



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

OTEVŘENÁ DATA A OSOBNÍ ÚDAJE

OPEN DATA AND PERSONAL DATA

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Lukáš Cu

VEDOUCÍ PRÁCE

SUPERVISOR

JUDr. MgA. Jakub Míšek, Ph.D.

BRNO 2020

Bakalářská práce

bakalářský studijní program **Informační bezpečnost**

Ústav telekomunikací

Student: Lukáš Cu

ID: 197767

Ročník: 3

Akademický rok: 2019/20

NÁZEV TÉMATU:

Otevřená data a osobní údaje

POKYNY PRO VYPRACOVÁNÍ:

V bakalářské práci Otevřená data a osobní údaje se bakalant bude zabývat problematikou právní úpravy informací veřejného sektoru a možnostem jejich následného využití v kontextu střetu s právem na ochranu osobních údajů. V obecné části teoretické práce se bakalant zaměří na právní úpravu práva na informace a otevřených dat, představí základní právní rámec jejich poskytování a popíše výzvy v podobě ochrany osobních údajů. Cílem praktické části bakalářské práce bude vytvořit funkční aplikaci, která bude pracovat s vybranými datovými sadami obsahujícími osobní údaje poskytovanými jako otevřená data orgány veřejné správy. Ve zvláštní části teoretické práce proto bakalant popíše principy fungování připravované aplikace, a provede právní analýzu použití dat, která si k vypracování praktické části vybral.

DOPORUČENÁ LITERATURA:

- [1] MÍŠEK, Jakub. Právní aspekty otevřených dat [online]. Brno, 2019, 172 s. Rigorózní práce, Masarykova univerzita, Právnická fakulta, [vid. 30. 6. 2019]. Dostupné z: <https://is.muni.cz/th/sqe7a/>
- [2] KASL, František. Osobnost, soukromí a osobní údaje v moderní společnosti. In: POLČÁK, Radim et al. Právo informačních technologií. Wolters Kluwer, 2018, 391-485. ISBN 978-80-7598-045-8.

Termín zadání: 3.2.2020

Termín odevzdání: 8.6.2020

Vedoucí práce: JUDr. MgA. Jakub Míšek, Ph.D.

doc. Ing. Jan Hajný, Ph.D.
předseda rady studijního programu

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Bakalářská práce se zabývá problematikou otevřených dat v rozporu s právem na soukromí a ochranou osobních údajů. Čtenáři přibližuje základní pojmy s tím spojené a jejich usazení v české legislativě. Praktickým výstupem této práce je aplikace pracující s datovými sadami poskytovanými jako otevřená data obsahující anonymizované osobní údaje. Těmito sadami jsou konkrétně záznamy o počtech exekucí na území České republiky a rozdělení populace podle ukončeného vzdělání. Použité datové sady jsou v druhé kapitole analyzovány.

KLÍČOVÁ SLOVA

aplikace, exekuce, informační zákon, mapa, osobní údaje, otevřená data, vzdělání

ABSTRACT

This bachelor thesis deals with the issue of open data and its legal collision with rights for privacy and protection of personal data. It's purpose is to familiarize the reader with some essential terms regarding this topic and their place in Czech legislation. Practical part of this work is an application working with multiple data sets provided as open data. These data sets contain information about the number of Czech people under distraint and the ratio of people based on their education. These data sets are analyzed in the second chapter.

KEYWORDS

application, distraint, information law, map, personal data, open data, education

LUKÁŠ, Cu. *Otevřená data a osobní údaje*. Brno, 2019, 46 s. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: prof. Ing. Míšek Jakub, CSc.

PROHLÁŠENÍ

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

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

Brno

.....

podpis autora

PODĚKOVÁNÍ

Zde chci mnohokrát poděkovat JUDr. MgA. Jakubovi Míškovi, Ph.D. za vedení, podporu, podněty k práci a trpělivost, kterou se mnou v průběhu jejího vývoje měl. Hlavně však za umožnění dělat práci která mě bavila způsobem, který mě bavil.

Obsah

| | |
|--|-----------|
| 1 Úvod | 13 |
| 2 Právo na informace a otevřená data | 15 |
| 2.1 Zákon o svobodném přístupu k informacím | 15 |
| 2.2 Otevřená data | 16 |
| 2.2.1 Otevřená data v Českém právu | 17 |
| 2.2.2 Stupňování kvality otevírání dat | 18 |
| 2.3 Nejčastější právní překážky pro otevírání dat | 20 |
| 2.3.1 Práva duševního vlastnictví | 20 |
| 2.3.2 Ochrana soukromí a osobních údajů | 21 |
| 3 Webová aplikace a analýza dat | 27 |
| 3.1 Původní verze programu | 27 |
| 3.1.1 Funkcionalita původního programu z pohledu uživatele | 27 |
| 3.1.2 Popis funkcionality původního programu | 28 |
| 3.1.3 Cíle pro výstup bakalářské práce | 29 |
| 3.2 Nynější verze programu | 30 |
| 3.2.1 Funkcionalita nynějšího programu z pohledu uživatele | 30 |
| 3.2.2 Popis funkcionality nynějšího programu | 31 |
| 3.2.3 Dosažení cílů programu | 31 |
| 3.3 Překážky při tvorbě aplikace | 32 |
| 3.4 Analýza použitých dat | 34 |
| 3.4.1 Autorskoprávně chráněné dílo | 35 |
| 3.4.2 Výskyt osobních údajů | 35 |
| 3.5 Podmínky užití datových sad | 36 |
| 3.6 Podmínky užití datových sad vzdělání | 36 |
| 3.7 Podmínky užití knihoven AnyChart | 37 |
| 4 Závěr | 39 |
| Literatura | 41 |
| Seznam symbolů, veličin a zkratk | 45 |

1 Úvod

V dnešní době naše společnost každou sekundu vyprodukuje obrovské množství informací. To se týká i orgánů státní správy. Bez jakékoliv direkce nebo řádu by šlo jen o chaotickou zmiř bezcenných údajů. Se systematickým přístupem tyto informace však nabývají na hodnotě. Výběrem konkrétních dat a jejich propojováním se dají zjistit důležité informace. Informace zpracované a vyprodukované orgány státní správy jsou užitečné nejen občanům, ale i těmto orgánům samotným.

Lidé s přístupem ke správné kombinaci informací mohou zjednodušit svůj každodenní život nebo zlepšit své povědomí o okolí. Mohou také lépe vidět do dění na úřadě samotném. Orgány státní správy zase mohou sledovat efektivitu svojí práce nebo spolupracovat mezi sebou prostřednictvím vzájemného sdílení těchto informací. Efektivní informovanost o činnosti orgánů státní správy je tedy žádoucí ze všech stran, a právě k tomu slouží otevřená data. Jejich účel je jednoduchým způsobem zprostředkovávat informace veřejné správy v co nejlepší kvalitě.

Úřady však nemálo pracují s údaji, které mohou být předmětem právní ochrany. Vzniká tedy kolize a otázka, zda tyto informace přesto poskytovat. Právními překážkami pro otevřená data bývají nejčastěji práva duševního vlastnictví a práva na ochranu soukromí a osobních údajů. [1, s. 14]

Tato práce se zabývá právě problematikou konfliktu otevírání dat a případnou právní ochranou jejich obsahu, zejména jde o ochranu soukromí a osobních údajů.

Cílem teoretické části této práce je čtenáři přiblížit problematiku otevřených dat, osobních údajů a jejich vzájemného vztahu v kontextu právní úpravy v České republice. Jsou zde uvedeny subjekty povinné poskytovat informace a vymezení kvality otevřených dat.

Jako právní překážky jsou uvedeny již zmíněná práva duševního vlastnictví a práva na ochranu soukromí a osobních údajů.

Část věnovaná osobním údajům je rozšířená o popis povinností správce těchto údajů a popis právního titulu jako důvodu vzniku práva na jejich zpracovávání. Poslední podkapitola teoretické části pojednává o procesu zvaném anonymizace osobních údajů. Dle obecného nařízení o ochraně osobních údajů (dále jen GDPR) je tento proces vnímán jako možné řešení kolize otevírání dat a ochranou osobních údajů. [1, s. 103]

Praktickým výstupem této bakalářské práce je prototyp webové aplikace pracující s vybranými datovými sadami poskytovanými jako otevřená data. Těmito datovými sadami jsou data o počtu exekucí na území České republiky poskytována exekutorskou komorou České republiky a data o dosaženém vzdělání v jednotlivých krajích české republiky poskytována českým statistickým úřadem (dále jen ČSÚ). Důvodem pro zvolení těchto sad je právě kontext, který při spojování informací

vzniká. Z tohoto kontextu je možné najít případné spojitosti těchto datových sad. Počet exekucí nemusí být přesně dán dosaženým vzděláním, ale cílem bylo zjistit, zda se zde nachází nějaká korelace.

Celá druhá kapitola je dedikována této aplikaci. Popisuje její funkcionalitu, různé problémy při jejím vývoji a způsob jejich řešení. Dále obsahuje analýzu použitých datových sad, jejich podmínky užití a podmínky užití částí kódu třetí strany použitých v programu.

2 Právo na informace a otevřená data

Cíl této kapitoly je čtenáři přiblížit problematiku otevřených dat. Je zde popsán pojem otevřená data a také pojmy s nimi úzce spojené. Pojednává o důvodu existence otevřených dat a o jejich úpravě v zákoně. V neposlední řadě se zde uvádí případné právní překážky pro otevírání dat. Zejména je probírána ochrana osobních údajů.

2.1 Zákon o svobodném přístupu k informacím

V odst. 5 čl. 17 Listiny základních práv a svobod stojí: „*Státní orgány a orgány územní samosprávy jsou povinny přiměřeným způsobem poskytovat informace o své činnosti. Podmínky a provedení stanoví zákon.*“

Tím se rozumí zákon č. 106/1999 Sb. o svobodném přístupu k informacím. „*Povinnými subjekty¹, které mají podle tohoto zákona povinnost poskytovat informace vztahující se k jejich působnosti, jsou státní orgány, územní samosprávné celky a jejich orgány a veřejné instituce. Povinnými subjekty jsou dále ty subjekty, kterým zákon svěřil rozhodování o právech, právem chráněných zájmech nebo povinnostech fyzických nebo právnických osob v oblasti veřejné správy, a to pouze v rozsahu této jejich rozhodovací činnosti².*“ Povinné subjekty poskytují informace na základě žádosti nebo zveřejněním³. Poskytované informace jsou sdíleny všem žadatelům za stejných podmínek⁴.

Fungování orgánů veřejné správy je financováno z veřejných prostředků, což jsou daně občanů. Jako zdroj financí pro tuto činnost by občané měli mít k informacím vytvořeným jejich činností přístup⁵. [2, s. 17-18]

Svobodný přístup k informacím funguje jako mechanismus uvnitř i vně veřejné správy. Je to neodmyslitelný prvek fungování demokratického státu jako takového. [3]

Právo na informace může být omezeno, a to z různých důvodů. V čl. 17 pátém odstavci Listiny základních práv a svobod stojí, že „*Svobodu projevu a právo vyhledávat a šířit informace lze omezit zákonem, jde-li o opatření v demokratické společnosti nezbytná pro ochranu práv a svobod druhých, bezpečnost státu, veřejnou bezpečnost, ochranu veřejného zdraví a mravnosti*“. § 7 až § 11 zákona č. 106/1999 Sb. vyjmenovává případy, kdy povinný subjekt již nemusí poskytovat informace vztahující se k jejich působnosti. Jedná se například o ochranu utajovaných informací, ochranu obchodního tajemství, nebo ochranu důvěrnosti majetkových poměrů. Pro téma této

¹Povinný subjektem se zde rozumí orgán či osoba, která je povinna poskytovat informace vztahující se k jejich působnosti v rámci informačního zákona.

²Zákon č. 106/1999 Sb. § 2.

³Zákon č. 106/1999 Sb. § 4.

⁴To tedy platí i pokud je žadatelem povinný subjekt (Zákon č. 106/1999 Sb. § 4).

⁵Je vhodné mít transparentní činnost vůči občanům jako zdroji financování.

práce je však nejdůležitější § 8a a 8b. Ty totiž popisují zacházení s informacemi, které mají povahu osobních údajů.

Dle § 8a zákona č. 106/1999 Sb. platí, že „*Informace týkající se osobnosti, projevu osobní povahy, soukromí fyzické osoby a osobní údaje povinný subjekt poskytne jen v souladu s právními předpisy, upravujícími jejich ochranu.*“

Těmito předpisy se v tomto případě rozumí GDPR⁶. Úřady by tedy měly poskytovat veškeré informace kromě těch, které jsou zmíněny v § 7 až § 11 Zákona o svobodném přístupu k informacím. To platí i pro osobní údaje. Pokud tedy poskytování informací zahrnuje osobní údaje, je nutné aby se jednalo podle GDPR.

První pravidla pro poskytování informací pochází ze směrnice Evropského parlamentu a Rady 2003/98/ES. Ta udávala, že poskytování informací musí být co nejvíce uzpůsobené pro jejich následné užití veřejným sektorem. Nežadávala však povinnost tato data poskytovat. Tato směrnice byla novelizována směrnicí 2013/37/EU. Ta zadávala povinnost poskytování informací podávat ve vybraných formátech. Důvodem pro tuto úpravu je usnadnění automatizovaného zpracování. [1, s. 24-28] Směrnice bude v červenci roku 2021 nahrazena směrnicí EU 2019/1024. Jejím účelem je přizpůsobit poskytování informací k technologickému rozvoji. Klade důraz na zjednodušení automatického zpracování dat. Dále rozšiřuje působnost na veřejné podniky v odvětvích dopravy a veřejných služeb a na výzkumná data. Jde o data získaná výzkumem financového z veřejných prostředků. [4]

Tímto způsobem se postupně umožňuje využívání a poskytování otevřených dat. A právě otevřená data jsou jeden z nástrojů naplnění povinností vyplývajících z informačního zákona.

2.2 Otevřená data

Zákon o svobodném přístupu k informacím definuje otevřená data následujícím způsobem: „*Otevřenými daty se pro účely tohoto zákona rozumí informace zveřejňované způsobem umožňujícím dálkový přístup v otevřeném a strojově čitelném formátu, jejichž způsob ani účel následného využití není omezen a které jsou evidovány v národním katalogu otevřených dat.*“ [5]

Otevřená data jsou způsobem poskytování informací veřejného sektoru a tedy i nástrojem pro naplnění zákona o svobodném přístupu k informacím. Jejich smysl je usnadnění opětovného užití poskytovaných informací, a to za jakýmkoliv účelem. Otevírání dat nastavuje transparentnost, usnadňuje administrativu a tím nabízí celou škálu možností pro jejich další využití. [6, s. 8]

⁶Zákon 106/1999 Sb. odkazuje na zákon 101/2000 Sb. o ochraně osobních údajů, ale ten byl zrušen v dubnu roku 2019. Tento zákon je nahrazen obecným nařízením evropského parlamentu a rady EU o ochraně osobních údajů.

Čím více dat se zveřejní a propojí, tím ucelenější kontext se dá vytvořit. To znamená, že informační hodnota graduálně roste s množstvím nastřádaných informací a tím se i rozrůstají možnosti jejich využití. [7, s. 7]

Orgány veřejné správy svou každodenní činností generují a zpracovávají značné množství dat. Tato data jsou potřebná k vyplnění jejich zákonem daných povinností. Jedná se například o meteorologická měření nebo volební statistiky. [6, s. 8]

U většiny informací však jejich užitečnost nekončí pouze prvotním zpracováním. To otevírá možnost jejich využití v oblasti služeb. Vývojáři tedy vytvářejí aplikace pro zpracování datových sad vydaných orgány veřejné správy a jejich zprostředkování uživateli. Tím zjednodušují přístup k těmto datům nejen veřejnosti, ale i přímo orgánům, které tyto datové sady původně vydaly. Veřejnost tím může sledovat efektivitu práce úřadů nebo zlepšit přehled o dění v okolí. [2, s. 15-19]

Nabízí se možnost spojování více datových sad a vytváření dalšího užitečného kontextu. Například srovnání výsledků voleb s průměrným dosaženým vzděláním přidává informační hodnotu oběma použitým datovým sadám. Nemusí zde nutně vyplynout přímá závislost, ale z dlouhodobého statistického hlediska se dá zpozorovat trend, a tím tedy odhadovat preference voličů právě na základě jejich dosaženého vzdělání.

Povinné subjekty sdílející otevřená data však mohou zpracovávat údaje, které se dají považovat za údaje osobní. V tom případě spadají pod §8 a povinné subjekty, které s nimi pracují, se musí řídit povinnostmi a pravidly vyplývajícími z GDPR.

2.2.1 Otevřená data v Českém právu

Otevřená data jsou nástrojem vyplnění povinností plynoucích z odst. 5 čl. 17 Listiny základních práv a svobod. V České republice jsou otevřená data upravována zákonem č. 106/1999 Sb. o svobodném přístupu k informacím.

§ 4b tohoto zákona o poskytování informací zveřejněním tvrdí: *„Informace poskytovaná zveřejněním se poskytuje ve všech formátech a jazycích, ve kterých byla vytvořena; při zveřejnění takové informace v elektronické podobě musí být jeden z těchto formátů otevřený⁷ a, je-li to možné, též strojově čitelný.“*

Povinné subjekty podle § 4b *„zveřejňují informace obsažené v jimi vedených nebo spravovaných registrech, evidencích, seznamech nebo rejstřících, které jsou na základě zákona každému přístupné a které lze využít při podnikání nebo jiné výdělečné činnosti, ke studijním nebo vědeckým účelům anebo při veřejné kontrole povinných subjektů, jako otevřená data. Povinné subjekty zaevidují tyto informace v národním katalogu otevřených dat.“*

⁷Otevřený formát se dá číst pomocí volně dostupného softwaru.

Národní katalog otevřených dat je informační systém veřejné správy přístupný způsobem umožňujícím dálkový přístup sloužící k evidování informací zveřejňovaných jako otevřená data a jeho správcem je Ministerstvo vnitra⁸.

Seznam informací zveřejňovaných jako otevřená data je taxativně vypsán v Nařízení vlády č. 425/2016 Sb. o seznamu informací zveřejňovaných jako otevřená data. Jde o výčet dat, která se mají poskytovat v otevřené formě a která provádí ustanovení § 4b odst. 2 zákona č. 106/1999 Sb.

2.2.2 Stupňování kvality otevírání dat

Otevřená data jsou z definice dálkově dostupná a volná pro následné využití. Ne všichni však poskytují data ve stejné kvalitě. Tim Berners-Lee navrhl takzvané pětihvězdičkové schéma nasazení pro otevřená data. To hodnotí otevřená data jednou až pěti hvězdičkami podle jejich otevřenosti. Každý stupeň zahrnuje stupně předchozí a přidává novou vlastnost. [8]

V kontextu České republiky je vyžadován stupeň tři, nebo vyšší⁹ [9]. Následuje stručný popis pro každý stupeň tohoto schématu.

První stupeň (jedna hvězdička) – Data jsou dostupná na webu v libovolném formátu bez právních překážek. To mohou být například skeny dokumentů jako obrázky nebo PDF soubory.

Druhý stupeň (dvě hvězdičky) – Data jsou dostupná na webu ve strukturované, strojově čitelné podobě. Musí být tvořena znaky a dodržovat určitou strukturu, aby mohla být zpracovávána automaticky.

Třetí stupeň (tři hvězdičky) – Data jsou dostupná na webu ve strukturované, strojově čitelné podobě. Musí být tvořena znaky a dodržovat určitou strukturu, aby mohla být zpracovávána automaticky. Jsou také zpřístupněna v neproprietárním otevřeném formátu. S daty se tedy dá operovat pomocí volně dostupného softwaru.

Čtvrtý stupeň (čtyři hvězdičky) – Data jsou dostupná na webu ve strukturované, strojově čitelné podobě. Jsou v otevřeném formátu a je použito IRI k umožnění odkazování na jednotlivé prvky.

Pátý stupeň (pět hvězdiček) – Data jsou dostupná na webu ve strukturované, strojově čitelné podobě, je použito IRI a jsou na sebe nalinkovány prvky různých

⁸Zákon č. 106/1999 Sb. § 4c.

⁹Nicméně běžně se zde potkáme s otevřeností druhého stupně. [9]

sad pro poskytnutí kontextu.

Nezisková organizace Open knowledge foundation zase vytvořila deset principů otevřenosti dat, podle kterých se data také mohou definovat. Těchto deset principů udává přesné požadavky na otevřená data. Pokud je dodrženo alespoň prvních šest z těchto deseti požadavků, data se mohou považovat za otevřená. Následuje seznam těchto principů a jejich stručný popis. [10]

1. Data musí být úplná. – Musí tedy být publikována v podobě co nejvěrnější jejich původnímu obsahu. Míra této úplnosti může být dána zákonem.
2. Data musí být snadno dohledatelná. – Musí být dohledatelná pomocí běžně dostupné techniky a softwaru.
3. Data musí být strojově čitelná. – Musí být dostupná ve formátě, který má určitou strukturu. Může se s nimi tedy pracovat automatizovaně.
4. Data musí být dostupná v otevřeném formátu.– Formát, ve kterém jsou data dostupná, musí být čitelný softwarem s volnou distribucí.
5. Data musí být zpřístupněna s jasně definovanými podmínkami použití dat s co nejmenším omezením.– Podmínky pro data udávající rozsah jejich použití musí být jasně zformulované a zveřejněné.
6. Data musí být dostupná s použitím minimálního úsilí a nákladů. – Získání těchto dat musí být v co nejvyšší možné míře získatelná bez jakéhokoliv úsilí nebo potřebných finančních prostředků¹⁰. Pro poskytovatele dat¹¹ však může jejich poskytování znamenat nadměrnou zátěž. Může za ně tedy požadovat jednorázovou platbu úměrnou této zátěži.
7. Data jsou přístupná v podobě, která je totožná s podobou, ve které byla původně vytvořena.
8. Data jsou dostupná co nejdříve po jejich vzniku. Jediné zdržení je samotné jejich zpracování.
9. Data jsou dostupná všem stejným způsobem a ve stejné formě bez rozdílů.
10. Data jsou dostupná po dobu uvedenou poskytovatelem těchto otevřených dat.

Otevřená data představují praktický způsob poskytování informací veřejného sektoru. S jejich využitím se dá naplnit povinnosti vyplývající z informačního zákona a pokud jsou dodržena určitá pravidla, stává se z nich velice užitečný nástroj. S přístupem k otevřeným datům z mnoha zdrojů je možné vytvořit prostředí nabízející informace, které v originálních datech třeba vůbec nejsou.

¹⁰ Asi nejlepší způsob je volné stažení.

¹¹ Správce informačního systému veřejné správy nebo jiný povinný subjekt, který poskytuje otevřená data dle zvláštního právního předpisu.

2.3 Nejčastější právní překážky pro otevírání dat

Orgány státní správy mohou mít ze zákona povinnost poskytovat informace, které vytváří nebo zpracovávají. Nicméně tato povinnost může kolidovat s jinými zákony nebo lidskými právy. Nejčastěji se jedná o ochranu práv duševního vlastnictví a ochranu soukromí a osobních údajů. Oba případy jsou v následujících podkapitolách popsány.

2.3.1 Práva duševního vlastnictví

Poskytování informací veřejného sektoru, tedy i poskytování otevřených dat, se může dostat do kolize s právy duševního vlastnictví, zejména s autorskoprávní ochranou děl, případně ochranou zvláštních práv pořizovatele databáze. [1, s. 64]

Tato práva ošetřuje Zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (dále jen autorský zákon).

Práva duševního vlastnictví a otevřená data

Poskytování informací v podobě otevřených dat a právo duševního vlastnictví spolu v určité rovině kolidují. Základem ochrany práv duševního vlastnictví je kontrola nad danou informací. Nicméně princip otevírání dat je volně sdílet informace veřejného sektoru s širokou veřejností.[6, s. 18]

Dle autorského zákona je autorské dílo definováno v § 2 odst.1 jako „*dílo literární a jiné dílo umělecké a dílo vědecké, které je jedinečným výsledkem tvůrčí činnosti autora a je vyjádřeno v jakékoli objektivně vnímatelné podobě včetně podoby elektronické, trvale nebo dočasně, bez ohledu na jeho rozsah, účel nebo význam.*“ [11]

Obsah otevřených dat je zápis informací pro administrativní účely. Na jednotlivých prvcích datové sady většinou není nic jedinečného nebo unikátního a obsah samotný tedy až na výjimky není autorským dílem.

Dle § 2 odst.2 pod autorskoprávní ochranu však spadá i databáze. „*Databáze, která je způsobem výběru nebo uspořádáním obsahu autorovým vlastním duševním vytvořením a jejíž součástí jsou systematicky nebo metodicky uspořádány a jednotlivě zpřístupněny elektronicky či jiným způsobem.*“ Takové originální databáze se tedy považují jako souborné dílo a jejich struktura je chráněna. Zároveň však podle § 3 platí, že jestli je databáze úředním dílem, ochrana podle práva autorského se na ni nevztahuje¹².

¹²Podle § 3 se ochrana podle práva autorského nevztahuje na úřední dílo, jímž je právní předpis, rozhodnutí, veřejná listina, veřejně přístupný rejstřík a sbírka jeho listin, jakož i úřední návrh úředního díla a jiná přípravná úřední dokumentace, sněmovní a senátní publikace, pamětní knihy obecní a jiná taková díla, u nichž je veřejný zájem na vyloučení z ochrany.

Datová sada by tedy byla chráněna v případě, kdy by uspořádání jejích prvků bylo unikátní, a není dílem úředním.

Překážka otevírání dat je také ochrana zvláštních práv pořizovatele databáze. Podle § 88a autorského zákona platí: „*Zvláštní práva k databázi přísluší pořizovateli databáze, pokud porízení, ověření nebo předvedení obsahu databáze představuje kvalitativně nebo kvantitativně podstatný vklad bez ohledu na to, zda databáze nebo její obsah jsou předmětem autorskoprávní nebo jiné ochrany.*“^[11]

Zvláštní právo pořizovatele slouží jako ochrana jeho investice. Pořizovatel práva nabývá, pokud investoval podstatné množství prostředků do vzniku této databáze. Má právo požadovat peněžní kompenzaci za využití této databáze nebo její podstatné části třetí stranou. Tato ochrana tedy i motivuje k pořizování těchto databází. Pořizovatel je následně může využít například komerčně.

Pokud poskytování informací zasahuje do práv duševního vlastnictví třetích stran, je toto poskytování protiprávní. Pokud je však nositelem těchto práv právě poskytovatel, v poskytování informací mu nic nebrání. Jestli tak učiní, následné využití informací třetími stranami musí být řádně licencováno. To znamená, že následné využití stranami nedisponujícími duševním právem k obsahu informací musí být smluvně umožněno poskytovatelem pomocí licence. V kontextu otevřených dat je nejvhodnější veřejná licence¹³.^[1, s. 64-66]

Od prvního ledna 2017 se výjimka pro úřední dílo aplikuje i pro ochranu zvláštních práv pořizovatele databáze. ^[12]

Poskytování otevřených dat má v České republice relativně malé množství právních překážek týkajících se kolizí s ochranou duševního vlastnictví. Ty se týkají autorských děl a databází chráněných zvláštním právem pořizovatele databáze. Většinou se při poskytování informací vůbec nejedná o informace chráněné autorským právem, nebo naplňují definici takzvaného úředního díla, které má výjimku.

2.3.2 Ochrana soukromí a osobních údajů

Poskytování informací v podobě otevřených dat má s ochranou soukromí a osobních údajů v porovnání s ochranou duševního vlastnictví mnohem více překážek. Možnost data propojovat může vést k identifikaci subjektu těchto dat a to z nich dělá data osobní. ^[6, s. 20-21]

Právní úprava ochrany osobních údajů je obsažena v GDPR.

Pokud datová sada obsahuje osobní údaje, spadají na poskytovatele i příjemce povinnosti a odpovědnosti správce osobních údajů. Tyto povinnosti a odpovědnosti jsou podrobněji rozepsány v následujících podkapitolách.

¹³U veřejné licence nejsou poskytovatel a nabyvatel v přímém kontaktu. Licence je uzavřena poskytnutím informací nebo díla s příloženými licenčními podmínkami, které nabyvatel musí splnit.

Osobním údajem je jakákoliv informace použitelná k přímé či nepřímé identifikaci subjektu této informace. [13]

Škornicová uvádí, že „*mezi osobní údaje řadíme jméno, pohlaví, věk a datum narození, osobní stav, ale také IP adresu a fotografický záznam. Vzhledem k tomu, že se GDPR vztahuje i na podnikající fyzické osoby, řadíme mezi osobní údaje i tzv. organizační údaje, kterými jsou například e-mailová adresa, telefonní číslo či různé identifikační údaje vydané státem.*“ [16]

Informace vede k přímé identifikaci v případě, kdy k této identifikaci stačí právě tato informace samotná. K nepřímé identifikaci je nutné použití kontextu více informací, které by samy o sobě pro identifikaci nestačily. [1, s. 101]

Pojem osobní údaj je tedy velice rozsáhlý. Dá se totiž říct, že když informace, byť jen vágně, označuje konkrétní osobu, je možné tuto informaci použít k nepřímé identifikaci této osoby. Z toho by vyplývalo, že jakákoliv informace, která jen minimálně vypovídá o subjektu těchto údajů, je použitelná pro vytvoření kontextu a je tudíž osobním údajem.

Touto problematikou se zabýval Soudní dvůr evropské unie (dále jen SDEU) ve věci Patrick Breyer proti Bundesrepublik Deutschland v roce 2016. Německý občan Patrick Breyer navštívil několik webových stránek německých spolkových orgánů. Ty sloužily k podávání aktuálních informací svým občanům a byly volně přístupné. Mimo jiné však uchovávaly některé informace o svých návštěvnících, a to včetně IP adresy. Důvodem k tomu byla ochrana před případnými útoky na stránky. Pan Breyer se s žalobou u správního soudu dožadoval zákazu SRN uchovávání jeho IP adresy. Tato žaloba byla zamítnuta, a tak se pan Breyer odvolal. Rozhodnutí bylo odvolacím soudem pozměněno s konstatováním, že „*SRN neměla uchovávat IP adresu po ukončení internetového připojení Breyera, pokud byla tato IP adresa uchovávána ve spojení s datem otevření internetové stránky, k níž se IP adresa vztahuje, a pokud Breyer během svého připojení odhalil svou totožnost (prostřednictvím e-mailové adresy).*“ Dalším bodům žaloby pana Breyera vyhověno nebylo. Pan Breyer i SRN podali ke spolkovému soudnímu dvoru odvolání proti tomuto rozhodnutí. SDEU tedy řešil výklad směrnice 95/46/ES a její definici osobního údaje. [14] SDEU rozhodl, že dynamická IP adresa je osobním údajem a jeho definici upřesnil v bodech 44 – 49. Rozhodlo se, že dynamická IP adresa je osobním údajem, ale to pouze v případě, kdy by tato informace razantně zjednodušovala identifikaci subjektu. Dokud je množství úsilí či prostředků nutných k identifikaci dostatečně velké, o osobní údaj se nejedná. [15]

Z toho vychází i recitál 26 GDPR, kde stojí, že „*při určování, zda je fyzická osoba identifikovatelná, by se mělo přihlídnout ke všem prostředkům, jako je například výběr vyčleněním, o nichž lze rozumně předpokládat, že je správce nebo jiná osoba použijí pro přímou či nepřímou identifikaci dané fyzické osoby. Ke stanovení*

toho, zda lze rozumně předpokládat použití prostředků k identifikaci fyzické osoby, by měly být vzaty v úvahu všechny objektivní faktory, jako jsou náklady a čas, které si identifikace vyžádá, s přihlédnutím k technologii dostupné v době zpracování i k technologickému rozvoji. Zásady ochrany osobních údajů by se proto neměly vztahovat na anonymní informace, totiž informace, které se netýkají identifikované či identifikovatelné fyzické osoby, ani na osobní údaje anonymizované tak, že subjekt údajů není nebo již přestal být identifikovatelným. “[17] Je tedy nutné brát v potaz, nakolik je tato informace nápomocná a jaké náklady či úsilí je třeba vynaložit pro identifikaci subjektu s použitím této informace. Pokud je poměr míry rizika identifikace osoby k míře úsilí a finančních nákladů nutných pro tuto identifikaci dosti razantní, informace přestává být osobním údajem.

Ke zbavování dat osobních údajů slouží proces zvaný anonymizace údajů. Pokud se v otevřených datech vyskytují osobní údaje, poskytovatel i příjemce těchto dat se stávají jejich správci nebo zpracovateli a dopadají na ně povinnosti s touto pozicí spojené (viz 2.3.2). Je tedy nutná analýza dat pro vyhodnocení, zda jsou osobní údaje přítomny, nebo ne¹⁴. Při propojování datových sad, které takové údaje obsahují, totiž vzniká přídatný kontext a riziko identifikace subjektu roste. [1, s. 103]

Ze znění recitálu 26 GDPR vyplývá, že od určité míry anonymizace se data již nedají použít pro přímou či nepřímou identifikaci konkrétní fyzické osoby a přestávají tím pádem být osobními údaji.

Nestačí však pouze smazat jméno. Ostatní informace mohou tvořit jasný kontext ukazující na velmi malou množinu lidí či dokonce konkrétní osobu. Anonymizace dat se tedy dosahuje snižováním celkové informační hodnoty datové sady.

Anonymizace osobních údajů je tedy proces, při kterém data ztratí schopnost přímo či nepřímo identifikovat konkrétní osobu. To je dobrý způsob, jak umožnit zpracovávání těchto údajů a přitom se vyhnout případným právním překážkám spojeným s osobními údaji. Anonymizace se může dostáhnout různými způsoby. Někdy stačí údaje pouze zgeneralizovat¹⁵. Existují však i různé matematické algoritmy pro dosažení anonymizace¹⁶.

¹⁴Zjistit, zda jsou údaje použitelné k přímé či nepřímé identifikaci osoby.

¹⁵Už jen velká množina osob, na které mohou informace ukazovat může být dostačující. Při sdílení těchto informací je riziko identifikace minimální, protože k identifikaci téměř nepřispívají a identifikace samotná je velice složitá.

¹⁶Ty spočívají v alteraci obsahu těchto dat. Příkladem může být takzvaná k-anonymity. Ta sice také spočívá v generalizování a shlukování dat, ale je zde použit vzorec pro každý vstup. Jde o to, aby každý jednotlivý záznam měl minimálně k dalších shodných záznamů a data se generalizují nebo omezují o údaje, dokud tento požadavek nesplňují. k je předem určený koeficient.[21]

Povinnosti správce osobních údajů

Podle GDPR je správcem osobních údajů „každý subjekt, nerozhoduje jaké právní formy, který určuje účel a prostředky zpracování osobních údajů, provádí za jím stanoveným účelem jejich shromažďování, zpracování a uchování.“ Správce za zpracování osobních údajů plně odpovídá. Pro zpracovávání osobních údajů je nutný předpoklad alespoň jeden z právních titulů a dodržení principů pro jejich zabezpečení.[19]

Zásady zpracování osobních údajů jsou upravovány čl. 5 GDPR. Dle tohoto článku musí osobní údaje být ve vztahu k subjektu údajů zpracovávány korektně a zákonným a transparentním způsobem. Osobní údaje mohou být shromažďovány pouze s předem explicitně vysloveným účelem. Jakékoliv další zpracování s účelem neslučitelným tomu původnímu je neprávni. „další zpracování pro účely archivace ve veřejném zájmu, pro účely vědeckého či historického výzkumu nebo pro statistické účely se podle čl. 89 odst. 1 nepovažuje za neslučitelné s původními účely“ [17] Správce musí zpracovávat pouze osobní údaje potřebné k vyplnění původního účelu. Osobní údaje musí být přesné a v rámci možností aktualizované¹⁷. Ve formě, která umožní identifikaci subjektů těchto údajů, musí být údaje uloženy jen po dobu nutnou pro účely zpracování těchto údajů. Také je nutné osobní údaje zpracovávat způsobem, který tyto údaje zabezpečí před ztrátou nebo protiprávnímu použití.[20]

Všechny tyto zásady z čl. 5 odst. 1 Obecného nařízení musí správce osobních údajů nejen dodržet, ale musí být i schopen tuto skutečnost doložit¹⁸.

V čl. 24 Obecného nařízení je popsán princip odpovědnosti správce osobních údajů. Správce podle tohoto článku musí zavést vhodná technická a organizační opatření, aby zajistil a byl schopen doložit, že zpracování je prováděno v souladu s Obecným nařízením. Tato opatření musí být podle potřeby revidována a aktualizována.

Správci osobních údajů se v kontextu otevřených dat dělí na poskytovatele informací, kteří osobní údaje zpracovávají za účelem naplnění svých právních povinností a na ty, kteří data využívají ve vlastním zájmu. Nejčastěji to jsou tvůrci aplikací, které s daty pracují.[1, s. 109-110]

Právní titul

Podle GDPR je možné zpracovávat osobní údaje pouze na základě právního titulu¹⁹. Titul se stanovuje podle účelu tohoto zpracovávání a musí být stanoven a zajištěn před zpracováváním samotným. Zpracovávání osobních údajů bez právního titulu je tedy protiprávní. Zánik jednoho z více titulů neznamená nutnost údaje mazat²⁰. [22]

¹⁷To znamená, že při jejich nepřesnosti je buď opravit, nebo vymazat.

¹⁸GDPR čl. 5 odst. 2.

¹⁹Je možné tak činit i na základě více právních titulů. GDPR článek 6

²⁰Dokud správce disponuje alespoň jedním dalším

Právní tituly umožňující zpracování osobních údajů jsou vypsány v čl. 6 odst. 1 GDPR. Následuje výpis těchto titulů.

1. Souhlas se zpracováním osobních údajů – Platí, pokud sám subjekt souhlasí se zpracováním svých osobních údajů.
2. Plnění smlouvy nebo jednání o jejím uzavření – Platí, pokud je nutné zpracování osobních údajů pro plnění smlouvy, jejíž smluvní stranou je subjekt, nebo pokud je to nutné opatření před uzavřením smlouvy.
3. Plnění právní povinnosti – Pokud je zpracování osobních údajů součástí správcovi právní povinnosti, je toto zpracování nutné.
4. Životně důležitý zájem – V případě, kdy je zpracování osobních údajů v životně důležitém zájmu subjektu, je jejich zpracování nezbytné a v souladu s právem.
5. Veřejný zájem – Zpracování osobních údajů je nutné ke splnění úkolu, který je ve veřejném zájmu.
6. Oprávněný zájem správce – Platí, pokud je zpracování osobních údajů nutné pro naplnění oprávněného zájmu správce nebo třetí strany. Může se tedy jednat o téměř jakýkoliv rozumně vysvětlitelný důvod. Netýká se případů, kdy je důležitější zájem nebo základní práva a svobody subjektu pro ochranu těchto údajů.

V případě poskytovatele otevřených dat obsahujících osobní údaje je důležité rozeznat, zda data zpracovávají za účelem naplnění svých právních povinností, nebo tak činí z vlastní vůle. Jedná-li se o plnění jeho právních povinností, může data zpracovávat pod titulem plnění právních povinností. Pokud zpracovávání informací není právní povinností poskytovatele, tento právní titul není platný a pokud chce správce data dále zpracovávat, musí využít titulu jiného. [1, s. 110]

V rámci vytváření aplikace je nejvhodnější titul oprávněný zájem správce nebo třetí osoby, protože *„oprávněným zájmem může být výkon jakéhokoli práva, nebo činnosti, která je právem aprobovaná. Příkladem může být zajištění bezpečnosti sítě, jak je tomu v případě zpracování IP adres při řešení incidentů kybernetické bezpečnosti.“* [1, s. 123-124]

Tento právní titul je však použitelný pouze, když oprávněný zájem správce nepřevýší ochrana osobních údajů subjektů těchto údajů. Je tedy nutné před jejich zpracováním vzít v potaz, nakolik správce zasahuje do soukromí a zájmů jednotlivých subjektů a nakolik je oprávněný zájem správce důležitý. Čím méně se zasahuje do zájmů nebo základních práv a svobod subjektu, tím větší uplatnění tento titul může mít²¹. [1, s. 124]

Kromě dostatečného oprávněného zájmu je tedy i potřeba vzít v potaz, v jaké míře subjekt čeká následné zpracování osobních údajů²².

²¹Shodně viz recitál 47 GDPR.

²²GDPR recitál 47

Pokud jsou tedy data anonymizovaná, přestávají obsahovat osobní údaje. Potom platí, že povinný subjekt je může dle zákona č. 106/1999 Sb. volně zveřejnit na základě vlastního rozhodnutí a příjemce těchto dat s nimi může volně pracovat. Nemusí tedy plnit povinnosti správce osobních údajů a data může zpracovávat bez nutnosti právního titulu. Je však možné, že při nastrádání dostatečného množství dat a vytvoření dostatečného kontextu dojde k identifikaci subjektu. Tím se z těchto dat opět stávají osobní údaje a správce musí splnit povinnosti s tímto faktem spojené. [23, s. 113]

3 Webová aplikace a analýza dat

Výstupem bakalářské práce je aplikace pracující s konkrétními datovými sadami obsahující anonymizované údaje a prezentující informace z těchto datových sad vyvozené uživateli. Datové sady v této aplikaci zastupuje počet exekucí na území České republiky a rozdělení obyvatel podle ukončeného vzdělání. Výstupní aplikace je webová aplikace napsaná v jazycích JavaScript a HTML. Účel této části bakalářské práce je popsat čtenáři výstupní aplikaci a právní problematiku s ní spojenou. V následujících podkapitolách je přiblížena funkcionality aplikace z pohledu uživatele, technický popis procesů probíhajících v programu, seznámení s překážkami při tvorbě této aplikace, porovnání aplikace s plánovanou podobou a analýza použitých dat z právního hlediska.

3.1 Původní verze programu

Zde se nachází popis funkcionality programu odevzdaného jako semestrální práce. Slouží k porovnání se stavem programu nynějšího a poukázání na jeho vývoj. V následujících kapitolách je popsána nejnovější verze výstupní aplikace.

3.1.1 Funkcionality původního programu z pohledu uživatele

Datová sada o počtech exekucí je dostupná pro všechny, kteří mají přístup na internet. Z pohledu člověka je však velice nepraktické z této sady zjišťovat různé užitečné informace. Je v ní téměř jeden milion řádků údajů a manuální prohledávání je časově náročný úkon.

To byl důvod pro vytvoření této aplikace. Místo toho, aby člověk musel zdlouhavě prohledávat tak obsáhlý soubor, je mnohem jednodušší využít počítačový program, který z tohoto souboru zjistí skupinu pro uživatele podstatných informací a prezentuje je v grafické podobě. Uživatel se tedy z datové sady dozví veškeré informace, které sada obsahuje a nemusí je zdlouhavě hledat sám.

Aplikace je webová, a proto ji musí uživatel spustit přes webový prohlížeč. Po načtení stránky je uživatel vyzván k výběru vstupních dat z počítače pomocí tlačítka pro výběr souboru. Soubor, který uživatel vybere, musí být ve formátu CSV a musí mít stejnou strukturu jako soubor dostupný na webu exekutorské komory¹. Uživatel může soubor upravit. Důležité je, aby formát a struktura zůstaly stejné². Po vybrání

¹Tento soubor je zde dostupný ve formátu XLSX takže je nutné jej na počítači převést do formátu CSV. To se dá udělat například tak, že se soubor otevře v programu Microsoft Excel a při volbě „Uložit jako“ vybrat formát CSV(Comma-separated values).

²Může umazat, přidat, nebo upravit řádek.

správného souboru jsou data načtena do modulu mapy a na obrazovku se vykreslí mapa České republiky. Uživatel může přejíždět kurzorem myši přes jednotlivé kraje, a tím se mu zobrazí vybraná data o kraji. Konkrétně to je název kraje, celkový počet exekucí v tomto kraji, počet dlužníků v tomto kraji a jejich zastoupení ve vybraných věkových kategoriích. Kraje jsou na mapě také barevně označeny podle průměrného počtu exekucí na jednoho dlužníka. To je prozatím vše, co tento prototyp dokázal.

3.1.2 Popis funkcionality původního programu

Webová aplikace má uživateli zobrazit údaje o počtech exekucí v jednotlivých krajích. Jde o soubor funkcí, které vstupní údaje upraví do zpracovatelného tvaru a vloží je do mapy. Mapa je objekt vytvořený pomocí knihoven společnosti AnyChart. Jako první věc aplikace uživatele požádala o vložení vstupních dat. Tato data musela být ve formátu CSV a musí mít stejnou strukturu jako tabulky dostupné na stránkách exekutorské komory. Bohužel zpracování souboru ve formátu XLSX³ se nepodařilo v JavaScriptu zprovoznit. Text formátu XLSX totiž nemá podobu, ve které by mohl být převeden do JSONu a následně vepsán do tvaru zpracovatelného modulem mapy. Bylo tedy nutné stažená data na počítači do žádoucího formátu převést. CSV v textové podobě odděluje prvky předem definovaným symbolem⁴. V moment, co uživatel do funkce vloží soubor pomocí filechoosera se spustí funkce pro jeho prvotní kontrolu⁵. Následně je textový obsah nahrán do proměnné, která je přidána jako argument pro funkci sloužící ke kompresi. Původní soubor má téměř milion řádků. Vkládat tolik hodnot přímo do mapy je výpočetně náročné a zabralo by to tedy hodně času. Kompresní funkce celý obsah jednou přečte a vytvoří vstup pro mapu s již zpracovanými údaji⁶.

Při kompresi je také výstupem pole⁷ JSONů. JSON je formát, který je založen na kolekci páru klíč/hodnota. Bylo by možné udělat pouze pole polí. Pole může nést i další pole. Je tedy možné každému poli přiřadit soubor hodnot, které by sloužily jako údaje o jednotlivých krajích. Toto řešení by bylo výpočetně jednodušší a tedy praktičtější z časového hlediska. Použití pole JSONů místo pole polí má však v této aplikaci více výhod. Tím, že JSON identifikuje hodnoty podle jejich klíčů, případné úpravy pořadí nebo přidání dalšího údaje do datové sady neovlivní způsob, jakým budou hodnoty hledány. Klíče k hodnotám totiž zůstanou stejné,

³Formát, ve kterém je datová sada dostupná na stránkách exekutorské komory.

⁴V případě této aplikace jsou prvky oddělené čárkou. Funkce pro čtení souboru potom identifikuje prvky podle jejich rozdělení čárkami.

⁵Kontroluje se, zda je soubor vůbec čitelný

⁶To jsou již výsledné součty u jednotlivých hodnot (celkový počet exekucí v jednotlivých krajích, celkové počty dlužníků v jednotlivých věkových kategoriích, atd.).

⁷Pole je datová struktura, která je schopná držet konečný počet prvků stejného datového typu.

zatímco u přidání dalších polí by se změnilы indexy ukazující na požadované hodnoty a kód programu by se tomu musel přizpůsobit. Knihovna AnyChart také spoléhá na to, že vstupní data pro zápis budou ve formátu JSON. S tím, že těchto polí je vytvořeno relativně málo⁸, se časová prodleva z pohledu člověka ani tolik neprojeví. Vytvořené pole JSONů se dále použije jako argument pro načtení dat do mapy České republiky. Funkce z knihoven AnyChart vyhledá hodnoty z komprimovaného pole podle korespondujících klíčů a každému kraji podle něj přiřadí hodnoty. Poslední funkcí je zbarvování krajů podle počtu exekucí na dlužníka. Každé pole má totiž i JSON s hodnotou *value*, která se nezobrazí v tabulce, ale zbarví se podle ní příslušný kraj⁹.

Z logiky programu vyplývá, že spoléhá na vybrání souboru uživatelem. Kód se dá upravit tak, aby výběr dat proběhl automaticky. Tento způsob slouží pro ukázkou toho, že údaje nejsou v aplikaci pevně zakódované, ale při každém spuštění proběhne jejich načítání ze souboru. Změny hodnot v tomto souboru se tedy projeví na datech, která jsou načtena do mapy a následně zobrazena¹⁰.

3.1.3 Cíle pro výstup bakalářské práce

Z mapy se daly vyčíst údaje týkající se počtu exekucí v jednotlivých krajích. Aplikace splnila zadání, nicméně se jednalo pouze o kostru programu, který byl plánován k práci bakalářské. Následuje výčet původních návrhů a nápadů pro vylepšení této aplikace.

Prvním krokem bylo přidání podpory další datové sady. Například přidání demografických dat ČSÚ pro každý kraj by mohlo přinést zajímavé poznatky v kombinaci s informacemi o exekucích. Přidání další datové sady pravděpodobně bude znamenat i nutnost vytvoření další funkce pro její převod do zpracovatelného formátu. Funkce použitá pro zpracování dat o exekucích je zpracována na míru struktury tohoto souboru. Pokud tedy struktura přidané datové sady nebude stejná, nebude s nynější funkcí kompatibilní.

Další úpravy se týkaly úrovně, na které jsou data do mapy zapisována. V první verzi se údaje o exekucích zobrazovaly jen na úrovni krajů. Datová sada však obsahuje i údaje o konkrétních okresech. Mapa podporovala pouze vykreslení krajů, ale dá se vstoupit do jejího kódu a upravit jej tak, aby vykreslovala mapu podle okresů. To

⁸Jedno pole pro každý kraj, jedno pro údaje, které mají pro klíč kraj hodnotu „Neidentifikováno“ a jedno pro každý chybně zapsaný kraj při úpravě souboru.

⁹Barevné rozpětí je volitelné. V této aplikaci je vybrána světle modrá pro nejmenší počet exekucí na dlužníka a tmavě modrá pro počet nejvyšší.

¹⁰Změna však musí dodržovat původní strukturu souboru. Například přidání čárky do názvu kraje z něj z pohledu programu udělá dva prvky a naruší to přiřazování hodnot odpovídajícím klíčům.

by spočívalo v upravení hraničních bodů pro vykreslení a přidání identifikátorů pro vzniklé oblasti. Na webu data.gov.cz je také dostupná plně zakreslená mapa okresů.

S přidáním podpory další datové sady se manuální vybírání stává spíše překážkou. Informace o exekucích zastupuje jediný soubor, ale data o ukončeném vzdělání jsou rozdělena do čtrnácti souborů. Pro každý kraj je separátně vytvořena tabulka. Bohužel se nepodařilo získat informace na úrovni okresů. Už jen z praktických důvodů bylo tedy nutné tento krok automatizovat a vynechat lidský vstup.

Poslední plánovanou úpravou bylo vytvořit skript, který automaticky stáhne soubory datových sad z webu poskytovatele, ty převede do aplikací čitelného formátu a ta je následně zapíše do mapy. To by znamenalo, že uživatel nadále nemusí sám hledat datovou sadu potřebnou pro zápis, protože aplikace samotná by jej stáhla a zpracovala.

3.2 Nynější verze programu

V této kapitole je popsán nynější stav výstupní aplikace v porovnání se semestrální prací. Dále je zde shrnutí pro zjištění, zda byly plánované úpravy úspěšné.

3.2.1 Funkcionalita nynějšího programu z pohledu uživatele

Aplikace na výstupu funguje velice podobně jako předtím. Je zde tedy kladen důraz hlavně na věci, kterými se nynější podoba liší od verze předchozí.

Po spuštění webové aplikace se uživateli zobrazí mapa krajů s již zapsanými hodnotami. Mapa je zbarvená podle poměru dlužníků k počtu obyvatel. Při přejíždění nad jednotlivými kraji se však zobrazí navíc údaje o dosaženém vzdělání. Uživatel je tedy může rovnou porovnat s počtem exekucí v daném kraji. Dosažené vzdělání není nutně přímo závislé na exekucích nebo naopak. Nehledá se tedy kauzalita, ale případná korelace. Uživatel si stále může z počítače vybrat soubor exekucí a ten nahrát do map. To v případě, že má data buď aktuálnější, nebo chce porovnávat s daty staršími. Data vzdělání jsou převedena předem a z praktických důvodů se s nimi nijak jinak neoperuje. Je velice nepřehledné a nepraktické vybírat čtrnáct souborů pro každý kraj a to ještě ve správném pořadí. Navíc narozdíl od exekucí poměry vzdělání se nemohou tak rapidně měnit a není tedy potřeba pro jejich změnu. Dále uživatel může vybrat úroveň zobrazení dat. Je totiž implementována i mapa na úrovni okresů. Datová sada exekucí obsahuje i data na úrovni okresů, a tak bylo cílem zpřístupnit i jejich vyobrazení v grafické podobě. Bohužel data týkající se dosaženého vzdělání nesou pouze informace na úrovni krajů, a tak mapa okresů nese

pouze data o počtech exekucí. Tento problém by šlo vyřešit předáním stejných poměrů hodnot krajů. Problém je v tom, že tyto informace by mohly být nepřesné a tím by byla narušena jejich integrita.

3.2.2 Popis funkcionality nynějšího programu

Program v porovnání s verzí semestrální práce funguje téměř na stejném principu. Stále vypočítává komprimovaná data¹¹ a ta zapisuje do mapy.

Jednou z velkých změn je však podpora více datových sad. Nyní mapa nese a zobrazuje informace ze dvou na sobě nezávislých datových sad. Přidaná datovou sadou jsou informace o dosaženém vzdělání na úrovni krajů. Pro ty je vytvořen javascriptový soubor s již sečtenými hodnotami z tabulek. Důvod k tomu je ten, že čtení ze souboru je zbytečné přidávání výpočetní náročnosti programu. Informace o dosaženém vzdělání pochází z údajů poskytnutých ČSÚ v rámci sčítání lidu. Nejkaktuálnější tabulky jsou z roku 2011 a sčítání lidu probíhá jednou za deset let. Místo nahrávání informací uživatelem je tedy podstatně jednodušší je jednou za deset let aktualizovat na straně správce webu. Místo vypočítávání optimalizované tabulky ze čtrnácti různých souborů je tedy již tabulka předem vytvořena.

Další z úprav je přidání mapy zobrazující jednotlivé okresy. Původním plánem bylo převzít kód již implementované mapy krajů, připsat do něj dodatečné hraniční body a vzniklým celkům přidat identifikátory. Naštěstí na webu data.gov.cz se podařilo najít již zpracovanou mapu okresů. Bylo zapotřebí tuto mapu upravit tak, aby na ni bylo možné aplikovat funkce z knihoven AnyChart. Ty vyžadují, aby jednotlivé celky měly danou unikátní hodnotu *id*. stačilo tedy tyto hodnoty v tabulkách přiřadit a mapa byla připravena. Program pro zápis hodnot do mapy může zůstat stejný, protože použitá data jsou ten stejný soubor a struktura tedy zůstává nezměněna.

Poslední velkou úpravou je podpora XLSX formátu. Na fórech StackOwerflow se podařilo dohledat knihovnu umožňující převod XLSX tabulky do JSONu, tedy do formátu zpracovatelného naší aplikací. Uživatel tedy již nemusí stažený soubor otevřít a ručně převést do formátu CSV. Pro nahrání vlastního souboru jej akorát musí mít na počítači.

3.2.3 Dosažení cílů programu

Aplikace nyní podporuje další datovou sadu a je schopna zobrazovat údaje ze dvou různých datových sad zároveň. První cíl je tedy splněn.

Program je také schopen zobrazit mapu nejen na úrovni krajů, ale i na úrovni okresů. Problém je ten, že datová sada týkající se dosaženého vzdělání nese informace

¹¹Nyní i data týkající se dosaženého vzdělání v jednotlivých krajích.

pouze na úrovni krajů. Mapa okresů tedy zobrazuje pouze údaje o exekucích. Mapa okresů je však plně funkční a je připravena na případné užití dalších sad, které ponese i informace pro jednotlivé okresy. To znamená, že i tohoto cíle bylo dosaženo.

Také není nutné, aby uživatel data vybíral z počítače. Je vytvořen soubor s předem spočítanou tabulkou ve správném formátu, a tak si ji program nemusí počítat sám a má hned po spuštění odkud nahrát údaje do mapy. Sníží se tím výpočetní náročnost, a to povede ke kratší časové odezvě programu. Nabídka výběru souboru je stále přítomna pro případ, kdy uživatel bude chtít zobrazit údaje z datové sady například pro jiný rok, nebo by chtěl hodnoty nějakým způsobem pozměnit. Tím však donutí program k novému výpočtu tabulky. Celý krok zapsání do mapy může proběhnout bez interakce uživatele a je plně automatizován.

Poslední plánovanou úpravou bylo vytvoření skriptu, který by automaticky stáhl nejaktuálnější verzi přímo od zdroje. V případě naší aplikace se jedná o web exekutorské komory a ČSÚ. Bohužel tento skript se nepodařilo vytvořit. Stránky jsou totiž chráněny proti automatickému stahování souborů. V javascriptu se tedy nepodařilo najít způsob, jakým tuto překážku obejít.

Až na jednu výjimku se podařilo docílit všech plánů pro vylepšení výchozí aplikace a došlo i k některým neplánovaným úpravám. Například podpora původního formátu umožnila vynechání uživatelského vstupu při spuštění programu. Nicméně kvůli ochraně stránek exekutorské komory a ČSÚ proti automatickému stahování souborů bylo nutné předpočítat tabulku, kterou následně program zanesl do mapy krajů a okresů.

Aplikace splňuje zadání, dosáhla téměř všech plánovaných cílů a byla upravena i mimo ně.

3.3 Překážky při tvorbě aplikace

Původním plánem bylo využít modulu Google Maps¹² a data zapsat s jeho pomocí. Je možné v něm mapovat podle souřadnic a nabízí platformu pro projekci dalších informací. Pro použití tohoto modulu stačilo vygenerovat API klíč¹³ a ten pak použít ve svém kódu. Nicméně v průběhu vývoje funkcí pro převod dat do mapy přestal být API klíč podporován¹⁴.

Bylo tedy třeba vymyslet jiný způsob podávání dat uživateli. Pro splnění zadání byla preventivně vytvořena velice jednoduchá aplikace, která zobrazí obrázek mapy a tabulku s údaji pro jednotlivé kraje dle výběru uživatele¹⁵. Naštěstí se podařilo

¹²Jde o open source modul pro zobrazování a manipulaci s mapami.

¹³Dá se chápat jako unikátní klíč pro zpřístupnění modulu na stránkách majitele klíče.

¹⁴Důvodem k tomuto je pravděpodobně změna podmínek užití modulu Google map.

¹⁵Tato aplikace byla velice primitivní a nepodporovala změnu souboru.

nalézt alternativu pro modul map na webu www.anychart.com. Tento modul je určený pro vyjadřování dat v závislosti na lokaci. Jedná se například o zobrazení různých demografických údajů, což ideální pro použitou datovou sadu.

Další komplikací bylo vytvoření mapy okresů. Tato činnost spočívala v převzetí mapy krajů, přidání hraničních bodů okresů do tabulky a vzniklým celkům přiřadit vlastní identifikátory. Manuálně vytvořit několik set bodů a umístit je ve správném celku a pořadí je časově až nesmyslně náročné. Na webu data.gov.cz je však dostupná již hotová tabulka hraničních bodů okresů. Tato skutečnost ušetřila mnoho hodin manuální práce.

Choropleth mapa okresů z webu data.gov.cz však není mapou od AnyChart, a tak neobsahuje některé údaje potřebné k práci s používanými knihovny. Bylo tedy nutné buď jednotlivým okresům manuálně přiřadit identifikátor, nebo pro tento účel napsat příkaz. Je tedy vytvořena funkce přiřazující jednotlivým okresům identifikátory. Identifikátory okresů jsou shodné s jejich jménem.

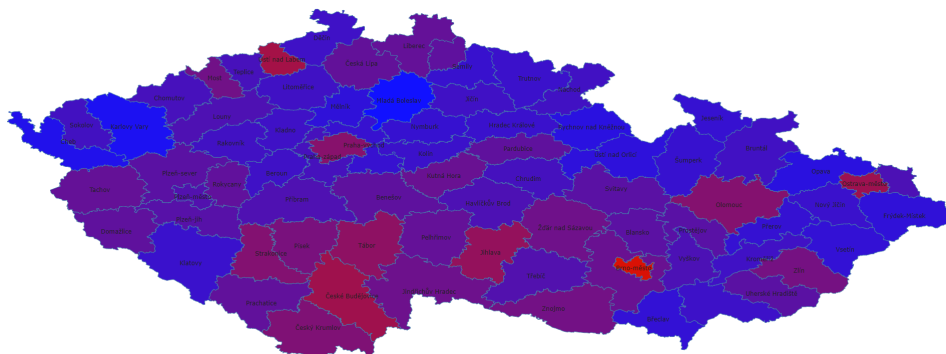
Funkce dále považovala kraj a okres Hlavní město Praha za stejný prvek nehledě na mapu. Proto hodnoty, které měly být zapsány do mapy okresů zapsala do mapy krajů. To vedlo k tomu, že kraj Hlavní město Praha měl dvojnásobné hodnoty a okres Hlavní město Praha neměl hodnoty žádné. Ve funkci se tedy jednotlivým oblastem muselo udat jejich druh. Funkce nyní kontroluje, o jaký druh oblasti se jedná, a až následně do ní zapisuje. Shoda jmen tedy již není problém.

Mapa okresů z webu data.gov.cz je mapovaná podle skutečných geografických souřadnic. Dále narozdíl od mapy krajů nemá udané žádné pravidlo pro projekci. Jedná se o to, jakým druhem zobrazení se mapa zakreslí. Země je kulatá a obrazovka počítače plochá. To vede ke zkreslení při přenášení zaobleného povrchu na povrch plochý. Různé druhy projekce tedy slouží k různým věcem¹⁶. Obě mapy se tedy zobrazovaly jinak a z důvodu vzhledu a rozmístění prvků na obrazovce bylo oběma mapám přiřazeno Mercatovo zobrazení¹⁷ (viz. obr. 3.1 a obr 3.2) . Obě mapy nyní mají stejné poměry stran a vzdáleností. Rozlišují se v tom, jak detailně jsou hranice vykreslené, ale to ve splnění účelu této aplikace nehraje roli.

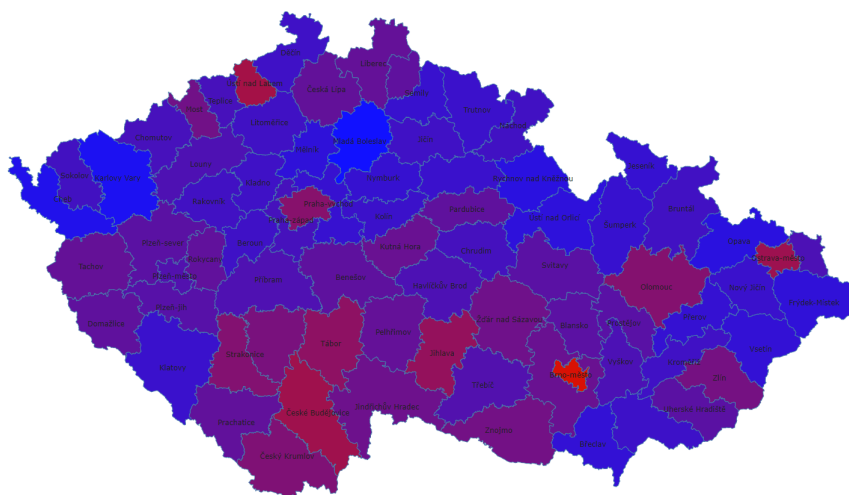
V původní verzi programu odevzdaného jako semestrální práce bylo nutné, aby uživatel manuálně stáhl datovou sadu, převedl ji do formátu CSV na svém přístroji a až tu pak předal aplikaci pro zapsání hodnot do mapy. To sice není nutně úplně problém a aplikace funguje i tak. Záměrem ale bylo omezit manuální činnost uživatele a byly přidány další datové sady. Na fórech StackOwerflow se podařilo najít knihovny umožňující převádět přímo soubor formátu XLSX. Datové sady se tedy

¹⁶Například u kuželového zobrazení nezáleží na relativním zachování vzdáleností, ale na udržení rovnosti jednotlivých poledníků. To logicky vede ke zkreslení tvarů.

¹⁷Při každé instanci vykreslování mapy, nehledě na její druh, jí tato projekce byla zadána jako pravidlo.



Obr. 3.1: Mapa okresů bez určeného zobrazení.



Obr. 3.2: Mapa okresů s Mercatovým zobrazením.

nemusi nijakým způsobem upravovat a program je umí rovnou zpracovat.

3.4 Analýza použitých dat

Použitá datová sada je skupina souborů nesoucí Použitá datová sada je převedena do formátu CSV. Důvod k tomuto převodu je jednoduchý. Formát CSV totiž jednotlivé prvky odděluje čárkou¹⁸. Datová sada nese informace o počtu exekucí v jednotlivých krajích a okresech. U každého záznamu je také uveden věk dlužníka.

¹⁸Nemusi být odděleny zrovna čárkou. Čárka může být nahrazena jakýmkoliv jiným znakem. Ten se dá jednoduše nastavit. Důležité ale je, aby se tento znak shodoval se znakem, podle kterého rozdělují data aplikace a aby se tento znak nevyskytoval jinde než na místě pro rozdělení.

3.4.1 Autorskopravně chráněné dílo

Autorským právem může být chráněna datová sada jako celek. To by platilo v případě, kdy by datová sada byla „*způsobem výběru nebo uspořádáním obsahu autorem vlastním duševním výtvorem a jejíž součásti jsou systematicky nebo metodicky uspořádány a jednotlivě zpřístupněny elektronicky či jiným způsobem.*“¹⁹ Jednalo by se tak o dílo souborné. Uspořádání nebo obsah ani jednoho ze souborů však není nijakým způsobem jedinečný. O díla souborné se tedy nejedná.

Všechny použité datové sady splňují definici úředního díla. Případné zvláštní právo pořizovatele databáze tedy podléhá výjimce.

Exekutorská komora i ČSÚ mají podmínky užití prezentované na svém webu. V podkapitole 3.5 jsou tyto podmínky uvedeny. Aplikace má tedy rozšíření sloužící k jejich splnění.

Jednotlivé prvky nenesou znaky autorského díla. Celý obsah je podán ve všeobecné formě. O prvcích se tedy dá říci, že nejsou výsledkem tvůrčí činnosti. Obsah datových sad není licencovatelný. Není tedy vůbec nutné licenci řešit.

3.4.2 Výskyt osobních údajů

Datová sada pocházející z webu exekutorské komory obsahuje záznamy týkající se každého dlužníka v exekuci na území České republiky. Každý záznam obsahuje informaci o kraji a okresu trvalého bydliště, věk dlužníka a počet exekucí dlužníka²⁰. Z definice pojmu osobní údaj vyplývá, že prvky použité datové sady buď jsou, nebo někdy byly osobními údaji.

Ve výstupní aplikaci jsou použita data ze čtrnácti souborů stažených z webu ČSÚ. Jsou zde zaznamenány počty obyvatel rozděleny podle věku, rodinného stavu a podle dosaženého vzdělání. V aplikaci jsou však použity pouze údaje o dosaženém vzdělání. U těchto dat by se tedy také dalo polemizovat o tom, zda se jedná o osobní údaje.

Jestli se o osobní údaje stále jedná je velice důležité, protože pokud to osobní údaje stále jsou, spadají na správce povinnosti správce osobních údajů (viz 2.3.2). Datová sada exekucí obsahuje záznamy s věkem a bydlištěm dlužníků s přesností na okres. Tabulky stažené z webu ČSÚ obsahují záznamy o rozdělení obyvatel podle věku, rodinného stavu a ukončeném vzdělání. Tyto údaje by se mohly považovat za údaje osobní. Hraje zde však roli anonymizace těchto údajů. Jednotlivé záznamy jsou

¹⁹Zákon č. 121/2000 Sb. § 2 odst. 2.

²⁰Jak se ale na webu exekutorské komory píše, tyto informace jsou zatíženy jistou mírou chybovosti a nejsou tedy naprosto přesné. Nemusí být uvedeny místa bydliště (okres nebo kraj) a věk nebo počet exekucí dlužníka může být nepřesný nebo chybět úplně. Tato nepřesnost má však i vedlejší efekt v podobě další anonymizace dat.

ve velkém shluku²¹ a jakýkoliv záznam označuje velkou množinu lidí²². Identifikace osoby, která má záznam v této datové sadě je s jejím použitím téměř nemožná. Ani propojení těchto dvou sad nevede k identifikaci subjektu nebo subjektů. Data se tedy ani po propojení nestala osobními údaji.

O datech použitých datových sad se dá říct, že již nesplňují definici osobních údajů a tudíž se na ně nevztahují zásady ochrany osobních údajů. Na zprostředkovatele těchto dat tedy nespádají povinnosti spojené se zpracováváním osobních údajů.

3.5 Podmínky užití datových sad

Datová sada týkající se počtu exekucí v České republice pochází z webových stránek exekutorské komory. Pokud je chce uživatel šířit (kopírovat, sdělovat veřejnosti), využívat nebo citovat, musí dodržet následující podmínky:

- jako zdroj dat musí být uvedena Exekutorská komora ČR (a to způsobem, který nevzbuzuje dojem, že Komora podporuje uživatele či způsob využití otevřených dat)
- v případě dalšího šíření dat je uživatel povinen vhodným způsobem uvést znění podmínek či odkázat přímo na tuto webovou stránku
- data mohou být použita pouze způsobem, který nijak neupravuje jejich význam nebo smysl

„Uživatel bere na vědomí, že data obsahují i údaje poskytnuté Komoře třetí osobou a jsou vytvořena i zpracováním údajů poskytnutých třetí osobou. Komora proto žádným způsobem neodpovídá za správnost, úplnost a pravdivost údajů poskytnutých třetí osobou, ani za případnou újmu vzniklou dalším užitím nebo zpracováním dat či způsobenou případnými chybami, které by data mohla obsahovat. Komora si vyhrazuje právo i bez předchozího upozornění ze závažných důvodů (technických, právních aj.) dočasně omezit přístup k datům.“[24]

3.6 Podmínky užití datových sad vzdělání

Údaje o dosaženém vzdělání pochází ze souborů dostupných na webu ČSÚ. V podmínkách užití stojí, že statistické informace je možné volně šířit, využívat a citovat v jiných dílech, nebo využívat komerčně. Je zde však vyžadováno dodržení následujících podmínek:

²¹Záznamů exekucí je dohromady téměř milion a soubory ČSÚ zahrnují celou populaci.

²²Nejpřesnější vyfiltrování je zjištění věku jedince v okrese. Hledat konkrétního jedince s konkrétním věkem v konkrétním okrese je stále téměř nemožné.

- ČSÚ musí být uveden jako zdroj těchto dat.
- U použitých dat nesmí být pozměněn nebo upraven jejich význam.

Pro co nejlepší dodržení těchto podmínek je v aplikaci implementována samostatná část. Vedle výběru souboru je také přítomen odkaz na stránku určenou pro splnění podmínek.

3.7 Podmínky užití knihoven AnyChart

AnyChart je soubor knihoven umožňující načtení dat do vybraného objektu v jazyce JavaScript. Tímto objektem mohou být grafy, tabulky nebo v případě výchozí aplikace mapy. Následně obsahuje funkce pro jejich vykreslení na stránce. Výchozí aplikace této práce využívá neplacenou verzi těchto knihoven. EULA společnosti AnyChart má pro uživatele této verze speciální podmínky užití.

„The trial version of this software may be used for evaluation purposes only at the user’s own risk. You are not allowed to use the Trial Version for any application deployment, commercial or production purpose. You may not remove any Trademarks, Watermarks or alter in any way²³.“ [25]

Po komunikaci s vývojářem těchto knihoven bylo schváleno jejich použití v rámci této práce za podmínky, že nebudou odebrána loga připnutá k jednotlivým vykresleným objektům.

²³Zkušební verze tohoto softwaru může být použita pouze pro hodnotící účely na uživatelské vlastní riziko. Zkušební verzi nejste oprávněn používat pro nasazení aplikace, komerční nebo produkční účely. Nesmíte odebrat žádné ochranné známky, vodoznaky, nebo je nijak upravovat.

4 Závěr

Cílem této bakalářské práce bylo přiblížit problematiku otevřených dat a osobních údajů. Praktickým výstupem této práce je webová aplikace, která uživateli zprostředkovává informace z datových sad obsahujících anonymizované údaje.

V teoretické části je stručný popis pojmů potřebných k pochopení implementace aplikace a jejich vztahů. V kontextu této práce je nejrozvedenější konflikt otevírání dat a práva na ochranu osobních údajů.

Praktická část obsahuje analýzu použitých datových sad¹. Je v ní zmíněn formát, ve kterém jsou dostupné a různé další technické parametry. Hlavně se však zabývá právními aspekty použití těchto dat. To tedy znamená, zda data neobsahují právně chráněné údaje. Z analýzy dat vyplynulo, že datové sady neobsahují žádné prvky, které by mohly spadat pod ochranu autorského práva.

Co se osobních údajů týče, jednotlivé záznamy obsahují prvky, které by mohly splňovat definici pojmu osobní údaj. Díky anonymizaci² této sady je poměr rizika identifikace a úsilí, kterého by se k této identifikaci muselo vynaložit, natolik jednostranný, že údaje obsažené v těchto sadách přestávají být údaji osobními. Na tvůrce webu s touto aplikací jako zpracovatele těchto dat tedy nespádají povinnosti správce osobních údajů a není třeba právního titulu.

Dále je popsána funkcionalita aplikace nejen z pohledu uživatele, ale i nejdůležitějších metod pro zpracování dat samotných. Jedna podkapitola je také věnována problémům při vývoji výstupní aplikace.

Při dokončení semestrální práce bylo určeno několik cílů pro výstup bakalářské práce.

Prvním z nich bylo implementovat podporu více datových sad. Aplikace podává informace o počtu exekucí a o poměrech ukončeného vzdělání v České republice. Tato kombinace byla zvolena z důvodu případné korelace mezi těmito dvěma údaji. První cíl byl tedy splněn.

Datová sada exekucí nese informace nejen na úrovni krajů, ve které byla mapa v semestrální práci, ale i na úrovni okresů. Byla proto implementována choropleth mapa z webu data.gov.cz a aplikace nyní podporuje i mapu okresů. Tato úprava také dopadla úspěšně.

Dalším krokem bylo eliminovat vstupy uživatele a automatizovat co nejvíce kroků programu. To se podařilo. Uživatel již nemusí z počítače vybírat soubor pro zakreslení. Původně k tomu mělo docházet tím, že skript tento soubor stáhnul sám přímo od zdroje a potom si jej přeložil do komprimované tabulky čitelné zakreslovací funkcí. Bohužel webové stránky exekutorské komory i ČSÚ byly chráněny před

¹Počet exekucí na území České republiky a rozdělení obyvatel podle ukončeného vzdělání.

²Data jsou generalizována a stěží ukazují na jednu osobu.

takovým vstupem a data se tedy automaticky stáhnout nepodařilo. Byla tedy uložena již předem vypočítaná tabulka a z té nyní program vychází. Možnost vybrání souboru je stále přítomna. Je možné že uživatel by chtěl zakreslit data například z jiného roku.

Až na skript automaticky stahující potřebné soubory bylo všech plánů docíleno.

Literatura

- [1] MÍŠEK, J. *Právní aspekty otevřených dat* [online]. Brno, 2018, 172 s. [cit. 19.11.2019]. Rigorózní práce. Masarykova univerzita, Právnická fakulta. Dostupné z URL: <https://is.muni.cz/th/sqe7a/rigo_opendata_final.pdf./>
- [2] BOČEK, J., ČEPICKÝ, J., MRÁČEK, J. *Jak otevírat data?* [online]. Praha: Fond Otakara Motejla, 2014, 113 s. 2004 [cit. 2.12.2019]. ISBN 978-80-87725-15-3. Dostupné z URL: <<https://osf.cz/publikace/jak-otevirat-data/>>
- [3] RIGEL, F. *Subjekty povinné poskytovat informace ve světle judikatury* [online]. 2010, poslední aktualizace 26.2.2010 [cit. 11.11.2019]. Dostupné z URL: <<https://bit.ly/38JIVmB>>
- [4] KUBÁŇ, M. *Evropská unie rozšiřuje pravidla přístupu k datům novou směrnici o otevřených datech* In: Portál otevřených dat [online]. [cit. 14.5.2019]. Dostupné z URL: <<https://data.gov.cz/2019/11/18/sm%C4%9Brnice-o-otev%C5%99en%C3%BDch-datech.html>>
- [5] *Zákon č. 106/1999 Sb. Zákon o svobodném přístupu k informacím* In: Zákon pro lidi [online]. [cit. 7.11.2019]. Dostupné z URL: <<https://www.zakonyprolidi.cz/cs/1999-106#f1945950>>
- [6] GÁLIKOVÁ, M. *PRÁVO NA INFORMACE A OTEVŘENÁ DATA* [online]. Brno 2019, 39 s. [cit. 19.12.2019]. Bakalářská práce. Vysoké učení technické v brně, Fakulta elektrotechniky a komunikačních technologií. Dostupné z URL: <https://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=193528>
- [7] MÍŠEK, J. *Otevřená data - Současný stav a doporučení pro otevírání dat* [online]. 2018, 35 s. [cit. 9.11.2019]. Dostupné z URL: <https://digistrategie.cz/opendata_2018.pdf>
- [8] BERNERS-LEE, Tim. *5 star open data* [online]. [cit. 19.11.2019]. Dostupné z URL: <<https://5stardata.info/cs/>>
- [9] KLÍMEK, J. *Důležité pojmy v oblasti otevřených dat* [online]. 2019 [cit. 1.12.2019]. Dostupné z URL: <<https://bit.ly/2syn0ho>>
- [10] OPEN KNOWLEDGE FOUNDATION. *Open Definition 2.1* [online]. 2004 [cit. 2.12.2019]. Dostupné z URL: <<http://opendefinition.org/od/2.1/en//>>

- [11] *Zákon č. 121/2000 Sb. Zákon o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon)* [online]. 2004 [cit. 2. 11. 2019]. Dostupné z URL: <<https://www.zakonyprolidi.cz/cs/2000-121>>
- [12] *Stanovení podmínek užití otevřených dat* In: Opendata.gov [online] poslední revize 15. 3. 2019 [cit. 16. 12. 2019]. Dostupné z:<<https://opendata.gov.cz/cinnost:stanoveni-podminek-uziti>>
- [13] *Zákon o zpracování osobních údajů* [online]. 2019 [cit. 2. 12. 2019]. Dostupné z URL: <<https://www.zakonyprolidi.cz/cs/2019-110>>
- [14] *C-582/14 - Patrik Breyer* [online]. 2019 [cit. 2. 6. 2019]. Dostupné z URL: <http://ictjudikatura.law.muni.cz/wiki/C-582/14_-_Patrik_Breyer>
- [15] *ROZSUDEK SOUDNÍHO DVORA - PATRIK BREYER* [online]. 2019 [cit. 24. 5. 2019]. Dostupné z URL: <<http://curia.europa.eu/juris/document/document.jsf?text=&docid=184668&pageIndex=0&doclang=CS&mode=lst&dir=&occ=first&part=1&cid=8975913>>
- [16] ŠKORNIČKOVÁ, E. *Co považuje GDPR za osobní údaje* [online]. [cit. 18. 12. 2019]. Dostupné z URL:<<https://www.gdpr.cz/gdpr/osobni-udaje/>>
- [17] *Nařízení Evropského parlamentu a Rady (EU) 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES* In: EUR-Lex [online] 2016 [cit. 18. 12. 2019]. Dostupné z:<<https://en.wikipedia.org/wiki/K-anonymity>>
- [18] *Povinnosti správce dle zákona o ochraně osobních údajů* [online]. [cit. 2. 11. 2019]. Dostupné z URL: <<http://www.oou.cz/povinnostispravcu>>
- [19] *Ochrana osobních údajů a poskytování otevřených dat* [online]. [cit. 2. 6. 2020]. Dostupné z URL: <<https://opendata.gov.cz/dokumenty:ochrana-osobn%C3%ADch-%C3%BAtaj%C5%AF-a-gdpr>>
- [20] DIBLÍK, J., JAROŠ, J. *Povinnosti správců a zpracovatelů ve světle Obecného nařízení o ochraně osobních údajů (GDPR), část I.* [online]. 2018 [cit. 18. 12. 2019]. Dostupné z URL: <<https://bit.ly/2s7Nc2H>>
- [21] *k-anonymity* In: Wikipedia [online] poslední revize 17. 11. 2019 [cit. 16. 12. 2019]. Dostupné z:<<https://en.wikipedia.org/wiki/K-anonymity>>

- [22] ŠKORNIČKOVÁ, E. *Právní důvody zpracování osobních údajů* [online]. [cit. 4.12.2019]. Dostupné z URL: <<https://www.gdpr.cz/gdpr/heslo/pravni-duvody-zpracovani-ou/>>
- [23] MÍŠEK, J. *Osobní údaje v čase a prostoru Role performativní regulace v ochraně osobních údajů* [online]. Brno, 2019, 231 s. [cit. 16.12.2019]. Disertační práce. Masarykova univerzita, Právnická fakulta. Dostupné z URL: <https://is.muni.cz/th/wpa9m/dis_final_03.pdf>
- [24] *Podmínky užití dat* In: EXEKUTORSKÁ KOMORA ČESKÉ REPUBLIKY [online]. [cit. 13.12.2019]. Dostupné z URL: <<http://statistiky.ekcr.cz/zakladni-data-o-exekucich.html>>
- [25] *AnyChart End User License Agreement* In: AnyChart [online]. [cit. 13.12.2019]. Dostupné z URL: <https://www.anychart.com/buy/eula/anychart_eula.pdf>
- [26] MÍŠEK, J. *Ochrana osobních údajů a poskytování otevřených dat* [online]. 2019 [cit. 4.12.2019]. Dostupné z URL: <<https://bit.ly/2reJXWI>>

Seznam symbolů, veličin a zkratek

| | |
|-------------|---------------------------------------|
| API | Application Programming Interface |
| CSV | Comma-Separated Values |
| EULA | End User License Agreement |
| GDPR | General Data Protection Regulation |
| HTML | Hypertext Markup Language |
| IRI | Internationalized Resource Identifier |
| JSON | JavaScript Object Notation |
| PDF | Portable Document Format |
| XLSX | Formát specifikace Office Open XML |
| XML | Extensible Markup Language |

