanent“ zároveň obdržíte „refresh_token“ společně s access tokenem. Tento refresh token můžete použít pro získání nového access tokenu.
- `Scope` – skupiny pravidel, které vaše aplikace bude vyžadovat. Následují některé skupiny pravidel (scopes).
 - `Identity` – přístup k Reddit uživatelskému jménu a k datu registrace,
 - `History` – přístup k historii hlasování, komentářů, uložených a skrytých příspěvků,
 - `Read` – přístup k příspěvkům a komentářům skrz účet,
 - `Edit` – úpravy a mazání komentářů a příspěvků uživatele,
 - `Save` – ukládání a odstraňování z uložených příspěvků,
 - `Subscribe` – správa členství v subredditech, správa přátel. [9,15]

2) Endpoint - www.reddit.com/api/v1/access_token

Tento endpoint slouží k získání access tokenu uživatele. Na tento endpoint je potřeba poslat POST požadavek, proto jsou argumenty potřeba vložit do těla POST požadavku. Argumenty jsou následující: `grantType` je potřeba nastavit na „authorization_code“ a „code“ na kód, který jsme získali z endpointu „/api/v1/authorize“. A jako poslední je třeba nastavit „redirect_uri“ opět na stejnou stránku jako v Reddit aplikaci. [15]

3) Endpoint - oauth.reddit.com/api/v1/me

Endpoint pro získání informací o uživateli. Umožněn rámcem „Identity“. Volá se pomocí metody GET. Jako argument je potřeba přiložit „Authorization“ a do jeho hodnoty vložit „Bearer [access token]“. [7]

4) Endpoint - [oauth.reddit.com/user/\[username\]/saved](http://oauth.reddit.com/user/[username]/saved)

Endpoint získá uložené příspěvky uživatele. Umožněn rámcem „History“. Volá se pomocí metody GET. Jako argument je potřeba přiložit „Authorization“ a do jeho hodnoty vložit „Bearer [access token]“. Druhý argument je součástí endpoint adresy, do které musíte vložit uživatelské jméno (místo [username]).

Do těla URL je potřeba ještě vložit parametr „limit“ a volitelně také „after“. Limit je přirozené číslo, které určuje kolik příspěvků se má vrátit, přičemž maximální hodnota je 100. Pokud je potřeba získat více než 100 příspěvků, tak se musí poslat dotaz s argumentem „after“. Ten určuje, od jakého id příspěvku se budou příspěvky vracet. Pokud je tedy cílem získat všechny příspěvky, tak je třeba po prvním dotazu bez „after“, poslat dotaz s „after“ parametrem nastaveným na id posledního vráceného příspěvku z minulého dotazu. [7]

2.1.2 Vývojové prostředí

Jako IDE bylo použito Microsoft Visual Studio Community 2022 (64-bit) Version 17.4.1. Pro účely verzování bylo při vývoji využito GitHub. K ovládání githubu byl použit GitHub Desktop Version 3.2.1 (x64). Ke správě databáze byl použit Microsoft SQL Server Management Studio 19.0.1. Častému využití se také dostalo webové stránce <https://jsoneditoronline.org/>. Tato stránka byla používána pro nahlížení a kontrolu JSON souborů, převážně při integrování domén příspěvků.

2.2 Databáze

Pro uchovávání dat aplikace bylo využito SQL serveru verze 15.0.4153. Ke správě této databáze bylo použito Microsoft SQL Server Management Studio 19.0.1.

Výsledná práce využívá pět tabulek a jednu automaticky generovanou tabulku pomocí Microsoft Entity Framework (dále jen MEF). Tento framework je používán pro vytváření, modifikaci a migraci tabulek. MEF je balíčkem pro .NET, který poskytuje jednoduché nástroje pro vytváření a modifikaci tabulek. Respektive pomocí tohoto frameworku je možné vytvořit třídu v jazyku C# a následně deklarovat, že má být tabulka této entity v databázi. MEF se již potom postará o všechno sám (vytvoření tabulky, zaznamenání migrací, možnost callbacku). MEF také poskytuje anotace v C#, aby mohl vývojář specifikovat požadavky v kódu a nezasahovat manuálně do databáze.

Využívané tabulky jsou v práci následující (Název tabulky – uchovávaný druh dat – důležité informace v tabulce):

- „DownloadHistories“ – záznamy historie stahování – uživatel, seznam id příspěvků obsažených ve stahování,
- „Downloads“ – každý záznam představuje jedno stahování (běžící/dokončené – soubory stále na serveru) – uživatel, datum zahájení, progress, je dokončeno, je dostupné ke stáhnutí,
- „Statistics“ – záznamy statistiky ke každému uživateli a ke speciálnímu uživateli „SYSTEM“ představující statistiku celé aplikace. – uživatel, statistické data,
- „Templates“ – šablony – uživatel, nastavení formulářů,
- „Users“ – uživatelé aplikace – uživatelské jméno, Reddit id, přístupový token.

2.3 Stránky

V této kapitole jsou informace o jednotlivých stránkách aplikace. Také je zde zmíněn jejich obsah a úkol.

Všechny stránky jsou složeny ze sdíleného souboru „_Layout.cshtml“. Tato stránka obsahuje záhlaví a zápatí webové stránky. Importuje některé javascript skripty a také stylování CSS. Pro stylování stránky byl primárně použil Bootstrap a to konkrétně „bootswatch_cyborg.css“. V neposlední řadě soubor „_Layout.cshtml“ zobrazuje ostatní stránky v html tagu „main“.

Každá stránka aplikace je rozdělena na dva soubory. Soubor „cshtml“ a soubor „cshtml.cs“. Toto rozdělení je určené použitím technologie *Razor Pages*. V souboru „cshtml“ je html kód stránky společně s kódem C#, který je zde pro dynamické generování stránky. V souboru „cshtml.cs“ je potom funkční kód, jako u kontroleru v modelu MVC.

Header obsahuje odkazy na ostatní stránky aplikace a možnost se přihlásit/odhlásit.



Obrázek 6 - záhlaví stránek bez přihlášeného uživatele



Obrázek 7 - záhlaví stránek s přihlášeným uživatelem

2.3.1 Index

Index je domovskou stránkou aplikace. Tato stránka se lehce liší pokud uživatel je nebo není přihlášen. V obou případech jsou na stránce základní informace. Pokud uživatel není přihlášen, tak se zde nachází výzva k přihlášení. Pokud je uživatel přihlášen, tak se zde nachází navigace

na „Download“ a „Progress“. Zároveň při zobrazení této stránky se pošle dotaz GET na stránku „Download/Loading“. Tento dotaz způsobí, že začne proces získávání příspěvků z Redditu.

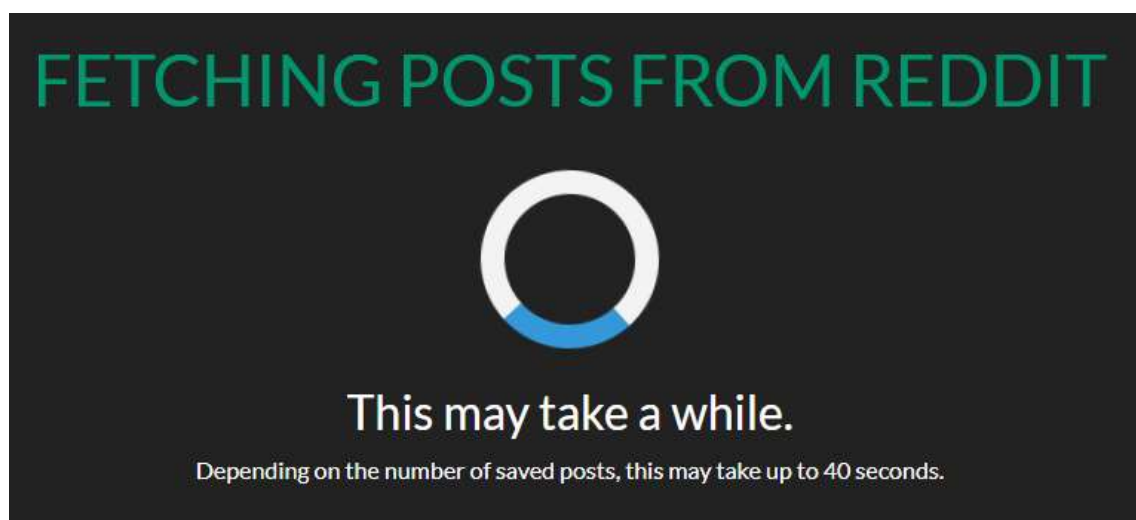


Obrázek 8 - domovská stránka aplikace s přihlášeným uživatelem

2.3.2 LoggedIn

LoggedIn je pomocná stránka, která se uživateli ani nezobrazí, jestliže proběhne přihlášení v pořádku. Stránka zpracovává odpověď z reddit přihlášení. Vytvoří nebo aktualizuje uživatele v databázi. Získá a uloží uživatelovo id a uživatelské jméno. Pokud je přihlášení úspěšné, uživatele přesměruje na domovskou stránku aplikace. Pokud přihlášení selže, tak se zobrazí chybová hláška.

2.3.3 Download/Loading



Obrázek 9 - stránka Download/Loading

Tato stránka slouží jako portál na stránku „Download/Select“. Na stránce se nachází načítací obrazovka. Tedy výzva k čekání a načítací kolečko. Na pozadí se při vstupu na stránku zkontroluje zda jsou příspěvky z Redditu již získané. Pokud ano, tak hned přesměruje

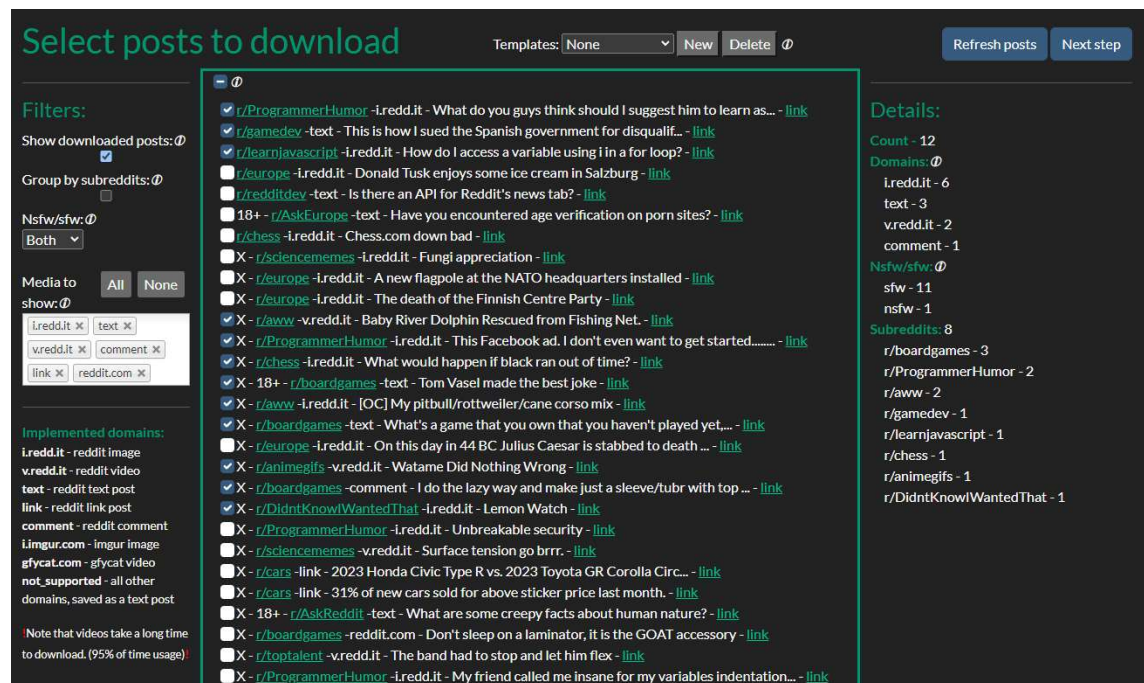
na „Download/Select“. Pokud příspěvky získané nejsou, tak se stránka vykreslí. V javascriptu se zavolá metoda z této stránky, která zkontroluje, jestli pro tento session se již příspěvky získávají. Pokud ne, tak je začne získávat. Pokud se pro tuto relaci již příspěvky získávají, začne čekat na dokončení získávání příspěvků. Nakonec je uživatel přesměrován na stránku „Download/Select“

2.3.4 Download/Select

Tato stránka slouží pro vybrání uložených příspěvků, které chce uživatel uložit. Na levé straně je formulář na filtrování příspěvků. V centru jsou zobrazeny příspěvky. Nad nimi je správa šablon. Na pravé straně jsou detaily o vybraných příspěvcích.

V tento moment jsou zobrazovány proměně entit „Post“, které jsou uloženy v relaci. Samotné stažení medií se začne provádět až ze stránky „Download/Structure“.

Na straně serveru se při vyžádání stránky načtou potřebné informace z databáze a z relace. Filtrování probíhá na straně uživatele. Více o filtrování lze najít v kapitole 2.7. Část javascript kódu dokáže pomocí AJAX volat metody u této stránky. Konkrétně se jedná o vytvoření a odstranění šablon. Zbytek logiky stránky je kód javascript, který se stará o filtrování, vypisování, počítání detailů.



Obrázek 10 - stránka Download/Select

2.3.5 Download/Structure

Tato stránka slouží pro nastavení podoby stažených souborů. Jejich pojmenování a rozložení do složek. Rozložení stránky se podobá „Download/Select“. Ve středu opět máme zobrazené příspěvky, teď ovšem zobrazené tak, jak budou následně staženy. Vlevo je formulář s nastavením stahování. Nahoře je opět správa šablon. A vpravo jsou detaily stejné jako na minulé stránce.

Na straně serveru se spravují šablony, získávají data z databáze a relace, obstarává se tlačítko na zahájení stahování. Stahování se zahájí jenom v případě, pokud uživatel nemá více než tři právě se stahující procesy.

Struktura příspěvků se generuje v rámci javascript, tato struktura se neposílá do C#. V C# se vytváří struktura znova. Tento postup je sice více duplicitní, ovšem používání a přesunování struktury z Javascriptu do C# by bylo velmi náročné, nepřiliš efektivní a nebezpečné.

The screenshot shows a web interface titled "Naming/Foldering". At the top, there are "Templates: None" and buttons for "New" and "Delete". On the right, there are buttons for "Previous step" and "Start download process".

Names:

- Numbering: \emptyset
- Ids
- Standard
- None

Name contents: \emptyset

y/n Priority

Subreddit:

Domain:

Title: \emptyset

Folders:

Group by: \emptyset

y/n Priority

Subreddits:

Domains:

Remove empty folders: \emptyset

Split by nsfw/swfw: \emptyset

Details:

- Count:** 38
- Domains:** \emptyset
- i.redd.it - 20
- text - 8
- v.redd.it - 6
- link - 2
- comment - 1
- reddit.com - 1
- Nsfw/swfw:** \emptyset
- swfw - 35
- nsfw - 3
- Subreddits:** 22
- r/ProgrammerHumor - 4
- r/europe - 4
- r/aww - 4
- r/boardgames - 4
- r/chess - 3
- r/sciencememes - 2
- r/cars - 2
- r/gamedev - 1
- r/learnjavascript - 1
- r/redditdev - 1
- r/AskEurope - 1
- r/animegifs - 1
- r/DidntKnowIWantedThat - 1

Main Content:

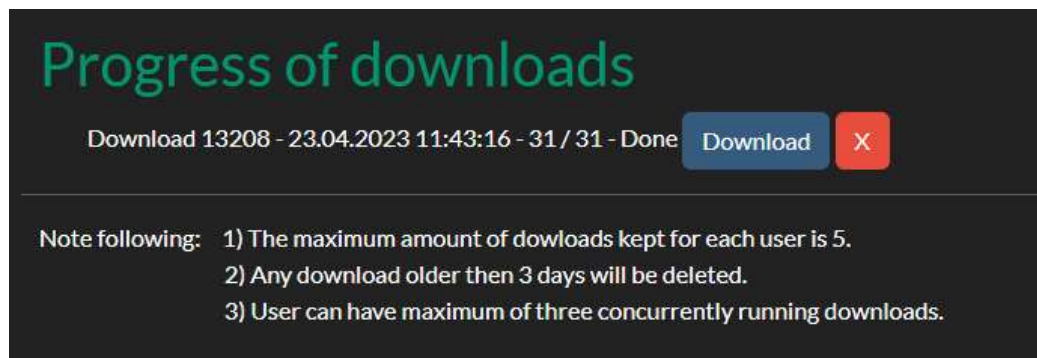
- r/ProgrammerHumor**
 - i.redd.it_What do you guys think should I suggest him to learn as thei
 - i.redd.it_This Facebook ad. I don't even want to get started.... *Sig
 - i.redd.it_Unbreakable security
 - i.redd.it_My friend called me insane for my variables indentation...
- r/europe**
 - i.redd.it_Donald Tusk enjoys some ice cream in Salzburg
 - i.redd.it_A new flagpole at the NATO headquarters installed
 - i.redd.it_The death of the Finnish Centre Party
 - i.redd.it_On this day in 44 BC Julius Caesar is stabbed to death by 30
- r/aww**
 - v.redd.it_Baby River Dolphin Rescued from Fishing Net.
 - i.redd.it_[OC] My pitbull/rottweiler/cane corso mix
 - v.redd.it_Yesterday our indoor cat escaped for the first time. 12 hour
 - i.redd.it_[OC] My little Subwoofer is 12 this week!
- r/boardgames**
 - text_Tom Vasek made the best joke
 - text_What's a game that you own that you haven't played yet, beca
 - comment_I do the lazy way and make just a sleeve/tubr with top and b
 - reddit.com_Don't sleep on a laminator, it is the GOAT accessory
- r/chess**
 - i.redd.it_Chess.com down bad
 - i.redd.it_What would happen if black ran out of time?
 - i.redd.it_Chess.com CEO reveals that their investigation into Hans is
- r/sciencememes**
 - i.redd.it_Fungi appreciation
 - v.redd.it_Surface tension go brrr.
- r/cars**

Obrázek 11 - stránka Download/Structure

2.3.6 Download/Progress

Tato stránka slouží pro zobrazení a správu stavu stahovacích procesů a to včetně možnosti stáhnout dokončené procesy. Ty se zde zobrazují v listu. Pod tabulkou s procesy se nachází nápověda o omezení počtu stahování.

Logika stránky je velice jednoduchá, jelikož všechna potřebná data jsou již připravená z ostatních stránek, tudíž se zde jenom načítají a zobrazují.



Obrázek 12 - stránka Download/Progress

2.3.7 Statistics

Jedná se o stránku, na které se zobrazuje statistika stahování a to jak pro daného uživatele, tak i pro celou aplikaci. Na levé straně stránky jsou data pro celou aplikaci a na pravé pro uživatele.

Na straně serveru se získávají data z databáze a některé části statistiky se počítají přímo při dotazu na stránku. Tato stránka má rozdílné zobrazení pro uživatele přihlášené a nepřihlášené. Pokud uživatel není přihlášen, tak se zobrazí statistika pouze o aplikaci.

Zobrazované údaje jsou

- Společné (pro aplikaci i uživatele) – celkový počet stažení, celkový počet stažených příspěvků, průměrný počet příspěvků na jedno stahování, rozdělení domén,
- Unikátní pro aplikaci – počet registrovaných uživatelů, počet aktivních stažení (běžící i dokončené),
- Unikátní pro uživatele – rozdělení subredditů.

2.4 Flow aplikace

Zamýšlený popstup uživatele při používání aplikace je následující: při prvním navštívení aplikace se uživatel ocitne na stránce „Home“, zde se může dočíst jak má zacházet s webovou stránkou. Dále uživatel pokračuje kliknutím na „Login with Reddit“, což ho přesměruje na stránku Redditu s přihlášením a s potvrzením práv aplikace. Následně je přesměrován zpět na „Home“ stránku (ve skutečnosti je uživatel nejdříve přesměrován na stránku „LoggedIn“, na té se ale nezastaví, pokud přihlášení bylo úspěšné). Odtud se uživatel může rozhodnout jít buď do „Download“ a zahájit nový process nebo do „Progress“, aby dokončený process

stáhnul. Pokud je zde uživatel poprvé, tak zamíří do „Download“, tedy na stránku „Download/Select“ přes stránku „Download/Loading“, kde vybere, které příspěvky chce stáhnout. Pomocí tlačítka „Next step“ se posune na další stránku a to na stránku „Download/Structure“, kde uživatel určí jak mají být soubory pojmenované a do jakých složek mají být rozdělené. Při stisku tlačítka „Start Download“ je přesměrován zpět na domovskou stránku se zprávou, která obsahuje jméno nového procesu. Odtud může uživatel znova jít buď do „Download“ nebo do „Progress“. Aby uživatel stáhnul soubory na svoje zařízení, zamíří do „Progress“, kde po dokončení stahování může stáhnout archiv se stáhnutými příspěvky.

Při opakovaném navštěvování může uživatel také využít šablony na vrchní straně stránek „Download/Select“ a „Download/Structure“. Toto uživateli umožní načítat opakovaně používané nastavení a tak zkrátí dobu, kterou musí strávit s nastavováním aplikace.

V přepisu (za předpokladu nového uživatele):

Home > Login with Reddit > Home > Download (Select) > Download (Structure) > Home > Progress

Při prvním pohledu na flow aplikace se může zdát, že je zbytečné se tolikrát vracet na hlavní stránku, bohužel je toto nevyhnutelné, jelikož každé vrácení na hlavní obrazovku představuje nejistotu s uživatelským záměrem. Bez známého záměru jsem se rozhodl přesměrovat uživatele na „Home“ jako domovskou stránku, aby se sám rozhodl co chce dělat. Například pokud by uživatel chtěl zahájit ještě jedno stahování před samotným stažením souborů, je nevhodné uživatele posílat rovnou na stránku „Progress“.

2.5 Přihlášení

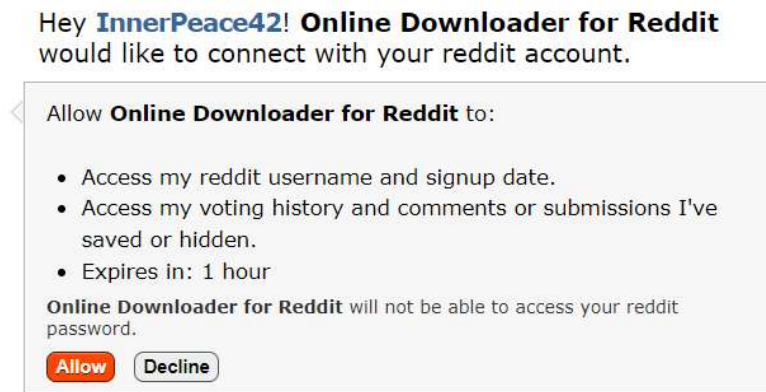
Přihlášení do aplikace probíhá přes Reddit. Tudíž aplikace se nikdy nedostane k heslu a proto se také hesla v aplikaci neuchovávají. Při prvním přihlášení do aplikace se vytvoří záznam v databázi pro daného uživatele (více o podobě databáze bude níže). Aby se uživatel přihlásil klikne v aplikaci na následující odkaz:

```
<a href="https://www.reddit.com/api/v1/authorize?client_id=9RevD-RRIRmNcGc3nsu-  
pg&response_type=code&state=test1&redirect_uri=https://localhost:44335/LoggedIn&duration=temporary&scope=identity,read,history">Login with Reddit</a>
```

Tento odkaz přesměruje uživatele na stránky Redditu s určenými argumenty. Aplikace má následující argumenty: [15]

- Client_id – id Reddit aplikace, která byla zaregistrována na Redditu,
- Response_type – „code“ nutná složka pro aplikaci typu „web“,
- State – tento parametr není využíván, proto může být libovolným řetězcem,
- Redirect_uri – uri LoggedIn stránky, která zpracuje výsledek z pokusu o přihlášení,
- Duration – pro zvýšenou bezpečnost aplikace bylo zvoleno „temporary“, tedy aplikace získá pouze 1 hodinu platný access token a o prodloužení žádat nemůže,
- Scope – pro účely aplikace budou stačit práva: identity a history.

Na této obrazovce se uživatel přihlásí do Redditu (pokud již není přihlášen) a jakmile je přihlášen, zobrazí se mu výzva, zda chce aplikaci „Online Downloader for Reddit“ udělit práva. Uživatel má teď možnost zvolit, zda práva udělí nebo odmítne, jakmile se toto stane, Reddit přesměruje uživatele na „LoggedIn“ stránku.



Obrázek 13 – Reddit formulář na schválení práv

Stránka „LoggedIn“ zpracuje výsledek z pokusu o přihlášení. Logika programu vypadá zhruba takto: nejdřív se kód získá z odpovědi na požadavek, tento kód se použije jako ověření při požádání o access token. Následuje tedy volání na endpoint „https://www.reddit.com/api/v1/access_token“. Z odpovědi na tento požadavek se získá access token. Následně se tento token použije pro získání informací o uživateli Reddit účtu. Získané informace se uloží do relace a zkontroluje se zdali uživatel již aplikaci využil. Na základě tohoto faktu se informace o uživateli buď aktualizují nebo se vytvoří zcela nový uživatel. Obě tyto změny se ukládají do databáze do tabulky Users. Pokud jsou všechny operace

úspěšné, tak je uživatel přesměrován na „Home“ stránku, pokud se něco nepodaří, tak se vypíše chybová hláška.

	Id	UserName	AccessToken	Registered	LastLogin	RedditId
1	6	InnerPeace42	2266675398600-6XIOtr...	2023-04-16 14:56:51.4207908	2023-04-23 11:36:55.8956576	sxao0b0o
2	7	Jose946	162712379165-ntJfVAN...	2023-04-16 15:59:06.2295626	2023-04-23 11:31:37.9563554	██████████

Obrázek 14 – příklad dat v tabulce Users

2.6 Získání příspěvků

Příspěvky ze stránky Reddit se začínají získávat hned po přihlášení uživatele. Přesněji hned po vykreslení domovské stránky se zavolá pomocí technologie AJAX metoda ze stránky „Download/Loading“ a ta na základě pár podmínek může zavolat metodu ve službě „RedditAPI“ jménem „GetAllSavedPosts“ se dvěma parametry. Pokud se uživatel pokusí vstoupit na stránku „Download“ a příspěvky ještě nejsou načtené, tak se uživateli zobrazí načítací obrazovka z „Download/Loading“ a bude na stránce čekat dokud se příspěvky nenačtou. Pokud jsou příspěvky již získané v okamžiku, kdy se uživatel pokusí vstoupit na stránku „Download“, je uživatel ihned přesměrován na „Download/Select“.

Na stránce „Download/Select“ je také možnost obnovit příspěvky přes tlačítko „Refresh posts“. Toto tlačítko vymaže aktuální získané prvky z relace a pošle uživatele na stránku „Download/Loading“, kde se uživateli zobrazí načítací stránka a příspěvky se začnou načítat.

2.6.1 Metoda GetAllSavedPosts

V této kapitole je detailně popsán praktický postup při získávání příspěvků ze sociální sítě Reddit za pomoci Reddit API.

Jak bylo zmíněno, metoda má dva parametry, a to access token, potřebný pro všechny dotazy směrem na API a uživatelské jméno z Redditu, což je specificky potřebné pro dotaz na získání uložených příspěvků uživatele.

Po inicializaci potřebných proměných následuje cyklus, ve kterém se získávají uložené posty uživatele. V tomto cyklu se musí rozlišit první průběh od každého následujícího, jelikož při prvním průchodu se nesmí uvádět parametr „after“. Tento parametr je potřeba až pokud je prvních sto příspěvků získáno. Parametr „after“ je nastaven vždy na id posledního vráceného příspěvku. Zároveň ke všem těmto dotazům je třeba přidat parametr „limit“, který je nastavován na maximální hodnotu, což je hodnota 100. Každá odpověď s příspěvků je kompilována do jednoho JSON pole. Cyklus končí ve chvíli, když se z dotazu nevrátí sto příspěvků, protože

program narazil na konec, nebo ve chvíli, když se vrátí desátý dotaz, jelikož se ukládá maximálně tisíc příspěvků.

```
//Send requests until returned number of posts are not 100 or until all posts were returned. And make all these pos
while (returnedNumber == 100 && count < 10)
{
    if (after != "")
    {
        request = new HttpRequestMessage();
        request.RequestUri = new Uri($"https://oauth.reddit.com/user/{userName}/saved?limit=100&after={after}");
        request.Method = HttpMethod.Get;
        request.Headers.Add("Authorization", $"Bearer {token}");
    }
    else {
        request = new HttpRequestMessage();
        request.RequestUri = new Uri($"https://oauth.reddit.com/user/{userName}/saved?limit=100");
        request.Method = HttpMethod.Get;
        request.Headers.Add("Authorization", $"Bearer {token}");
    }
    response = await client.SendAsync(request);
    contents = await response.Content.ReadAsStringAsync();
    jsonData = JObject.Parse(contents);
    list = "" + jsonData.data.children;
    list = list.Substring(1, list.Length - 2);
    if (after != "")
    {
        sb.Append(", " + list);
    }
    else
    {
        sb.Append(list);
    }
    returnedNumber = jsonData.data.dist;
    after = jsonData.data.after;
    length = length + returnedNumber;
    count++;
}
sb.Append("[, \"NumberOfPosts\": " + length + "];");
```

Zdrojový kód 3 - část metody *GetAllSavedPosts* odpovědná za získání příspěvků z uživatelského účtu

Následuje druhý cyklus, který se provede pro každý prvek v JSON array. Pro každý příspěvek se kontrolují specifické informace, které určují, jak se daný prvek převede do entity „Post“. Toto je potřeba například proto, že různé druhy příspěvků mají URL na svoje médium (pokud nějaké médium vůbec mají) na různých místech. Také se může jednat o uložený komentář. Všechny tyto a další eventualy se zde kontrolují a řeší tak, aby výsledkem byl list entit „Post“.

Zde následují všechny možné operace, které se provádí příspěvku získanému z Redditu:

- Pokud je příspěvek odstraněn (moderace, smazání autorem), pak se tento post nepřevede na entitu „Post“.
- Pokud se jedná o komentář, tak následuje specifické převedení na entitu „Post“ s doménou nastavenou na „comment“.
- Pokud je doména „reddit.com“, tak se uloží všechny url medií ze specifického místa.
- Pokud je doména „v.redd.it“, tak se zkontroluje, zdali se jedná o crosspost (příspěvek, který je sdílený z jednoho subredditu do jiného). Pokud se o crosspost jedná, tak se URL získá z originálního příspěvku. Pokud příspěvek crosspost není, tak se url získá ze specifického místa příspěvku.

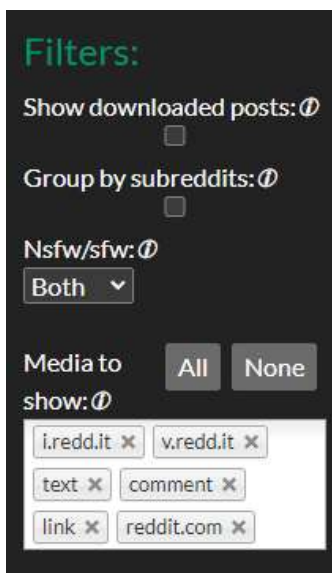
- Pokud je doména „gfyca.com“, tak se médium získá ze specifické cesty.
- Pokud doména začíná na „self“, tak se doména nastaví na „text“ (pro rozumné zacházení s příspěvky během filtrování a stahování, jinak by každý subreddit měl vlastní „self.[jméno subredditu]“ doménu).
- Pokud post_hint je „link“, tak se nastaví doména na „link“. Opět pro rozumné zacházení při filtrování a stahování (každý link má doménu stránky, ze které link pochází).
- Pokud je doména jiná než jedna z podporovaných domén, tak se nastaví na „not_supported“.

Po těchto operacích se již příspěvek převede na entitu „Post“ a přidá do listu. Toto se provede pro každý prvek. Jakmile je tento cyklus dokončen, tak se list „Postů“ vrací odkud byl zavolán.

2.7 Filtrování a zobrazení příspěvků

Jelikož uživatel potřebuje pravidelně rychlou odezvu na změnu filtrů v aplikaci, tak bylo filtrování a zobrazení příspěvků implementováno v javascriptu na straně klienta. Toto zobrazování a filtrování probíhá na stránkách „Download/Select“ a „Download/Structure“.

2.7.1 Filtrování na stránce „Download/Select“



Obrázek 15 - formulář na filtrování příspěvků

Jednotlivé položky formuláře mají následující funkce:

- „Show downloaded posts“ – zobrazí již stažené příspěvky,
- „Group by subreddits“ – ukáže příspěvky seskupené podle subredditu. (přidává možnost zaškrtnout celé subreddity),
- „Nsfw/sfw“ – rozhoduje, které příspěvky se budou ukazovat na základě věkové přístupnosti (Both – všechny příspěvky, Nsfw – příspěvky 18+, sfw – dostupné pro mladší 18-ti let),
- „Media to show“ – vybrané domény se budou zobrazovat.

Filtrování na straně klienta probíhá ve funkci nazvané „redraw“, ta se spustí při každé změně ve formuláři. Tato funkce zaznamená všechny vybrané příspěvky a potom vyfiltruje příspěvky viz zdrojový kód 4. Následně se dostává na generování html pro seznam příspěvků, přičemž pokud se vyfiltrují všechny příspěvky, tak se místo seznamu zobrazí upozornění o tom, že uživatel pravděpodobně má příliš přísné filtry. Pokud nějaké příspěvky zůstaly po filtrování, tak se na základě prvku „Group by subreddits“ rozhodne, zda se příspěvky mají zobrazit v seznamu nebo seskupené podle subredditů.

```
//Filtering posts
let filteredPosts = posts;
if(!v_showDownloaded == true){
  filteredPosts = filteredPosts.filter(post => !downloadedIds.includes(post.Id));
}
if(v_nsfw == 'nsfw'){
  filteredPosts = filteredPosts.filter(post => post.Over18);
}else if(v_nsfw == 'sfw'){
  filteredPosts = filteredPosts.filter(post => !post.Over18);
}
if(v_multipleSelect == null){
  filteredPosts = [];
}else{
  filteredPosts = filteredPosts.filter(post => v_multipleSelect.includes(post.Domain));
}
filteredLength = filteredPosts.length;
```

Zdrojový kód 4 - filtrování příspěvků v Javascriptu

2.7.2 Zobrazení na stránce „Download/Structure“

Names:

Numbering: \emptyset

Ids
 Standard
 None

Name contents: \emptyset

	y/n	Priority
Subreddit:	<input type="checkbox"/>	<input checked="" type="radio"/>
Domain:	<input type="checkbox"/>	<input type="radio"/>

Title: \emptyset

Folders:

Group by: \emptyset

	y/n	Priority
Subreddits:	<input checked="" type="checkbox"/>	<input checked="" type="radio"/>
Domains:	<input type="checkbox"/>	<input type="radio"/>

Remove empty folders: \emptyset

Split by nsfw/sfw: \emptyset

Obrázek 16 - formulář na pojmenování a strukturu složek

Jednotlivé položky v formuláři mají následující funkce:

- „Numbering“ – styl číslování příspěvků. (Ids – id příspěvku z Redditu, Standard – numbering in each folder 1,2,3, None – no numbering)
- „Name contents“ – co má být obsahem názvu před nadpisem z příspěvku
 - „y/n“ pomocí tohoto zaškrtačacího políčka dáte stránce vědět, že tento prvek (název subredditu, název domény) má být součástí názvu
 - „Priority“ – zaškrtnuté pole bude v názvu uvedeno jako první
- „Title“ – číslo v tomto poli určuje, kolik znaků z nadpisu příspěvku se bude vyskytovat ve jméně souboru.
- „Group by“ – podle čeho se budou příspěvky seskupovat do složek. Tento prvek technicky funguje stejně jako „Name contents“. Přičemž priorita zde určuje podle kterého prvku (domény/subredditu) se bude rozdělovat do složek jako první.
- „Remove empty folders“ – odstraní všechny složky, které by zůstaly prázdné (například při rozdělení podle nsfw a zároveň rozdělení podle subredditů)
- „Split by nsfw/sfw“ – pokud je pole zaškrtnuté, tak vzniknou složky sfw a nsfw, do nichž budou příspěvky rozděleny podle proměné „Over18“ z entity „Post“.

Výpis příspěvků opět probíhá v javascriptu ve funkci „redraw“. Tato funkce používá řadu dalších funkcí, přičemž jsou všechny zaměřené na vytváření html pro seznam příspěvků podle

zadaných parametrů z formuláře. Funkce „redraw“ je spuštěna při každé změně formuláře. Proces vytváření této struktury slouží jenom pro zobrazení podoby stažených souborů. Při samotném stažení se používají proměnné, které byly zadané na této stránce – „Download/Structure“ a tak se struktura generuje znovu.

```
//function that redraws the posts
function redraw(){
  masterIndex = 0;
  postList.innerHTML = "";
  offset = 0;
  if(v_split == true){
    postList.appendChild(makediv1("sfw"));
    offset++;
    let remove = drawPart('sfw');
    offset--;
    if(v_empty && remove){
      postList.removeChild(postList.lastElementChild);
    }
    postList.appendChild(makediv1("nsfw"));
    offset++;
    remove = drawPart('nsfw');
    offset--;
    if(v_empty && remove){
      postList.removeChild(postList.lastElementChild);
    }
  }else{
    drawPart('none');
  }
}
```

Zdrojový kód 5 - funkce redraw ze souboru „Structure.cshtml“

Na zdrojovém kódu 5 vidíte funkci redraw ze souboru „Structure.cshtml“. Proměnná offset značí, kolik tabulátorů se má použít při vypsání daného příspěvku. Při této funkci se zjistí, zda se mají příspěvky rozdělovat na NSFW/SFW nebo ne. Následně se zavolá funkce „drawPart“ s argumentem, který značí, jakou část příspěvků je třeba vypsát. Zároveň se při rozdělení kontroluje jestli nějaká ze složek není prázdná, pokud je, tak se odstraní.

V rámci funkce „drawPart“ se zjistí, jak se mají příspěvky rozdělit do složek. Následně se začnou vypisovat příspěvky pomocí dalších pomocných funkcí. Při samotném vložení příspěvku do html kódu se volá funkce „createName“, ta vytváří jméno příspěvku podle vrchních tří prvků ve formuláři na stránce „Download/Structure“, včetně kontrolování duplicity ve složce.

2.8 Stahování příspěvků

Principiálně stahování probíhá následovně: podle parametrů ve formuláři na stránce „Structure“ se stahují příspěvky lokálně na server. Takže při stažení jsou už správně pojmenované a rozřazené do složek. Poté co jsou všechny příspěvky úspěšně staženy, tak se složka přidá do archivu zip. Původní složka se smaže a archiv je připraven ke stažení uživatelem.

V rámci následujících odstavců se používá slovo stahování ve dvou možných významech a to stahování jako process serveru a stahování související s uživatelským ukládáním stažených souborů.

Stahování spravuje služba „DownloadManager“. Hlavní funkce pro stahování je zavolána z metody pro obsluhu tlačítka „Start Download“ na stránce „Structure“. Tato metoda s názvem „NewDownloadProcessAsync“ vyžaduje hned několik argumentů: uživatele, který stahování zahájil, list příspěvků ke stažení, list id všech získaných příspěvků z Redditu a objekt „downloadParameters“, který je vytvořen z informací formuláře na stránce „Structure“. Tento objekt obsahuje informace jako je pojmenování souborů i styl struktury složek.

Metoda „NewDownloadProcessAsync“ provádí následující:

- Vytvoří nový záznam v tabulce „Downloads“. Tento záznam představuje proces stahování na serveru. V C# programu se tento záznam jmenuje „DownloadProcess“. Záznam obsahuje nejen datum zahájení a ukončení, ale i uživatele, který stahování zahájil, progress ve stahování a zda je stahování dokončené a přístupné ke stáhnutí.
- Vypočítá interval. Tento interval značí, kolik příspěvků se musí stáhnout před tím, než bude aktualizovaný progress vložený do databáze a tím i aktualizovaný na stránce „Download/Progress“.
- Vytvoří nové vlákno, které bude vykonávat práci stahování.
- Vytvoří a přidá objekt „DownloadHistory“ do tabulky „DownloadHistories“ v databázi.
- Odstaruje vytvořené vlákno.

Odstartované vlákno obsahuje volání metody „DoWork“, která vykoná následující kroky. Vytvoří složku označenou s id download procesu. Připraví si předem list se všemi unikátními doménami a subreddity. Podle parametru „Split“ z „DownloadParameters“ se program dozví, zda má rozdělit příspěvky podle přístupnosti (podle obsahu nad 18). Pokud se má obsah rozdělit, tak se postupně vytvoří dvě složky, vyfiltrují se požadované příspěvky a zavolá se metoda „SavePart“ pro obě části. Pokud se obsah nemá rozdělit, tak stačí zavolat metodu „SavePart“. Jakmile se dokončí metoda „SavePart“, tak jsou příspěvky již stažené. Nicméně proces stahování ještě hotový není. Následuje vytvoření archivu ze stažené složky. Potom se použitá složka smaže. Do záznamu v tabulce „downloads“ se zapíše, že už je toto stažení připravené ke skutečnému stáhnutí. Pokud je uživatel autorem více než pěti stažení, tak se odstraní jeho nejstarší stažení. Následně se odstraní nepotřebné záznamy z tabulky „downloadHistories“. Tímto je stažení kompletní.

Metoda „SavePart“ má následující argumenty:

- „DownloadProcess“ což je třída ze služby „DownloadManager“, která odpovídá jednomu stahovacímu procesu,
- List filtrovaných příspěvků,
- Aktuální cesta, na kterou se mají ukládat soubory,
- Seznam všech unikátních domén,
- Seznam všech unikátních subredditů.

Zde se vykonání programu dostane do jedné z pěti větví. Rozdělují se podle parametrů „DomainFolder“ a „SubredditFolder“, které představují požadavek dělit podle domén a podle subredditů. Obě možnosti mohou být vypnuté, jedna z nich může být zapnutá nebo můžou být obě možnosti zapnuté, pokud jsou obě možnosti zapnuté, tak ještě rozhoduje parametr „FolderPriorityIsSubreddit“. Na základě požadavků se vytvoří složky (pokud jsou třeba) a následně se zavolá „DownloadPosts“ pro filtrované listy příspěvků.

Metoda „DownloadPosts“ slouží pro stažení skupiny filtrovaných příspěvků. Sama příspěvky nestahuje, jelikož na to jsou určeny jiné pomocné metody. Tato metoda má na starost postupně volat metodu „SavePost“. Jedním z argumentů, které metodě poskytne je path. Ten je vytvořen z předchozí cesty a výsledku metody „generateName“, která jméno souboru vytvoří podle všech parametrů uživatele z formuláře na stránce „Download/Structure“. Zároveň se v tomto cyklu kontroluje, zda není vyžádáno zastavení vlákna. Pokud je, tak pomocí příkazu „return“ vrátí ovládání postupně zpět až do „DoWork“ metody, která se tím zastaví. Zároveň metoda „DownloadPosts“ kontroluje, zda se dosáhlo intervalu pro aktualizování databáze a pokud se tak stane, tak záznam v databázi aktualizuje.

Metoda „SavePost“ slouží jenom jako „rozcestník“ pro spouštění kontrétních stahovacích metod jako jsou: SaveImage, SaveMultipleImages, SaveVideo, SaveLinkPost, SaveComment, SaveTextPost a SaveNotSupported. Metoda se rozhoduje na základě domény příspěvku.

Jednotlivé metody jsou zodpovědné za stahování specifických domén a tím i medií.

```
private async Task SaveImage(Post post, string path, int id) {  
    Stream stream = await client.GetStreamAsync(post.Urls[0]);  
    using (var fileStream = File.Create(path + Path.GetExtension(post.Urls[0])))  
    {  
        stream.CopyTo(fileStream);  
    }  
}
```

Zdrojový kód 6 - metoda pro ukládání domén i.redd.it a i.imgur.com


```

private void SaveTextPost(Post post, string path, int id)
{
    using (StreamWriter sw = File.CreateText(path + ".txt"))
    {
        sw.WriteLine("Title: " + post.Title);
        sw.WriteLine("Subreddit: " + post.Subreddit);
        sw.WriteLine("PermaLink: www.reddit.com" + post.PermaLink);
        sw.WriteLine("Text: \n" + post.SelfText);
    }
}

```

Zdrojový kód 7 - metoda pro ukládání textových příspěvků

Samotné metody pro stahování jsou krátké, jelikož se naprostá většina logiky odehrála již před těmito metodami. Například argument „path“ už obsahuje správné pojmenování souboru i správnou cestu, na kterou je potřeba soubor uložit.

Každé stahování má svůj záznam v tabulce „Downloads“ v databázi. Každý takový záznam má svoje soubory na serveru a to buď již stažené nebo ještě se stahující. Proto je třeba uživateli omezit počet stahování, aby nemohl zahltit server. Z tohoto důvodu byla zavedena následující pravidla:

- 1) Uživatel může mít pouze tři najednou se stahující procesy,
- 2) Uživatel může mít pouze pět stáhnutých procesů,
- 3) Každý proces starší třech dnů bude smazán.

2.8.1 Historie stahování

Jelikož primární záměr aplikace je stahování uložených příspěvků, přičemž uživatelé pravděpodobně nechtějí stahovat dvakrát ty stejné příspěvky, tak je potřeba zachovávat, které příspěvky si již uživatelé nechali stáhnout a které ještě ne.

Historie stahování se zapíše do databáze vždy při zahájení stahování, tedy při zmáčknutí tlačítka „Start download proces“ na stránce „Download/Structure“. Pokud je ovšem stahování přerušeno před jeho dokončením je historie tohoto stahování odstraněna.

Systém funguje tak, že se do tabulky „DownloadHistories“ vloží záznam s konkrétním uživatelem a seznamem id příspěvků, které byly obsaženy v daném stažení. Tento záznam se odstraní pokud stažení je zrušeno před jeho dokončením. Jestli stahování doběhne do konce, tak se z tabulky „DownloadHistories“ odstraní nepotřebné záznamy. Nepotřebný záznam je takový záznam, který je buď podmnožinou nových stažení nebo obsahuje id příspěvků, které nebyly získány z Redditu (id jsou již moc staré – už jsou za tisícem nebo byly odstraněny).

Při vypisování příspěvků uživateli se na serveru v souboru „Select.shtml.cs“ vytvoří seznam všech stažených id ze záznamů v databázi a ten se pošle do kódu Javascriptu, aby mohl dynamicky vypisovat příspěvky.

2.9 Šablony

Šablony (v aplikaci zvány jako Templates) je systém, pomocí kterého si mohou uživatelé ukládat nastavení na stránce „Download/Select“ a „Download/Structure“, aby stejné nastavení uživatelé nemuseli vyplňovat vícekrát.

Interakce s tímto systémem je tedy možná jenom na stránkách „Download/Select“ a „Download/Structure“. Na obou je tedy možnost jak vytvořit novou šablonu, tak ji i smazat. Při vytváření šablony má uživatel možnost šablonu pojmenovat.

Jelikož filtrování příspěvků pro zobrazení je řešeno na straně uživatele, tak se při změnách filtrů neposílá žádný dotaz směrem na server. Toto ovšem znamená problém pro ukládání změn v šablonách. Existují tři možnosti řešení. První řešení je při každé změně filtrů poslat dotaz, ve formě poslání Post, na server. Druhé řešení je při každé změně filtru poslat dotaz na server z kódu Javascript pomocí AJAX. Třetí je dotazy neposílat při změnách filtrů a dotazy posílat pouze u vytvoření šablony a posunutí uživatele na další krok (na stránce „Select“ to je „Next step“ a na stránce „Progress“ to je „Start Download“). Po důkladném zvážení všech možností se jako nejlepší jeví třetí řešení, tedy neposílat dotazy. Tato možnost je méně náročná pro server. A zároveň jenom malý počet uživatelů si všimne tohoto „nedostatku“.

Šablony jsou uchovávány v databázi aplikace pomocí balíčku EntityFramework of Microsoftu. V kódu C# šabloně odpovídá model „Template“.

Id	UserId	ShowD...	GroupBy...	Nsfw	DomainsForm	Numberi...	SubredditName	DomainNa...	PriorityNa...	Title	SubredditFo...	DomainFo...	PriorityFo...	Empty	Split	Name
1	1029	22qyqza5	0	0	ns...	[!redd.it",v.re... ids	0	0	1	60	1	0	1	1	1	All-NSFW
2	1030	22qyqza5	0	0	sfw	[!redd.it",v.re... ids	0	0	1	60	1	0	1	1	1	linkText
3	2026	sxao0b0o	0	0	both	[!redd.it",v.re... none	1	0	1	60	0	0	1	1	1	Standard

Obrázek 17 – ukázka šablon v tabulce „Templates“

Přidávání a odstraňování šablon probíhá skoro symetricky na stránkách „Select“ a „Structure“. V reakci na zmáčknutí tlačítkla se zavolá javascript funkce, která pomocí AJAX zavolá metodu v „c.shtml.cs“ souboru. Tato metoda potom přidá/odebere šablonu z databáze.

2.10 Statistika

Statistika má za úkol stráždat informace o stahování, jak pro jednotlivé uživatele, tak pro aplikaci. Tyto nastřádané informace jsou potom k nalezení na stránce „Statistics“.

Ve statistice se uchovává rozdělení stahovaných domén, rozdělení stahovaných subredditů, počet stažení a počet stažených příspěvků. Pro všechny uživatele dohromady se nezaznamenává rozdělení subredditů, jelikož subreddity jsou soukromá záležitost. Pro aplikaci se také zobrazuje počet registrovaných uživatelů a počet aktivních stažení, nicméně ani jedna z těchto informací není zaznamenána v tabulce „Statistics“, jelikož se tyto informace získávají přímo při vytváření stránky z jiných tabulek databáze. Zaznamenávání nových informací do statistiky probíhá při dokončení stahování v službě „DownloadManager“.

ZÁVĚR

Výsledkem práce je webová aplikace, dovolující uživatelům hromadně stahovat příspěvky z aplikace Reddit. Aplikace umožňuje uživateli se přihlásit pomocí jeho Reddit účtu. Následně dovoluje uživateli filtrovat a vybírat ze svých uložených příspěvků. Aplikace také umožňuje uživateli si zvolit jak se mají příspěvky pojmenovávat a nastavit jak mají být rozděleny do složek. Tyto nastavení si uživatel může uložit do šablon pro opětovné použití. Na stránce „Progress“ si potom uživatel může stáhnout dokončené procesy. Na stránce „Statistics“ uživatel najde statistiku o svém používání aplikace tak jako i statistiku celé aplikace.

Osobním přínosem práce je získání více zkušeností s většími a komplexnějšími projekty. Prohloubení znalostí s technologiemi *Razor Pages* s *ASP.NET Core*, C#, API a dalších. Realizace této práce mi také pomohla ocenit zavedení systému do práce. V mém případě se jednalo o excel tabulku, do které jsem si zaznamenával, jak dlouho jsem na bakalářské práci daný den pracoval, co jsem za ten den udělal a co bych měl dělat další dny. Tato tabulka mi sloužila jako repozitář poznámek o bakalářské práci.

Aplikace má také ale některé nedostatky, a je zde tedy prostor pro zlepšení. Optimalizace rychlosti získávání příspěvků z Redditu by byla užitečným krokem, bohužel bylo z testů zjištěno, že většina času je strávena čekáním na odpověď od API, tudíž prostor pro zrychlení není tak velký, jak by se mohlo na první pohled zdát. Dalším užitečným vylepšením by byla schopnost stránky „Download/Progress“ periodicky kontrolovat stav procesu stahování a pokud by byl dokončen, tak dovolit uživateli stáhnout nyní již dokončený proces. V neposlední řadě by bylo ideální pro statistiku vytvořit „Opt-out system“ (dát uživateli možnost nastavit si, že nechce, aby se o něm sbírala statistika). Tento systém by byl zároveň rozšířen i na zbytek aplikace, což by znamenalo, že by po uživateli po dokončení stahování na serveru žádná data nezbyla.

POUŽITÁ LITERATURA

- [1] KEMP, Simon. DIGITAL 2023 DEEP-DIVE: HOW MUCH TIME DO WE SPEND ON SOCIAL MEDIA?. *Datareportal.com* [online]. 2023, 26 JANUARY 2023 [cit. 2023-05-01]. Dostupné z: <https://datareportal.com/reports/digital-2023-deep-dive-time-spent-on-social-media>
- [2] KARTHICK, Manoj. ReddSaver. In: *Github.com* [online]. 2021 [cit. 2023-04-26]. Dostupné z: <https://github.com/manojkarthick/reddsaver/>
- [3] Introduction to ASP.NET Core. *Aspnetcore.readthedocs.io* [online]. Redmond, Washington: Microsoft, ©2017 [cit. 2023-04-27]. Dostupné z: <https://aspnetcore.readthedocs.io/en/stable/intro.html>
- [4] Introduction to Razor Pages in ASP.NET Core. *Learn.microsoft.com* [online]. Redmond, Washington: Microsoft, 2023 [cit. 2023-04-27]. Dostupné z: <https://learn.microsoft.com/en-us/aspnet/core/razor-pages>
- [5] *What is an API?* [online]. Armonk, New York: IBM, ©2023 [cit. 2023-04-28]. Dostupné z: <https://www.ibm.com/topics/api>
- [6] *Reddit by the Numbers* [online]. San Francisco: Reddit [cit. 2023-04-28]. Dostupné z: <https://www.redditinc.com/press>
- [7] Reddit API documentation. *Reddit.com* [online]. San Francisco: Reddit, ©2023 [cit. 2023-04-26]. Dostupné z: <https://www.reddit.com/dev/api/>
- [8] Reddit-archive: reddit/API. *Github.com* [online]. San Francisco: Reddit, 2015 [cit. 2023-04-26]. Dostupné z: <https://github.com/reddit-archive/reddit/wiki/API>
- [9] *Api/v1/scopes* [online]. San Francisco: Reddit, ©2023 [cit. 2023-04-26]. Dostupné z: <https://www.reddit.com/api/v1/scopes>
- [10] WATSON, M. What is C# used for?. *Stackify* [online]. Huntington Beach, California, 2020, 18 Sep 2020 [cit. 2023-04-26]. Dostupné z: <https://stackify.com/what-is-c-used-for/>
- [11] C# documentation. *Learn.microsoft.com* [online]. Redmond, Washington: Microsoft, ©2023 [cit. 2023-04-26]. Dostupné z: <https://learn.microsoft.com/en-us/dotnet/csharp/>

[12] Introduction to Razor Pages in ASP.NET Core. *Learn.microsoft.com* [online]. Redmond, Washington: Microsoft, 2023 [cit. 2023-04-27]. Dostupné z: <https://learn.microsoft.com/en-us/aspnet/core/razor-pages>

[13] BRIND, Mike. *Learn Razor Pages* [online]. 18 Feb 2021 [cit. 2023-04-27]. Dostupné z: <https://www.learnrazorpages.com/>

[14] HUGHES, Adam, STEDMAN, Craig, ed. Microsoft SQL Server. *Techtarget.com* [online]. Newton, Massachusetts, June 2019 [cit. 2023-04-27]. Dostupné z: <https://www.techtarget.com/searchdatamanagement/definition/SQL-Server>

[15] Reddit-archive: reddit/OAuth2. *Github.com* [online]. San Francisco: Reddit, 2016 [cit. 2023-04-26]. Dostupné z: <https://github.com/reddit-archive/reddit/wiki/OAuth2>

SEZNAM PŘÍLOH

Příloha A: Uživatelský návod

Příloha B: Zdrojový kód aplikace Reddit Online Downloader

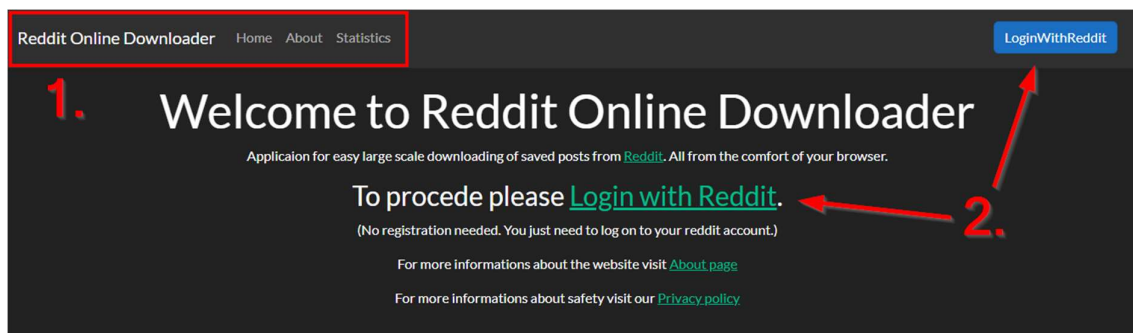
Příloha C: DDL skript na vytvoření databáze

PŘÍLOHA A: Uživatelský návod

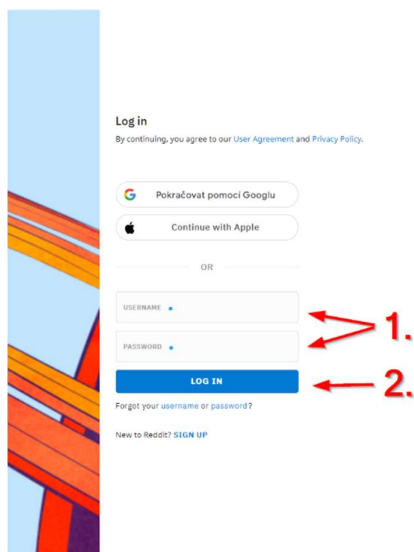
V tomto návodu se dozvíte jak se můžete přihlásit do aplikace, jak zahájit proces stahování, jak stáhnout dokončený process a jak pracovat se šablonami. V následujícím návodu bude označováno Reddit Online Downloader jako „ROD“.

Stáhnutí příspěvků

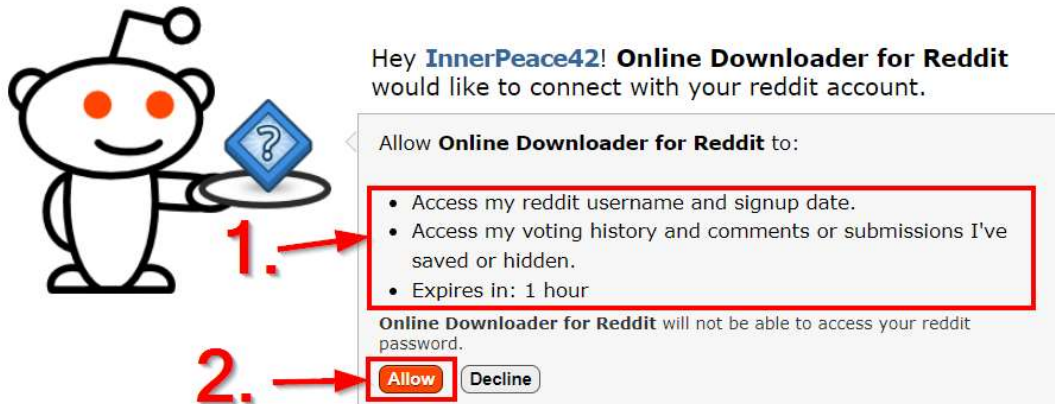
1. Při prvním otevření stránky se ocitnete na domovské stránce ROD. Na vrchní straně stránky se nachází navigační panel - 1. Tento panel slouží pro navigaci mezi stránkami aplikace. Klikněte na některé z tlačítek „Login with Reddit“ - 2.



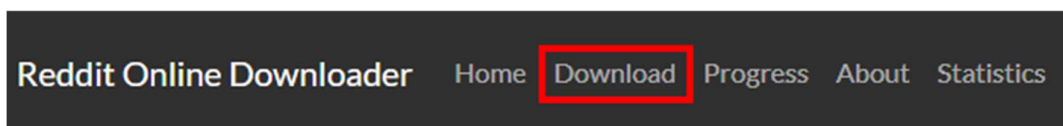
2. Pokud nejste přihlášení ocitnete se na přihlašovací obrazovce Redditu. Pokud přihlášení jste ,můžete pokračovat na krok 4. Pokud se na této stránce ocitnete, tak zde vyplňte vaše přihlašovací údaje k účtu Reddit – 1 a následně klikněte na tlačítko „Login“ - 2. (Tyto informace nejsou dostupné ROD)



3. Na této stránce vás žádá ROD o některé práva k vašemu účtu. Tyto práva zkontrolujte - 1 a povolte aplikaci práva dle svého uvážení pomocí tlačítka „Allow“ – 2 (bez udělených práv nebude aplikace fungovat).



4. Po zdařilém přihlášení se dostanete zpět na domovskou stránku. V navigačním okně teď klikněte na záložku „Download“.



5. Zobrazí se vám načítací obrazovka. Teď jen čekejte na dokončení získávání příspěvků. (může trvat až 40 sekund).



- Právě se nacházíte na stránce, kde si můžete vybrat, které z uložených příspěvků chcete stáhnout. V centrální části stránky můžete vidět seznam příspěvků. U každého příspěvku najdete zaškrťovací box na vybraní daného příspěvku - 1. Na vrchní straně seznamu také můžete najít zaškrťovací pole na oznašení/odznačení všech příspěvků - 2. Filtrovat seznam můžete pomocí formuláře na levé straně obrazovky - 3. Jakmile máte příspěvky vybrané, klikněte na tlačítko „Next step“ v pravém horním rohu obrazovky - 4.

The screenshot shows the 'Reddit Online Downloader' application interface. At the top, there is a navigation bar with 'Home', 'Download', 'Progress', 'About', and 'Statistics'. The user is logged in as 'InnerPeace42'. The main heading is 'Select posts to download'. Below this, there are several sections:

- Filters:** A section on the left with a search box and a list of filters including 'Show downloaded posts', 'Group by subreddits', 'Nsfw/sfw', and 'Media to show'. A red arrow labeled '3' points to this section.
- Post List:** A central list of posts, each with a checked checkbox and a link. A red arrow labeled '2' points to the top of the list, and another red arrow labeled '1' points to a checkbox in the list.
- Details:** A section on the right showing statistics for the selected posts, including 'Count - 31', 'Domains', 'Nsfw/sfw', and 'Subreddits'. A red arrow labeled '4' points to the 'Next step' button at the top right of this section.

At the top right of the main content area, there are buttons for 'Refresh posts' and 'Next step'. The interface is dark-themed with green accents.

7. Na této obrazovce si můžete zvolit, jak chcete, aby se vaše příspěvky uložily. V centrální části stránky vidíte seznam příspěvků ve formátu, ve kterém budou uloženy. Na levé straně najdete formulář, pomocí kterého si můžete zvolit parametry uložení - 1. Jakmile jste s podobou příspěvků spokojeni, tak klikněte na tlačítko „Start download process“ -2.

Reddit Online Downloader Home Download Progress About Statistics InnerPeace42 Logout

Naming/Foldering

Templates: None New Delete

Previous step **Start download process**

Names: 1. ←

Numbering: ∅
 Ids
 Standard
 None

Name contents: ∅

	y/n	Priority
Subreddit:	<input type="checkbox"/>	<input checked="" type="radio"/>
Domain:	<input type="checkbox"/>	<input type="radio"/>

Title: ∅
60

Folders: 1. ←

Group by: ∅

	y/n	Priority
Subreddits:	<input checked="" type="checkbox"/>	<input checked="" type="radio"/>
Domains:	<input type="checkbox"/>	<input type="radio"/>

Remove empty folders: ∅

Split by nsfw/swf: ∅

Details: 2. ↗

Count: 31
Domains: ∅
i.reddit - 16
v.reddit - 6
text - 5
link - 2
comment - 1
reddit.com - 1

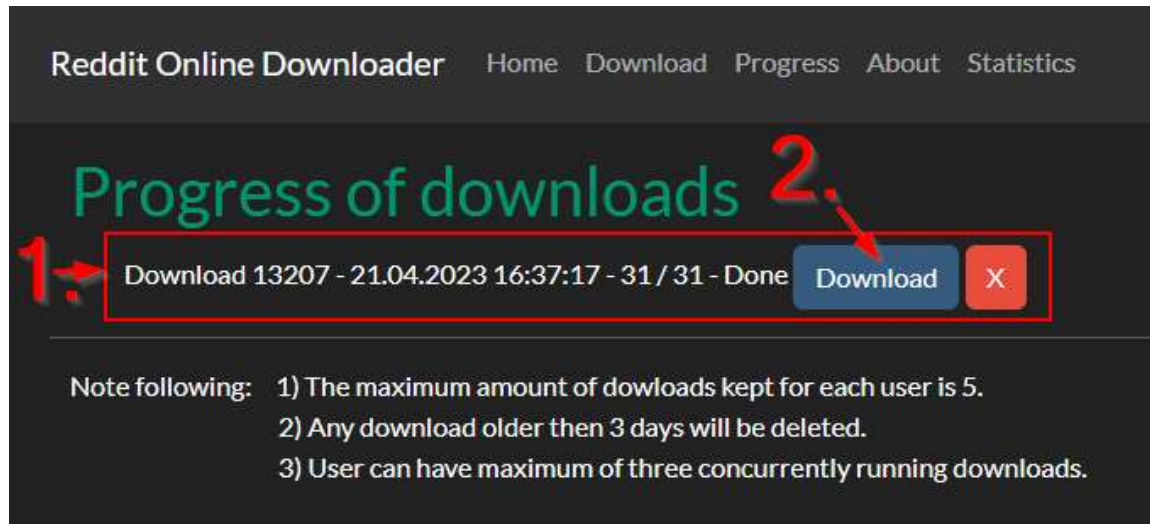
Nsfw/swf: ∅
swf - 29
nsfw - 2

Subreddits: 18
r/aww - 4
r/boardgames - 4
r/europe - 3
r/ProgrammerHumor - 3
r/sciencememes - 2
r/chess - 2
r/cars - 2
r/animegifs - 1
r/DidntKnowIWantedThat - 1
r/AskReddit - 1
r/TopTalant - 1
r/MealPrepSunday - 1
r/oddlyspecific - 1
r/books - 1
r/CozyPlaces - 1
r/simpleliving - 1

8. Byli jste přesměrování zpět na domovskou stránku aplikace. Teď klikněte na záložku „Progress“ v horním navigačním okně.

Reddit Online Downloader Home Download **Progress** About Statistics

9. Pokud jste předchozí kroky provedli správně, tak na této stránce uvidíte nad rozdělovací čarou položku, která představuje váš stahovací proces - 1. Pro aktualizaci stavu procesu musíte aktualizovat stránku – stačí znovu kliknout na záložku „Progress“. Jakmile bude stahovací process dokončen, tak se vám u něj zobrazí modré tlačítko „Download“. Klikněte na něj, aby jste zahájili stahování souboru na vaše zařízení - 2.



10. Po kliknutí na tlačítko „Download“ se vám otevře průzkumník, ve kterém si vyberte, kde chcete mít uložený soubor s příspěvkem. Tímto krokem jste úspěšně hromadně stáhli příspěvky z Redditu.

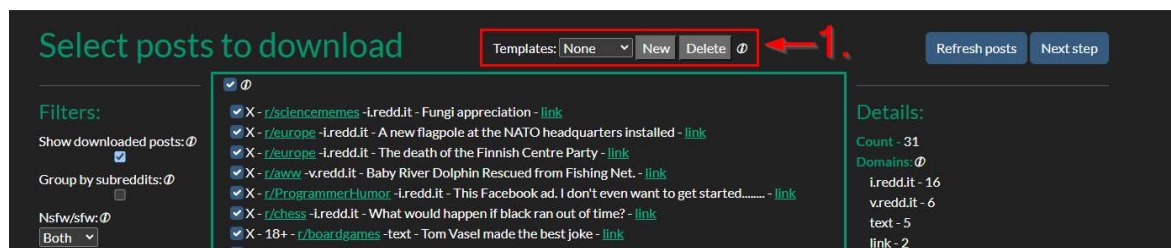
Šablony

Využit šablony můžete následovně:

1. V horním navigačním okně klikněte na „Download“.

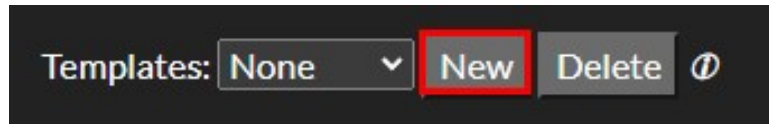


2. V horní části obrazovky vidíte „Templates:“ - 1. Zde můžete spravovat svoje šablony.



3. Vytvoření nové šablony

- a. Klikněte na tlačítko „New“.



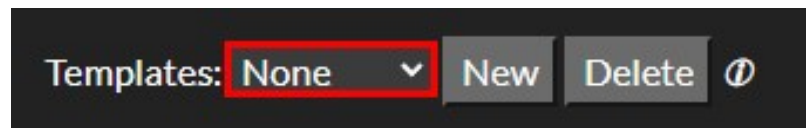
- b. Do dialogu vyplňte název nové šablony - 1. (název musí být unikátní) a následně dialog potvrďte - 2.



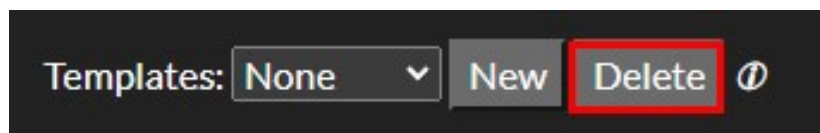
- c. Po úspěšném přidání nové šablony se zobrazí název šablony v poli vpravo od nápisu „Templates:“.



4. Přepnutí šablony - klikněte do pole vpravo od nápisu „Templates:“ zde si můžete vybrat z vašich šablon. Pokud nechcete mít vybranou žádnou šablonu tak zvolte „None“.



5. Odstranění šablony - kliknutím na tlačítko „Delete“ odstraníte aktuálně vybranou šablonu.



PŘÍLOHA B: Zdrojový kód aplikace Reddit Online Downloader

Tato příloha obsahuje archiv s projektem z Visual Studio 2022 jménem RedditOnlineDownloader.

PŘÍLOHA C: DDL skript na vytvoření databáze

Tato příloha obsahuje DDL skript pro vytvoření databáze pro aplikaci Reddit Online Downloader.