



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV INFORMAČNÍCH SYSTÉMŮ

DEPARTMENT OF INFORMATION SYSTEMS

**APLIKACE PRO ANALÝZU DOPRAVNÍCH UDÁLOSTÍ
VE MĚSTĚ MOST**

APPLICATION FOR TRAFFIC EVENT ANALYSIS IN THE CITY OF MOST

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. MIROSLAV ŠTĚPÁNEK

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. JIŘÍ HYNEK, Ph.D.

BRNO 2025

Zadání diplomové práce



164723

Ústav: Ústav informačních systémů (UIFS)
Student: **Štěpánek Miroslav, Bc.**
Program: Informační technologie a umělá inteligence
Specializace: Informační systémy a databáze
Název: **Aplikace pro analýzu dopravních událostí ve městě Most**
Kategorie: Informační systémy
Akademický rok: 2024/25

Zadání:

1. Prostudujte problematiku zpracování, vizualizace a vyhodnocování dopravních dat v městech. Prostudujte dostupné zdroje dat (PČR, Opendata, Waze) a jejich parametry.
2. Prostudujte principy zpracování a vizualizace dat na webu. Seznamte se s technologiemi vhodnými pro tento účel.
3. Analyzujte aktuální stav a nedostatky existující aplikace Silniční doprava města Most. Prozkoumejte využitelnost dostupných dat pro tento účel.
4. Na základě výsledků analýzy navrhnete datový model pro reprezentaci dopravních událostí ve městě Most integrující dostupná data. Dále navrhnete aplikaci pro automatizovanou aktualizaci dat a jejich následnou vizualizaci.
5. Implementujte navrženou aplikaci.
6. Výslednou aplikaci otestujte, vyhodnoťte její použitelnost a navrhnete možná rozšíření.

Literatura:

- Abdulhafedh, A. (2017). Road traffic crash data: an overview on sources, problems, and collection methods. *Journal of transportation technologies*, 7(2), 206-219.
- Chand, A., Jayesh, S., & Bhasi, A. B. (2021). Road traffic accidents: An overview of data sources, analysis techniques and contributing factors. *Materials Today: Proceedings*, 47, 5135-5141.
- Esri. (2024). *ArcGIS documentation*. ArcGIS. <https://doc.arcgis.com/en/>.
- Handscombe, J., & Yu, H. Q. (2019). Low-cost and data anonymised city traffic flow data collection to support intelligent traffic system. *Sensors*, 19(2), 347.
- Chen, W., Guo, F., & Wang, F. Y. (2015). A survey of traffic data visualization. *IEEE Transactions on intelligent transportation systems*, 16(6), 2970-2984.
- Leduc, G. (2008). Road traffic data: Collection methods and applications. *Working Papers on Energy, Transport and Climate Change*, 1(55), 1-55.

Při obhajobě semestrální části projektu je požadováno:
Body 1 - 4.

Podrobné závazné pokyny pro vypracování práce viz <https://www.fit.vut.cz/study/theses/>

Vedoucí práce: **Hynek Jiří, Ing., Ph.D.**
Konzultant: Ing. Kamil Novák; Město Most
Vedoucí ústavu: Kolář Dušan, doc. Dr. Ing.
Datum zadání: 1.11.2024
Termín pro odevzdání: 21.5.2025
Datum schválení: 22.10.2024

Abstrakt

Cílem práce je vytvořit náhradu za již existující programy města Most pro získávání, zpracování, vizualizaci a analýzu dopravních dat. Aktuální verze těchto programů jsou zastaralé a nabízí jen omezené možnosti analýzy. Jejich uživatelé tak nemají přístup ke všem dostupným informacím a nemohou je plně využít. Skript v jazyce Python vytvořený v rámci této práce umožňuje automatické získávání dat dopravních nehod z webových stránek Policie ČR, jejich zpracování a vyfiltrování pro dodaný polygon. Skript je možné využít v rámci ekosystému ArcGIS, ale i mimo něj bez nutnosti vlastnit licenci. Rozdělením na několik plug-in modulů umožňuje jednoduchou údržbu a univerzalitu v případě jakýchkoliv změn. Webová aplikace vytvořená pomocí frameworku React provádí vizualizaci zpracovaných dat v rámci vybrané mapy a dodatečných grafů, které zobrazují trendy a statistiky silnic a dopravních kamer. Data je možné analyzovat pomocí několika zobrazení s množstvím upravitelných parametrů. Webovou aplikaci mohou využít obyvatelé ORP Most a jeho okolí pro lepší informovanost o dopravním stavu. Zaměstnanci magistrátu města Most tyto informace mohou využít pro identifikaci problematických míst, na jejichž úpravu pak mohou soustředit dostupné prostředky.

Abstract

The goal of this thesis is to create a replacement for the existing programs used by the city of Most for acquiring, processing, visualizing, and analyzing traffic data. The current versions of these programs are outdated and offer only limited analysis capabilities. As a result, users do not have access to all available information and cannot fully utilize it. Python script developed as part of this project enables the automatic retrieval of traffic accident data from the website of the Czech Police, its processing, and filtering based on a provided polygon. The script can be used both within the ArcGIS ecosystem and independently without the need for a license. By dividing the code into several plug-in modules, it ensures easy maintenance and versatility in the event of any changes. A web application built using the React framework visualizes the processed data on a selected map along with additional charts that display trends and statistics related to roads and traffic cameras. The data can be analyzed through several views with a variety of adjustable parameters. Residents of the Most administrative district and its surroundings can use the web application to stay better informed about traffic conditions. Employees of the Most city hall can use this information to identify problematic areas and focus available resources on improving them.

Klíčová slova

Analýza dopravních dat, Město Most, Dopravní nehody, GIS, ArcGIS, Dopravní data

Keywords

Traffic data analysis, The city of Most, Traffic accidents, GIS, ArcGIS, Traffic data

Citace

ŠTĚPÁNEK, Miroslav. *Aplikace pro analýzu dopravních událostí ve městě Most*. Brno, 2025. Diplomová práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Jiří Hynek, Ph.D.

Aplikace pro analýzu dopravních událostí ve městě Most

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením pana Ing. Jiřího Hynka Ph.D. Další informace mi poskytl referent města Most Ing. Kamil Novák. Uvedl jsem všechny literární prameny, publikace a další zdroje, ze kterých jsem čerpal.

.....
Miroslav Štěpánek
15. května 2025

Poděkování

Chtěl bych poděkovat Ing. Jiřímu Hynkovi Ph.D. za vedení práce, cenné rady a zpětnou vazbu, která mi pomohla při její tvorbě. Zároveň bych chtěl poděkovat referentovi města Most Ing. Kamilu Novákovi za jeho čas, ochotu a cenné rady, které pomohly při tvorbě programové části práce.

Obsah

1	Úvod	2
2	Dopravní data ve městech	3
2.1	Dopravní data	4
2.2	Dostupné zdroje dat	6
2.3	Předzpracování dopravních dat	9
2.4	Vizualizace dopravních dat	10
2.5	Geografické informační systémy	12
2.6	Vyhodnocování dopravních dat pomocí GIS	16
3	Zpracování a vizualizace dat na webu	18
3.1	Vizualizace dat	18
3.2	Vizualizace vícedimenzionálních dat	20
3.3	Technologie pro vizualizaci dat	22
3.4	Knihovny pro vizualizaci dat	23
4	Analýza aplikace Silniční doprava a skriptu pro aktualizaci dat	25
4.1	Uživatelé aplikace	25
4.2	Současná podoba webové aplikace a skriptu	26
4.3	Požadavky na nové programy	29
4.4	Sady dopravních dat	30
5	Návrh webové aplikace a skriptu pro aktualizaci dat	33
5.1	Architektura systému	33
5.2	Datový model	33
5.3	Skript pro automatickou aktualizaci dat	35
5.4	Webová aplikace pro vizualizaci dat	37
6	Implementace nástrojů	40
6.1	Skript pro aktualizaci dat	40
6.2	Webová aplikace pro vizualizaci a analýzu dat	44
7	Testování vytvořených nástrojů	49
7.1	Uživatelské testování	49
7.2	Možné úpravy a rozšíření aktuální implementace	51
8	Závěr	52
	Literatura	53

Kapitola 1

Úvod

Dopravní zácpy, nehody a další komplikace na silnicích jsou pravděpodobně součástí každodenního života velkého procenta řidičů. Může se například jednat o každodenní dopravní špičku, kterou není silniční infrastruktura schopna efektivně pojmout. Při dopravních nehodách dochází v lepších případech pouze ke škodě na vozidlech. V horších případech dochází ke zranění a v nejhorších scénářích i k úmrtí zúčastněných osob.

Podle Světové zdravotnické organizace ztratí celosvětově při dopravních nehodách život více jak 1,1 milionu osob. Dalších přibližně dvacet až padesát milionů osob pak utrpí zranění, které nejsou fatální. Mezi hlavní faktory, které mají vliv na nehodovost a následky nehod patří příliš vysoká rychlost jízdy, nepřítomnost bezpečnostních prvků nebo nesoustředěnost řidičů. Mimo chyb ze strany řidičů přispívá k vyššímu výskytu nehod a horším následkům také zastaralá infrastruktura, která nedisponuje dostatečnými bezpečnostními prvky. Negativní vliv může mít i nepřítomnost nebo nedostatečné vymáhání dopravních omezení nebo nedostupnost dostatečné lékařské péče [65].

V rámci České republiky došlo v posledních třech letech k nárůstu celkového počtu vozidel na silnicích o přibližně sto osmdesát tisíc [49]. Tento rostoucí trend představuje problém hlavně pro oblasti, které se s dopravními komplikacemi potýkají již dnes. Města tak potřebují nástroje, které jim pomohou dělat správná rozhodnutí při plánování a úpravách infrastruktury, tak aby se komplikace na silnicích a jejich následky co nejvíce omezily.

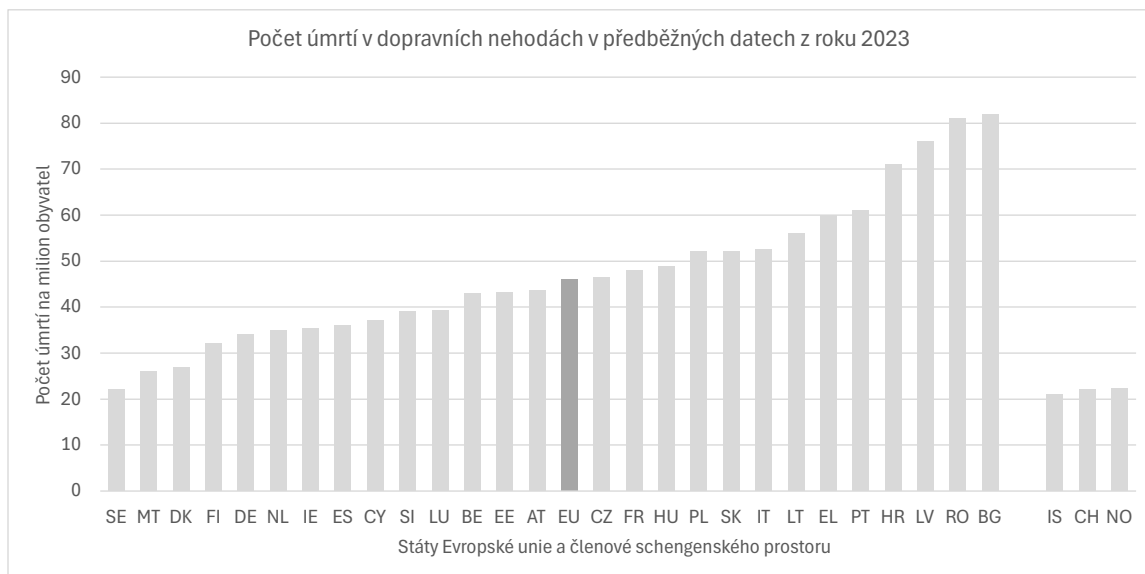
Tato práce se zabývá náhradou stávajících nástrojů města Most, které již nejsou schopny pracovat s novými daty dopravních nehod v plné kapacitě a umožňují nad nimi pouze omezenou analýzu. Jedná se o nahrazení skriptu pro zpracování dat, aby bylo možné pracovat s aktuálními a kompletními údaji. Následně se jedná o vytvoření nové webové aplikace, která bude nabízet rozšířené možnosti vizualizace a dovolí nad daty provádět pokročilejší analýzu, než je v aplikaci Silniční doprava možné. Nové nástroje umožní zaměstnancům magistrátu města Most lépe identifikovat problematická místa v silniční infrastruktuře a veřejnosti poskytnou informace o aktuálním stavu dopravní situace a jejím historickém vývoji.

Kapitola 2 se zabývá popisem dopravních dat, kategorizací zdrojů dat, možnostmi předzpracování, způsoby vizualizace a analýzou pomocí geografických informačních systémů. Kapitola 3 se zaměřuje na vizualizaci dat a webové technologie pro její tvorbu. Kapitola 4 se zabývá analýzou současných nástrojů města Most, jejich nedostatky a požadavky na jejich náhradu. Kapitola 5 se zabývá návrhem nových nástrojů na základě požadavků analýzy. Kapitola 6 popisuje implementaci navržených nástrojů. Kapitola 7 popisuje testování a využitelnost vytvořených řešení a navrhuje možné úpravy a rozšíření.

Kapitola 2

Dopravní data ve městech

Na základě dat z roku 2021 publikovala evropská komise studii [24] věnující se bezpečnosti na silnicích v členských zemích. V ní se Česká republika zařadila na sedmnáctou příčku v počtu úmrtí při dopravních nehodách. Urbanizované oblasti tvořily necelé dvě stovky takových případů, což byla přibližně třetina z celkového počtu. Při srovnání s údaji z předchozích let je vidět celkový klesající trend [67]. Ze současného umístění, které je vidět na obrázku 2.1, je však vidět, že v ohledu bezpečnosti na silnicích existují stále značné rezervy a umístění mezi ostatními státy EU by mohlo být lepší.



Obrázek 2.1: Předběžná statistika počtu úmrtí na území států Evropské unie a Schengenského prostoru v roce 2023 [23]. Hodnoty jsou přepočteny na milion obyvatel daného státu.

Jedním z hlavních faktorů, které se na této statistice negativně podílejí, je vysoká hustota dopravy. Když se na dopravní komunikaci vytvoří kolona nebo se alespoň značně sníží průměrná rychlost, začnou vznikat tendence dohnat časovou ztrátu. Ty následně vedou k rychlejší a obecně méně bezpečné jízdě, což zvyšuje šanci, že dojde k dopravní nehodě. Mimo zvýšenou šanci výskytu dopravních nehod má vysoká hustota dopravy negativní dopad také na množství vyprodukovaných emisí, spotřebované množství pohonných hmot a ztracený čas všech účastníků [17].

Aby se zamezilo vzniku dopravních zácp a udržela se plynulost provozu, spoléhají města na systémy řízení dopravy. Ty sbírají a ukládají data ze silničních senzorů, palubních mýtných přístrojů, semaforů a dalších zdrojů. Na základě těchto dat je možné identifikovat problematická místa [60], plánovat ideální trasy pro dopravní prostředky a předpovídat krátkodobý vývoj dopravní situace. Informace mohou řidiči obdržet například přes chytré telefony [17].

V případě Evropských států žije přibližně 74 % obyvatel ve městech a toto číslo se stále zvyšuje. Každým rokem je tak potřeba řešit stále složitější dopravní situaci. V tomto ohledu mohou pomoci omezení vjezdu do městských center, poplatky za parkování nebo finančně zvýhodněné využití městské hromadné dopravy [26]. Cílem je přesvědčit větší část obyvatel k takzvané multimodalitě. Multimodální lidé využívají více jak jeden dopravní prostředek pro dosažení své destinace. Toto by mělo snížit vyprodukované emise, zvýšit rychlost dopravy a snížit její celkovou cenu pro všechny účastníky dopravního provozu [38].

2.1 Dopravní data

Dopravní situaci byla věnována značná pozornost již od počátku minulého století. Ten byl ve Spojených státech amerických spojen s masovou produkcí automobilů, která snížila jejich cenu a tím je zpřístupnila většímu procentu populace. Výsledkem však byly každodenní dopravní zácpy, které navíc zpomalovaly tramvajovou dopravu. Komplikace působila i zaparkovaná auta, která zabírala velkou část ulic. Z těchto důvodů vznikla snaha řidiče kontrolovat, aby se předešlo negativním dopadům rostoucího počtu automobilů [44]. Hlavním zdrojem dat pro tyto rozhodnutí byly manuální počty. Tento způsob byl později doplněn nebo případně nahrazen mechanickými počítači, které zajišťovaly vyšší úroveň přesnosti a umožňovaly provádět dlouhodobější měření [3]. V Evropě se tyto problémy vyskytly až později, z důvodu vyšších cen automobilů a tím pádem i nižšího růstu vozidel na silnicích.

Charakteristika dopravních dat

Dopravní data je možné definovat jako datové sady, které byly získány pomocí senzorů a jiných měřicích zařízení. Tato zařízení se nachází v blízkosti dopravních komunikací nebo přímo v rámci sledovaných dopravních prostředků. Nejběžnějším druhem dopravních dat je sledování polohy dopravních prostředků. Z více záznamů polohy sledovaného objektu v čase je pak možné vytvořit jeho trajektorii [14].

Dopravní data je možné označit jako časoprostorová (ang. *spatio-temporal*). Tato kategorizace však nemusí být vždy vhodná pro všechny typy dopravních dat, a tak je možné data dále dělit do následujících podtypů podle obsažených údajů [66]:

- *Spatial-only data* obsahují pouze prostorové údaje a nenesou žádnou časovou informaci. Příkladem mohou být záznamy lokalit radarů ve městě nebo jiných významných míst.
- *Temporal-only data* obsahují pouze časové údaje a nenesou žádnou prostorovou informaci. Muže se jednat o souhrnné počty jako například počty dopravních nehod.
- *Spatio-temporal Static data* obsahují oba údaje, ty se však nijak nemění. Muže se jednat o popis bezpečnosti na daném úseku silnice nebo výskyt dopravní zácpy.

- *Spatial Static Temporal Dynamic data* mají neměnnou prostorovou a proměnnou časovou složku. Může se jednat o měření toku dopravy, průměrné rychlosti nebo sledování stavu počasí.
- *Spatial Dynamic Temporal Dynamic data* mají prostorovou i časovou složku proměnnou. Pomocí těchto dat je možné sledovat například souhrnné trajektorie vozidel na daných silnicích nebo rozsáhlejších oblastech.

Metody sběru dopravních dat

V dnešní době se ke sběru dopravních dat využívají intrusivní a neintrusivní metody. Intrusivní metody se přímo nachází na silnici a skládají se ze senzoru a záznamníku dat. Příkladem tohoto přístupu jsou:

- **Pneumatické silniční trubky** jsou položeny na povrchu silnice přes sledované pruhy. Při průjezdu vozidla je vytvořena vlna vzduchu, pomocí které čítač zaznamená průjezd vozidla. Tato metoda je určena spíše jako provizorní a levné řešení. Při nevhodném počasí nemusí tento systém fungovat spolehlivě. Navíc při pomalé průměrné rychlosti vozidel nemusí indikovat očekávanou hustotu dopravy [37].
- **Piezelektrické senzory** představují dlouhodobé měřicí řešení. V povrchu silnice je vytvořena drážka, do které jsou umístěny. Deformace tohoto materiálu při průjezdu vozidla vytvoří elektrický náboj, který zaznamená senzor vedle silnice [32]. Protože je vytvořený elektrický náboj úměrný způsobené deformaci, je možné měřit hmotnost a rychlost vozidel [37].
- **Magnetické smyčky** jsou určeny jako dlouhodobé řešení. Jedná se o zavedení drátu ve čtvercovém tvaru, který generuje magnetické pole, do povrchu vozovky. Jakmile vozidlo přejede přes tuto smyčku, způsobí kolísání frekvence, které zaznamená počítač vedle vozovky. Tento způsob měření je primárně využíván pro počítání průjezdů vozidel [32]. Nevýhodou tohoto zařízení je však nízká životnost, která je způsobená těžkými vozidly [37].

Intrusivní metody jsou založeny na vzdáleném pozorování vozovky. Jejich příkladem jsou:

- **Manuální počty** jsou využity v situacích, kdy není možné získat daný typ dat jiným způsobem. Pozorovatelé sledují vozidla a zaznamenávají chtěné údaje [37].
- **Mikrovlnný radar** zaznamenává pohyb vozidel a jejich rychlost. Z měření je možné získat také počet vozidel a jejich typ. Tento způsob funguje dobře v každém počasí [37].
- **Detekce obrazu z videa** se využívá pro automatickou identifikaci registrační značky. Pomocí té pak vyhledá všechny potřebné informace o vozidle ve státních databázích. Zároveň zaznamenává i počty vozidel. Pokud je navíc na trase daného vozidla dostatečný počet takovýchto kamer, dokáže určit nebo alespoň odhadnout jeho celkovou cestu po městě. Nevýhodou tohoto způsobu jsou vyšší pořizovací náklady kamer a nutnost instalace dostatečně velkého počtu [32].

Mimo zařízení, která byla přímo vytvořena pro účel sběru dat, mohou být tyto údaje získány i pomocí aplikací chytrých telefonů. Ty poskytují údaje o poloze mobilního zařízení, jeho rychlosti a poskytují data gyroskopu. Díky stále aktualizaci polohy je možné zrekonstruovat celou cestu daného vozidla [32].

2.2 Dostupné zdroje dat

Prvním krokem procesu získávání znalostí z dopravních dat je vyhledání vhodných zdrojů. Při výběru je vhodné hledět na kvalitu dat, které nabízí. Ideální data by měla co nej přesněji modelovat zachycenou situaci. Počet záznamů by měl být dostatečný a měly by být obsáhnuty všechny podstatné atributy. Záznamy by měly být konzistentní a důvěryhodné. Zároveň je vhodné, aby byl z dat co nejvyšší informační zisk. Pokud jde o provádění rozhodnutí, které budou mít do budoucna měřitelný vliv, jako například v případě úprav infrastruktury, je nutné, aby byla data aktuální a reprezentovala současnou situaci [33].

Data státních institucí a jejich časté nedostatky

Shromažďováním dopravních dat se obvykle zabývají státní instituce dopravy, které mají zpravidla nejkompletnější záznamy. Může jít ale také o zdravotnické, policejní nebo pracovní instituce, které sbírají data v odlišných okamžicích. Zdravotnické instituce tak mohou například sledovat pouze následky nehod po převezení do nemocnice, bez bližších detailů dopravní nehody [11]. Mimo případnou roztříštěnost dat se mohou vyskytnout i další problémy [11]:

- Data nemusí být vůbec shromažďována. V takových případech jsou skupiny, které s daty chtějí pracovat odkázány na odhady statistických modelů, které nemusí produkovat správná data.
- Hlášeno může být méně dopravních událostí, než se skutečně stalo. Jde o častý jev v odlehlých oblastech, kdy se méně závažné události vyřeší mezi účastníky. V případě lehčích zranění nemusí být v některých státech účastníci ochotní podstoupit lékařské ošetření z finančních důvodů, což může například znemožnit klasifikaci následků po nehodě. Problémem může být i samotná dokumentace událostí autoritami. Například pokud se některé instituce chtějí vyhnout kritice mohou hlásit méně incidentů.
- Data nemusí obsahovat všechny užitečné nebo očekávané atributy. V některých situacích mohou chybět kritické údaje a provádění rozhodnutí na základě takových dat nemusí být vhodné.
- Záznamy více státních institucí nemusí být vzájemně konzistentní.
- Shromažďovaná data nemusí být veřejně dostupná.
- Data mohou být publikována se značným zpožděním. Určitá míra zpoždění je očekávaná z důvodu zpracování a anonymizace dat. Problém však nastává při opoždění v rámci několika let, což znemožní reagovat na aktuální trendy.

Data státních institucí podle legislativy ČR

V případě České republiky se zmíněné problémy dat státních institucí z velké části nevyskytují z důvodu tuzemské a Evropské legislativy. Přesněji se jedná o zákon č. 106/1999 Sb. o svobodném přístupu k informacím¹, který zapracovává předpisy Evropské unie ze směrnice 2019/1024² o otevřených datech a opakovaném použití informací veřejného sektoru. Státní orgány a instituce tak mají povinnost zveřejňovat informace o své působnosti, data

¹<https://www.zakonyprolidi.cz/cs/1999-106>

²<https://eur-lex.europa.eu/legal-content/CS/TXT/?uri=CELEX%3A32019L1024>

registru a dynamická data, která nejsou v registrech obsažená. Mimo to jsou podle rozhodnutí 93/704³ Evropské rady členské státy povinny sbírat data o dopravních nehodách a sdílet tyto statistiky s orgány Evropské unie.

Otevřená data jsou pak podle § 3a odst. 5 zákona č. 106/1999 Sb. o svobodném přístupu k informacím⁴ definována jako „informace zveřejňované způsobem umožňujícím dálkový přístup v otevřeném a strojově čitelném formátu, jejichž způsob ani účel následného využití není povinným subjektem, který je zveřejňuje, omezen a které jsou evidovány v národním katalogu otevřených dat“. Podle [50] musí otevřená data splňovat následující požadavky:

- Přístupná v podobě datových souborů ve strojově čitelném a otevřeném formátu s úplným a aktuálním obsahem databáze nebo agregovanou statistikou.
- Obsahují neomezující podmínky užití.
- Evidovaná v Národním katalogu otevřených dat⁵.
- Opatřena dokumentací.
- Dostupná ke stažení bez technických překážek.
- Připravena s cílem co nejsnadnějšího strojového zpracování.
- Opatřena kontaktem na kurátora pro zpětnou vazbu.

Česká legislativa pak dodatečně definuje pět stupňů otevřenosti dat:

1. Datová sada musí být dostupná on-line s jasným vymezením podmínek užití. Mimo to nejsou kladeny žádné požadavky na formát dat.
2. Data musí být zveřejňována ve strojově čitelném formátu, který umožňuje co nejjednodušší zpracování běžně dostupnými programovacími jazyky a knihovnami. Nejsou kladeny žádné požadavky na datové formáty, a je tak možné využít i komerční formáty.
3. Specifikace datového formátu musí být otevřená a volně dostupná a pro její zpracování musí existovat volně dostupné programovací nástroje. Datová sada pak musí obsahovat pouze čistá data bez jakéhokoli formátování a případně doplňující údaje vyjadřující jejich sémantiku.
4. Při distribuci dat musí být identifikovány i entity, které jsou s daty spjaty.
5. Distribuce musí splňovat standardy propojených dat, které umožňují vyjádřit souvislosti mezi datovými sadami pomocí strojově zpracovatelných odkazů.

V rámci správních orgánů České republiky je požadován minimálně třetí stupeň otevřenosti. Data mohou být publikována i na nižších stupních, avšak pouze pokud existuje jejich publikovaná varianta alespoň na stupni otevřenosti tři [35].

Relevantní je také § 124 odst. 3 zákona č. 361/2000 Sb. o silničním provozu⁶. Ten po státních institucích vyžaduje, aby poskytly ministerstvu dopravy aktuální dopravní informace,

³<https://eur-lex.europa.eu/eli/dec/1993/704/oj/eng>

⁴<https://www.zakonyprolidi.cz/cs/1999-106#p3a-5>

⁵<https://data.gov.cz/>

⁶<https://www.zakonyprolidi.cz/cs/2000-361>

kteře mají vliv na bezpečnost a plynulost dopravy. Tyto informace pak musí ministerstvo zveřejňovat pro potřeby dopravní veřejnosti. Tuto roli aktuálně zastupuje Jednotný systém dopravních informací⁷.

Data mobilních aplikací

Existuje několik hlavních typů dopravních mobilních aplikací, které pomáhají jejich uživatelům šetřit čas nebo i peníze. Může se jednat o aplikace pro takzvaný *carpooling*, které umožňují jejich uživatelům cestovat ve větších skupinách s cizími uživateli za nižší cenu přepravy. Dalším typem mohou být aplikace, které uživatelům pomohou najít jejich zaparkované vozidlo nebo umožní analyzovat cestovní náklady [59].

Z pohledu zdrojů dopravních dat jsou však nejdůležitější aplikace, které pomáhají plánovat neoptimálnější cesty a informují řidiče o omezeních a dalších událostech, které mohly na trase nastat [59]. Populárními zástupci navigačních aplikací jsou Google Maps⁸ a Waze⁹. Statistiky počtu aktivních uživatelů nejsou zveřejňovány, ale podle dostupných dat¹⁰ je možné odhadovat, že se jedná o stovky milionů uživatelů měsíčně [55]. V případě České republiky je možné uvažovat i Mapy.cz¹¹, které v nejvytíženějších měsících využívá denně přes milion uživatelů [63].

Tento velký počet uživatelů umožňuje aplikacím využít takzvaných *crowdsourcing* metod pro nahlašování a moderaci událostí. Velkou výhodou je nízká cena tohoto přístupu, protože není nutné instalovat žádné senzory nebo jiné měřicí systémy. Instalované systémy navíc nemusí být schopné zachytit všechny druhy události, které mohou nastat a nemusí být schopny pokrýt stejně rozsáhlé oblasti jako tyto aplikace [39].

Značnou nevýhodou tohoto přístupu je však přítomnost nepravdivých, chybných nebo duplicitních hlášení od uživatelů. Aby bylo možné rozhodnout, zda jsou data pravdivá, jsou určeny indikátory, podle kterých je možné určit jejich validitu. Pro každého uživatele tak obvykle existuje jedna nebo více hodnot, které určují jeho důvěryhodnost a přesnost. Tato hodnota může platit pouze pro krátkodobé časové okno [42] nebo může být trvale spojená s uživatelským účtem, jako v případě aplikace Waze [28]. Pokud se více hlášení týká stejné události jsou tyto data agregována, aby byla dostupná jejich nejaktuálnější podoba [9].

Mimo manuálních hlášení sbírají tyto aplikace data i pasivně. Jak bylo zmíněno v sekci 2.1, jedná se o polohová data a informace ze senzorů zařízení, na kterém je aplikace nainstalovaná. Příklady takového sběru uživatelských dat je možné najít i ve smluvních podmínkách aplikací Waze¹² nebo Mapy.cz¹³. Historická data jsou pak například využita v kombinaci s těmi aktuálními pro automatické hlášení zvýšené hustoty dopravy v daném úseku [30].

Data zmíněných aplikací nejsou volně k dispozici. Pro Waze však existuje partnerský program Waze for cities¹⁴, který umožňuje přístup k datům z určené oblasti. Primárním účelem tohoto programu je poskytnutí informací pro budoucí plánování a úpravy infrastruktury státním institucím a územním celkům [29].

⁷Jednotný systém dopravních informací: <https://dopravniinfo.cz/cs/pages/jsdi>

⁸<https://www.google.com/maps>

⁹<https://www.waze.com/cs/live-map/>

¹⁰Přibližný počet aktivních uživatelů aplikace Waze: <https://support.google.com/waze/answer/6071177?hl=en#zippy=%2Chow-does-waze-work>

¹¹<https://mapy.cz/>

¹²Smluvní podmínky Waze: <https://support.google.com/waze/answer/12075406?hl=en#collected>

¹³Smluvní podmínky Mapy.cz: https://licence.mapy.cz/?doc=mapy_zzoo

¹⁴Partnerský program Waze for Cities: <https://www.waze.com/cs/wazeforcities>

2.3 Předzpracování dopravních dat

Před dalším zpracováním dat, je často nutné provést jejich předzpracování. Tento proces se dělí na přípravu dat a na jejich redukci. V rámci přípravy se provádí čištění, transformace a integrace dat. Redukce se zabývá výběrem vhodných atributů, výběrem reprezentativní podmnožiny a diskretizací hodnot. Výsledkem jsou úplná a spolehlivá data, která vyžadují nižší výpočetní výkon pro jejich zpracování [27].

Nutnost předzpracování dat je možné určit pomocí dimenzí kvality. Množství a typ dimenzí pro určení kvality dat se liší podle pole výzkumu a využití dat. Mezi dimenze, které se vyskytují častěji patří [58]:

- **Včasnost** určuje rozsah stárí dat, ve kterém jsou vhodné pro zpracování. Tato složka se kromě stárí dat zabývá také jejich mírou volatility. Jinak řečeno, jak často se daná data mění.
- **Aktuálnost** popisuje míru, v jaké data zahrnují nejnovější informace. Data jsou aktuální, pokud obsahují nejnovější informace i po možných změnách.
- **Úplnost** určuje míru, s jakou jsou data schopna reprezentovat danou skutečnost, případně zda mají dostatečnou šířku a hloubku [58]. Šířka zde popisuje dostupné atributy a hloubka dostupný počet hodnot [33].
- **Přesnost** určuje, do jaké míry jsou uložená data správná, spolehlivá a odpovídají skutečností reálného světa.

Čištění dat

Čištění dat zajišťuje odstranění chybějících hodnot, identifikaci odlehlých hodnot a jejich vyhlazení. Nepřítomnost očekávaných hodnot se často objevuje jako následek vadného senzoru, chyby v procesu nebo jiných limitací. Tento nedostatek je možné jednoduše odstranit, stačí zahodit záznamy s chybějícími hodnotami. Většinou se ale nejedná o vhodné řešení, protože by mohlo mít významný vliv na výsledky dalšího zpracování. Pro jejich doplnění je tak častěji využito pravděpodobnostních modelů a strojového učení [27]. Dalším řešením může být nahrazení průměrnou hodnotou atributu nebo průměrnou hodnotou záznamů, které spadají do stejné třídy záznamu [33]. Pro odstranění odlehlých hodnot se využívá filtrování nebo vyhlazení hodnot [27].

Transformace dat

Účelem transformace je úprava podoby dat tak, aby byla vhodná pro následné zpracování a neměla negativní efekt na jeho výsledky. Detailní údaje tak mohou být agregovány, kategorie atributů zobecněny, numerické atributy normalizovány do specifikovaného intervalu a kombinací více atributů mohou být vytvořeny nové, kterými se nahradí ty stávající [33].

Integrace dat

Pro lepší výsledky při zpracování dat je běžně použito více zdrojů dat. Může se jednat například o propojení záznamů dopravních nehod s daty počasí nebo dopravních kamer [31]. Možné je integrovat i stejný typ dat z více zdrojů, v tomto případě ale může dojít ke konfliktům. Může se jednat o konflikt schématu, při kterém záznamy různých zdrojů obsahují

odlišné atributy pro popis stejných skutečností. Dalším problémem může být konflikt hodnot, při kterém se hodnoty zdroje nachází mimo očekávanou doménu. Posledním typem je konflikt identifikace, při kterém jsou záznamy, které popisují stejnou událost, popsány odlišným identifikátorem. Zdroje mohou navíc obsahovat odlišně pojmenované atributy se stejným významem nebo atributy, které je možné odvodit z ostatních hodnot. Takovéto atributy jsou redundantní a je vhodné je odstranit [33].

Redukce dat

Zdroje dat mohou obsahovat velké množství údajů, které může být navíc rozšířeno integrací více různých zdrojů. Velké objemy dat ale nejsou ideální z pohledu potřebné výpočetní síly pro jejich zpracování. Aby v takových situacích bylo možné s daty dále pracovat v rozumné kapacitě, je nutné zredukovat jejich množství. Z tohoto důvodu se provádí následující redukce dimenzionality dat [27]:

- *Feature selection* označuje proces, jehož cílem je identifikovat podmnožinu potřebných atributů a ostatní zahodit. Získaná podmnožina by měla obsahovat všechny potřebné informace pro další zpracování. Tento proces je možné provést i před ostatními operacemi předpracování, čímž se dosáhne zrychlení i v tomto ohledu.
- *Space transformations* se na rozdíl od předešlé techniky snaží vytvořit z dostupných atributů atributy nové. Ty pak dokážou reprezentovat původní data pomocí menšího počtu atributů.

Mimo redukcí dimenzionality je možné provést i následující redukce instancí:

- *Instance selection* spočívá ve výběru minimální podmnožiny všech záznamů, která je schopna reprezentovat celek. Toto mimo jiné odstraní záznamy s odlehlými hodnotami a umožní práci s velkým množstvím dat.
- *Instance generation* kromě selekce podmnožiny dokáže i generovat data, kterými nahradí data původní. Výsledkem je opět redukovaný počet záznamů, který reprezentuje původní data.

Diskretizace umožňuje transformovat kvantitativní data na kvalitativní rozdělením hodnot do nepřekrývajících se intervalů. Toto může být vhodné, pokud další zpracování vyžaduje kvalitativní hodnoty. Výsledná data jsou jednodušší na pochopení člověkem a ulehčují další zpracování. Nevýhodou této transformace je částečná ztráta informací. Aby byla tato ztráta co nejnižší je třeba využít co nejoptimálnější algoritmus pro diskretizaci [27].

2.4 Vizualizace dopravních dat

Vizualizace dopravních dat je stále důležitější součástí systémů pro řízení dopravy. Pomáhá prezentovat a pochopit trajektorii vozidel a umožňuje nalézt dopravní, sociální a geografické vzory, které by jinak mohly zůstat skryté [14]. Podle [4] a [14] jsou definovány následující čtyři kategorie vizualizace dopravních dat podle zobrazených atributů.

Vizualizace prostorových údajů

Vizualizace prostorových údajů umožňuje zobrazit záznamy v mapě, identifikovat problematická místa nebo odhalit vzory mezi oblastmi. Spojením více prostorových záznamů

jednoho objektu je možné zobrazit jeho trajektorii [14]. Tento typ vizualizace se podle míry agregace dále dělí na [14]:

- **Vizualizace založené na bodech** zobrazují každý záznam samostatně bez jakékoliv agregace. Každý bod může mít odlišnou velikost a barvu podle hodnot záznamu, který je s ním spojen. Toto umožňuje například sledovat pozice dopravních prostředků v čase [14] nebo může pomoci odhalit místa častých dopravních nehod [4].
- **Teplotní mapy** se využívají v případech, kdy existuje příliš velký počet záznamů na to, aby je bylo vhodné zobrazit jako samostatné body. Pomocí tohoto zobrazení je možné přehledněji zobrazit hustotu bodů v různých oblastech. Tvorbu této mapy obvykle zajišťuje algoritmus *Kernel density estimation* (KDE). Zdroj [14] tento typ bere jako součást bodové vizualizace.
- **Liniové vizualizace** se využívají k zobrazení trajektorií a toku dopravy. Tento typ vizualizace nachází využití při doporučování tras nebo shlukování trajektorií. Pro zpřehlednění se při velkém počtu propojených lokalit využívá takzvaných *edge bundling* algoritmů. Nevýhodou je možná ztráta informací o směru trajektorie lokalitami [14].
- **Vizualizace založené na regionech** představují nejvyšší míru agregace. Toto jim umožňuje porovnat údaje městských částí [4], toky dopravy na silnicích, trajektorie mezi regiony [14] nebo odhalit další vzory na úrovni regionů.

Vizualizace časových údajů

Vizualizace časových údajů umožňují sledovat periodicky opakující se trendy nebo výskyty neobvyklostí [4]. Způsob vizualizace jde dále dělit podle účelu na [14]:

- **Vizualizace lineárního času** jsou definovány počátečním a koncovým bodem. Toto umožňuje sledovat vývoj hodnot a odhalit neobvyklé výkyvy. Běžně se pro tento účel používají spojnicové nebo sloupcové grafy.
- **Vizualizace periodického času** se soustředí na zachycení jevů, které se periodicky opakují. Tímto způsobem je například možné reprezentovat hustotu dopravy nebo počet nehod pro každý den v týdnu v určitý čas. Následně je možné určit, které části týdne jsou z těchto pohledů nejproblematictější. Pro vizualizaci periodických jevů je možné použít radiální graf.
- **Vizualizace větvičího se času** popisují strukturu vzniku, spojení a zániku událostí. Tento způsob se pro vizualizaci dopravních dat nevyužívá.

Vizualizace prostorových a časových údajů

Kombinace prostorových a časových údajů umožňuje porovnávat vývoj situace na daném místě v čase. Toto pomáhá rozlišit konstantně problematická místa s těmi, kde dochází k nehodám sporadicky [4]. V [14] je zmíněna metoda *Space-Time-Cube*, která umožňuje zobrazit trajektorii ve 3D prostoru pomocí souřadnic X, Y a Z. V rovině XY jsou zobrazena prostorová data, osa z pak popisuje jejich změnu v čase. Výsledkem je kostka, která popisuje celý průběh trajektorie objektu.

Vizualizace v kombinaci s ostatními atributy

Toto zobrazení se využívá v případech, kdy je cílem najít korelaci mezi danými atributy [4]. Zdroj [14] rozděluje zobrazované atributy do tří kategorií:

- **Kvantitativní atributy** reprezentují hodnoty jako například rychlost, zrychlení nebo váhu. Takovéto hodnoty je vhodné zobrazit například histogramem.
- **Kategorické atributy** popisují nějaký stav, ve kterém se objekt nachází. Příkladem takových atributů může být například typ nehody nebo vozidla. Na mapě jsou odlišné objekty barevně rozlišovány namapováním hodnot na specifické barvy.
- **Textové atributy** obsahují doplňující informace k zobrazeným záznamům. Tyto údaje jsou pak zakomponovány přímo do samotné mapy.

Při tomto typu vizualizace je nutné brát v potaz skutečnost, že větší množství zobrazených informací může pro pozorovatele představovat komplikace při snaze pochopit prezentovaná data a může je zmást [4].

2.5 Geografické informační systémy

Geografické informační systémy (GIS) představují nástroj pro tvorbu, vizualizaci a analýzu digitálních map. Mapy zobrazují prostorová data, která jsou obsažena v mapových vrstvách. Každá z těchto vrstev obvykle reprezentuje množinu nějakých prvků reálného světa. Tyto prvky se označují anglicky jako takzvané *features* a klasifikují se jako vektorová data. Druhým typem jsou rastrová data, která se využívají pro popis stále měnících se informací [53]. GIS nalézají využití v širokém spektru oborů jako například biologii, geografii, správě a plánování využití zdrojů [62], dopravě nebo při boji s přírodními katastrofami [53].

Prostorové referenční systémy

Prostorové hodnoty dopravních dat se odkazují na určité místo v reálném světě. Protože je ale cílem s daty pracovat v elektronické podobě, je nutné je převést do 2D. Pro tento účel se využívá prostorových referenčních systémů, které tuto projekci popisují. Bohužel neexistuje takový systém, který by dokázal identicky reprezentovat Zemi nebo alespoň její části. Vždy jsou tak přítomny některé nepřesnosti v podobě zkreslených úhlů, vzdáleností nebo oblastí. Každý systém balancuje poměry těchto nepřesností odlišně, a tak se různí jejich využití. Pomocí Mercatorova zobrazení je například možné správně zobrazit světové směry, což se využívá pro navigaci. Nevýhodou jsou ale nepřesné projekce oblastí a vzdáleností [53].

Prostorové referenční systémy se dále dělí na geografické referenční systémy a projektované souřadnicové systémy [53]:

- **Geografické referenční systémy** používají pro popis míst na Zemi zeměpisnou šířku a délku. Zeměpisná šířka rozděluje zeměkouli pomocí 180 rovnoběžek, které od sebe mají stejnou vzdálenost šedesáti námořních mil. Hodnoty severní polokoule se pohybují v intervalu 0 až 90 stupňů, pro jižní polokouli jde o hodnoty 0 až -90 stupňů. Zeměpisná délka rozděluje zeměkouli pomocí poledníků v rozmezí -180 až 180 stupňů. Vzdálenost mezi nimi se ale směrem od rovníku k pólům zmenšuje, až se v něm nakonec všechny protnou. Dostatečně správná reprezentace je tak možná jen při zobrazeních menších částí celé Země. Nejvyužívanějším systémem pro popis míst na Zemi podle zeměpisné délky a šířky je WGS 84.

- **Projektované referenční systémy** popisují místa na Zemi pomocí souřadnic X, Y a případně i Z. Hodnoty Y se zvyšují směrem od rovníku k pólům. Hodnoty X se na severní polokouli zvyšují směrem od nultého poledníku na východ a na jižní směrem na západ.

Vektorová data

Vektorová data slouží pro reprezentaci prvků reálného světa jako například stromů, domů nebo silnic. Tyto prvky obsahují numerické nebo textové atributy, které poskytují doplňující informace. Vektorový prvek je popsán v prostoru pomocí geometrie, která se skládá z jednoho nebo více vrcholů (ang. *vertices*). Každý z těchto vrcholů je popsán souřadnicemi X a Y. Pokud je cílem zachytit také výšku nebo hloubku prvku je použita i souřadnice Z. Prvky mohou být reprezentovány jedním z následujících tří typů geometrie [53]:

- **Bod** je popsán pomocí jednoho vrcholu. Může být například využit pro popis stromů nebo měst při větším oddálení mapy.
- **Křivka** se skládá ze dvou nebo více vrcholů. Tato geometrie nachází využití při popisu silnic, řek nebo leteckých trajektorií.
- **Polygon** popisují uzavřenou oblast, která je definovaná pomocí tří nebo více vrcholů, přičemž první a poslední vrchol jsou na stejném místě.

V GIS se tyto prvky obvykle seskupují do vrstev podle typu geometrie a podle prvků, které popisují. Při nepřesných souřadnicích může dojít k nechtěnému překrytí geometrie nebo naopak vytvoření nevyplněného prostoru [53].

GeoJSON

Jedním z formátů využívaných pro ukládání vektorových dat je GeoJSON. Tento formát ukládá geografická data pomocí struktury JSON. Obsah těchto souborů se skládá z typů *Feature* a *FeatureCollection*. *Feature* popisuje objekt, který obsahuje jednu z povolených sedmi typů geometrie [10]:

- **Point** popisuje polohu jednoho bodu pomocí jedné hodnoty souřadnic.
- **MultiPoint** popisuje polohu více bodů pomocí seznamu souřadnic.
- **LineString** popisuje polohu křivky pomocí seznamu dvou nebo více souřadnic.
- **MultiLineString** popisuje polohu více křivek pomocí seznamu *LineString* souřadnic.
- **Polygon** popisuje oblast pomocí čtyř nebo více souřadnic. První a poslední souřadnice musí mít stejnou hodnotu. Po směru hodinových ručiček se zapisují oblasti polygonu a proti směru jeho díry.
- **MultiPolygon** popisuje více polygonů pomocí seznamu *Polygon* souřadnic.
- **GeometryCollection** umožňuje popsat více různých druhů geometrie současně.

Hodnota atributu *coordinates*, který popisuje souřadnice, je seznam dvou až tří čísel, ve kterém první číslo popisuje zeměpisnou délku a druhé zeměpisnou šířku. Nepovinná třetí hodnota je určena pro popis nadmořské výšky. *FeatureCollection* pak obsahuje seznam *Feature* objektů [10]. Ukázku souboru GeoJSON, ve kterém je uložen objekt bodu, křivky a polygonu je možné vidět ve výpisu 2.1.

```

{
  "type": "FeatureCollection",
  "features": [{
    "type": "Feature",
    "geometry": {
      "type": "Point",
      "coordinates": [102.0, 0.5]
    },
    "properties": { ... }
  }, {
    "type": "Feature",
    "geometry": {
      "type": "LineString",
      "coordinates": [
        [102.0, 0.0],
        [103.0, 1.0],
        [104.0, 0.0],
        [105.0, 1.0]
      ]
    },
    "properties": { ... }
  }, {
    "type": "Feature",
    "geometry": {
      "type": "Polygon",
      "coordinates": [
        [
          [100.0, 0.0],
          [101.0, 0.0],
          [101.0, 1.0],
          [100.0, 1.0],
          [100.0, 0.0]
        ]
      ]
    },
    "properties": { ... }
  }]
}

```

Výpis 2.1: GeoJSON soubor obsahující objekty bodu, křivky a polygonu [10].

Rastrová data

Rastrová data se využívají v případech, kdy je cílem zobrazit větší oblasti, ve kterých se často mění hodnoty atributů dat. Tyto změny nejsou vektorové prvky schopny zachytit, a tak pro reprezentaci takových dat nejsou vhodné. Rastrová data jsou popsána maticí pixelů, které dohromady tvoří rastr. Každý z těchto pixelů popisuje stav v dané oblasti, kterou pokrývá. Tento typ dat je vhodný například pro reprezentaci hustoty porostu luk,

rizika ohně, teplotních trendů nebo intenzity srážek. Často se také využívají pro podkladové mapy, které pomáhají vektorová data lépe zasadit do kontextu dané oblasti.

Rastrová data jsou nejčastěji získávána z leteckých a satelitních snímků. Pixely těchto snímků mohou mít různé prostorové rozlišení (ang. *spatial resolution*), což umožní různé úrovně přiblížení mapy. Velikosti stran pixelů se mohou pohybovat v hodnotách od desítek centimetrů do stovek metrů. Menší rozlišení umožní zachycení vyšší úrovně detailu, avšak za cenu nárůstu celkové velikosti snímků [53].

GeoTIFF

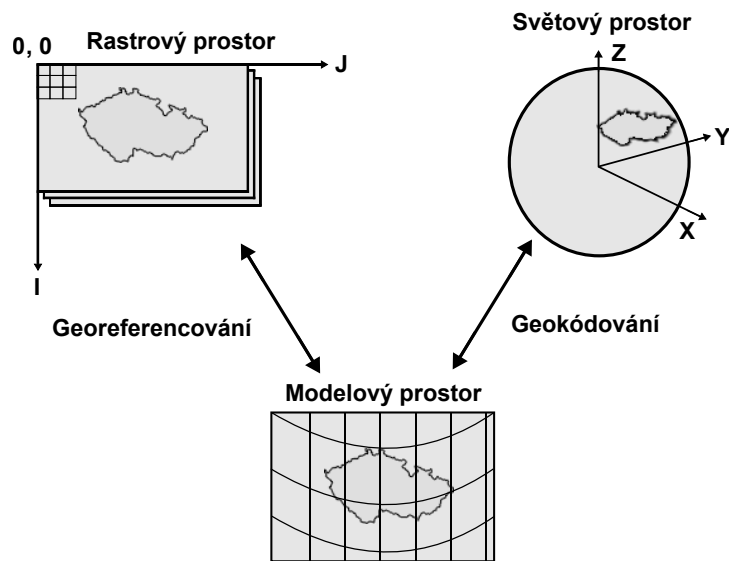
Rastrový formát *Tag Image File Format* (TIFF) byl vytvořen v osmdesátých letech minulého století pro práci s obrazovými soubory. Jednalo se o univerzální formát využívaný skenery, což do té doby nebylo běžné. Aktuálně je využíván pro ukládání obrázků a skenů ve vysoké kvalitě. To mu umožňuje bezztrátová komprese, po které zůstanou zachované detaily a barevná hloubka původního snímku [1].

S rostoucí popularitou mapovacích systémů začala narůstat i poptávka po formátu, který by umožnil propojit letecké, satelitní a další typy snímků s jejich přesnou polohou na Zemi. Tento proces se do té doby prováděl ručně, což bylo pomalé a vedlo k chybám nebo pomocí specializovaných programů, což vyžadovalo uživatelskou znalost. Z těchto důvodů byl vytvořen formát GeoTIFF, který toto propojení snímku s lokalitou zajišťuje [54].

Obsah souborů GeoTIFF rozšiřuje původní TIFF formát a zachovává jeho strukturu. GeoTIFF rozeznává tři souřadnicové systémy:

- Rastrový prostor (ang. *raster space*) popisující matici pixelů, které jsou popsány pomocí čísel řádku a sloupce.
- Modelový prostor (ang. *model space*) popisující 2D projekci povrchu Země pomocí zeměpisné délky a šířky.
- Světový prostor (ang. *world space*) popisující povrch Země pomocí souřadnic X, Y a Z.

Mezi rastrovým a modelovým prostorem se data převádí pomocí georeferencování a mezi modelovým a světovým prostorem pomocí geokódování. Při georeferencování je potřeba vytvořit takzvané kontrolní body v místech, které je možné přesně identifikovat jak v rastrovém, tak v modelovém prostoru. Pro každý z těchto bodů je třeba vybrat referenční souřadnicový systém a zadat pozici bodu. Na základě vybraných bodů a systému se provede transformace dat [20]. Geokódování označuje proces transformace lokality popsané pomocí souřadnic, adresy nebo jiným způsobem na místo na povrchu Země [22]. Aby tak bylo možné přiřadit data snímku odpovídajícímu místu na Zemi, je třeba provést oba převody. Schematická ukázka tohoto procesu je vidět na obrázku 2.2. Informace o souřadnicovém systému, poloze snímku v něm a další jsou uloženy v metadatech souboru [54].



Obrázek 2.2: Převod dat mezi rastrovým, modelovým a světovým prostorem pomocí georeferencování a geokódování. Ilustrace inspirována zdrojem [54].

2.6 Vyhodnocování dopravních dat pomocí GIS

GIS hrají významnou roli při analýze dopravních dat, bezpečnosti na silnicích, modelování dopravních toků a městského plánování [51]. Data je možné analyzovat na základě času, místa nebo jejich kombinace. Analýza na základě času se zaměřuje na počty výskytů určitých jevů v čase. Může jít například o množství dopravních nehod v různých částech dne. Prostorová analýza se zaměřuje na konkrétní místo jako například silnici nebo křižovatku. Jejich kombinace umožňuje odhalovat vzory na sledovaných komunikacích v čase [2].

Výhodou GIS je i možnost integrovat více zdrojů zároveň. Při analýze tak mohou být kromě dopravních dat brány v potaz i například data městské hromadné dopravy, lokality škol, demografická údaje nebo data o kvalitě ovzduší. Tyto doplňující informace umožní odhalit spojitosti a vzory, které by jinak nemusely být patrné. Zdroj [51] rozděluje analýzu dopravních dat do dvou kategorií. První je modelování dopravních dat, které má následující možnosti využití:

- Analýza dopravních toků, hledání problémových míst s vysokou hustotou dopravy a simulace potenciálních úprav stávající infrastruktury.
- Předvídání vytíženosti dílčích částí dopravní infrastruktury na základě údajů o obyvatelstvu a městské hromadné dopravě.
- Plánování efektivních tras městské hromadné dopravy a četnost jejich spojů.

Druhou je analýza dopravních nehod, která se jejich výskyt snaží omezit pomocí:

- Identifikace problematických míst, kde často dochází k dopravním nehodám. Pro tento účel je možné využít například algoritmu KDE nebo jiných shlukovacích metod.
- Srovnáním dopravních opatření na vybraných komunikacích a označení potenciálně nebezpečných úseků z pohledu nehodovosti.

V [2] jsou pro identifikaci nebezpečných částí silnic zmíněny následující metody:

- *Frequency Method* je založena na žebříčku bezpečnosti silnic v daných oblastech podle World Road Association. Jde o běžně využívanou metodu pro určení bezpečnosti daného místa. Nebere však v potaz různou hustotu dopravy, která má na nehodovost velký vliv.
- *Traffic Accident Rate Method* ohodnocuje různé oblasti podle počtu dopravních nehod na hustotu provozu nebo na vozidly najetou vzdálenost. Oblasti, které dosahují vyšších než očekávaných hodnot, jsou označeny za nebezpečné.
- *Combined Criteria Method* provádí kvantitativní analýzu a hledá místa s vyšší nehodovostí, než je průměr pro danou oblast. Po identifikaci problematických míst je využita odlišná metoda, aby se předešlo chybnému označení. Tento přístup by měl odstranit nebo alespoň zmírnit chyby způsobené limitacemi využitých metod.

GIS tak umožňují městům lépe plánovat vhodné úpravy dopravní infrastruktury, aby se na silnicích zvyšovala úroveň bezpečnosti a snižoval se počet míst s příliš vysokou hustotou dopravy.

Kapitola 3

Zpracování a vizualizace dat na webu

Práce s daty na webu představuje alternativu k tradičně lokálně nainstalovaným programům. Uživatelům odpadá nutnost program stahovat a lokálně instalovat, což snižuje bariéru pro vstup a zároveň umožňuje jeho využití i na dalších typech zařízení [56]. Webové aplikace navíc umožňují vývojářům provádět rychlé aktualizace, které pro koncového uživatele nemusí být vůbec znatelné.

I přestože že se jedná o populární způsob vizualizace dat, stále je nutné překonat několik problémů, které s sebou přináší. Především se jedná o nároky na výpočetní sílu, omezení velikosti uložených dat, rychlost odezvy na uživatelské akce a rychlost přenosu dat. Snahou je tyto problematické části stále vylepšovat a redukovat jejich negativní dopad, aby bylo možné pracovat s větším množstvím dat [25].

Vykreslování vizualizací na webu je běžně zajištěno pomocí Canvas API, SVG nebo WebGL. Každý z těchto způsobů vykreslování má své výhody a omezení, a je tak vhodný pro odlišné případy využití. Vzhledem k jejich rozšířenosti pro tvorbu 2D a 3D grafiky jsou tyto způsoby podporovány napříč všemi aktuálními verzemi nejpoužívanějších webových prohlížečů [45, 46, 47].

3.1 Vizualizace dat

Počátky vizualizace dat sahají až do druhého století našeho letopočtu, kdy sloužila pro zachycení historických událostí a popis známého světa [40]. Později vizualizace pomáhaly s navigací a objevováním neznámých území [13]. V dnešní době je vizualizace klíčovým nástrojem pro efektivní analýzu a pochopení velkých objemů dat. Člověku pomáhá odhalit skryté vzory a vztahy mezi daty a zároveň mu umožňuje identifikovat neobvyklé odchylky [40].

Vizualizaci dat je možné definovat jako prezentaci dat v obrazovém nebo grafickém formátu, která je vygenerována softwarovým nástrojem. Vizualizace umožňuje uživatelům interaktivně prozkoumávat a analyzovat data, čímž jim umožňuje identifikovat zajímavé vzory, korelace a příčiny [8]. Definice vizualizace dat není pevně stanovena a je tak možné najít i odlišné pohledy ostatních autorů [40].

Data, informace a znalosti

Pojmy jako data, informace a znalosti jsou zásadní součástí jakéhokoliv zpracování a vizualizace dat, i přesto pro jejich popis existuje více různých definic. Zdroj [7] data popisuje jako nezpracované hodnoty, které nemají mimo svoji existenci žádný význam. Dále zmiňuje, že data mohou existovat v jakémkoliv podobě. V těchto podobách mohou být data využitelná nebo nevyužitelná. Zdroj [40] data definuje jako nezpracované informace, které jsou uloženy ve formě čísel a binárního kódu.

Informace jsou podle [40] zpracovaná data, které mají pro člověka nějaký pochopitelný význam. Zdroj [7] informace označuje jako data, která získala význam relačním vztahem. Takto získané informace pak mohou nebo nemusí být užitečné. Informace tak oproti datům, které reprezentují pouze nějaký fakt, představují vztah, který například vyjadřuje příčinu a následek.

Podle [7] je kolekci užitečných informací možné označit jako znalost. Znalosti tak reprezentují určitý vzor, pomocí kterého je možné předpovídat nadcházející jevy. Získání nových a užitečných znalostí je hlavním cílem celého procesu zpracování a vizualizace dat.

Druhy vizualizací

Vizualizace je možné dělit na informační a vědecké. V případě informačních vizualizací se jedná o abstraktní entity jako jsou například business data, která popisují podniky a jejich chod [40]. Tyto věci nemají a nemohou mít přímý vztah s prvky fyzického světa a jejich význam musí být analyzován [13]. Cílem je efektivně prezentovat vybrané informace, které jsou přesné a nezaujaté, aby se na jejich základě mohli provádět rozhodnutí. Mezi způsoby vizualizace těchto informací patří tabulky, stromové vizualizace, mapy, korelační diagramy, diagramy nebo grafy [40]:

- **Tabulky** se skládají z řádků, které popisují záznamy dat a sloupců, které popisují jejich dimenze. Tabulky jsou také schopny reprezentovat hodnoty dat přesněji, než ostatní vizualizace a umožňují vizualizovat i takové hodnoty, které by pro ostatní druhy nebyly vhodné. Využití často nachází v podobě jízdních řádů a obdobných časových rozvrhů.
- **Stromové vizualizace** se skládají ze sítě uzlů. Ty jsou postupně seřazeny od kořenového přes vnitřní až po listové uzly. Kromě samotných dat je tak zobrazena i jejich hierarchie. Tento typ vizualizace umožňuje dohledat uzly na cestě od kořenového k cílovému uzlu. V dnešní době nachází tento typ využití například v investiční sféře, kde investorům pomáhá identifikovat vzory a tendence částí burzy.
- **Mapy** představují způsob vizualizace vztahů mezi místy v prostoru. Nejčastěji jsou využívány k vizualizaci silničních sítí, leteckých a námořních tras, sítí veřejné dopravy nebo turistických tras. S mapami je možné interagovat změnou přiblížení, jejich posunem nebo vyhledáním přesného místa.
- **Korelační diagramy** jsou často využívány pro zobrazení vícerozměrných dat ve dvou dimenzionálním prostoru. Tento typ zobrazení může mít významný pozitivní vliv na pochopení prezentovaných informací. Jednorozměrná verze korelačního diagramu může být využita například pro vizualizaci výsledků studií nebo hodnocení filmů.
- **Diagramy** popisují vztahy toku informací pomocí tvarů a jejich propojení. Alternativně mohou zobrazovat kvantitativní data například pomocí sloupcových grafů.

- **Grafy** slouží k podobnému účelu jako diagramy. Data propojují vzájemnými vztahy, což může být vhodné pro zkoumání jejich rysů. Nedokážou však přesně reprezentovat jejich samotné hodnoty.

Mapy je dále možné dělit do podkategorií podle specifického typu zobrazení [57]:

- **Kartogramy** popisují naměřené hodnoty v ohraničených oblastech pomocí barev, odstínů a vzorů. Využití nachází například při vizualizaci hustoty populace, hrubého domácího produktu nebo výsledků voleb.
- **Teplotní mapy** reprezentují hodnoty v různých barvách a intenzitách. Oproti diskrétním změnám barev oblastí kartogramu jsou změny mezi těmito hodnotami spojitě.
- **Bodové mapy** reprezentují údaje z oblastí pomocí bodů. Body jsou identické a bližší informace o oblasti, kterou reprezentují, je možné nalézt v dialogu daného bodu.
- **Hexagonální rozdělení** reprezentuje hodnoty oblastí pomocí šestiúhelníků. Výsledek je podobný kartogramu s tím rozdílem, že všechny oblasti jsou stejně velké.
- **Bublinové mapy** reprezentují hodnoty stejně jako bodové mapy s tím rozdílem, že body mění svoji velikost podle jejich hodnot. Body s vysokými hodnotami tak zabírají daleko více místa na mapě.

Vědecké vizualizace jsou založené na entitách fyzicky existujících v reálném světě a popisují například lidské tělo [40], meteorologické jevy nebo molekuly prvků [25]. Data mohou být reprezentována pomocí simulací, průběhů signálu nebo objemového vykreslování [40]:

- **Simulace** jsou využívány pro vizualizaci přírodních jevů v čase. Využití nalézají například při vizualizaci počasí nebo popisu teplot pomocí teplotních map.
- **Průběh signálu** se využívá pro reprezentaci signálů, kterými je možné popsat množství dat v určitém čase. Tento typ vizualizace je možné najít ve zdravotnictví pro sledování krevního tepu nebo například při sledování otřesů země.
- **Objemové vykreslování** projektuje a prezentuje data získaná ve tří dimenzionálním prostoru do prostoru dvou dimenzionálního. Touto metodou se vykreslují například skeny magnetické rezonance.

3.2 Vizualizace vícedimenzionálních dat

Data, která jsou popsána více jak jedním atributem se nazývají vícedimenzionální. V případě, že je těchto atributů pevně daný počet, je použít označení *n-dimenzionální*, kde je *n* nahrazeno počtem atributů. Vícedimenzionální data však představují značný problém pro lidské vnímání. To je schopné pochopit běžně používané dvě nebo tři dimenze a vyšší počty už představují problém. Pro vizualizaci vícedimenzionálních dat tak existují metody přímé vizualizace a metody projekce, které se tento problém pokouší řešit.

V následujícím textu je pro souhrnné označení záznamů, pozorování a instancí použito slovo objekt. Objekty jsou popsány pomocí rysů, které souhrnně označují atributy, parametry a vlastnosti. Zdrojem pro tyto odstavce a následující podsekcce je [18].

Metody přímé vizualizace

Metody přímé vizualizace zobrazují všechny dimenze původních dat pomocí specializovaných vizualizací. Ty je dále možné rozdělit na geometrické, ikonografické a hierarchické. Každá z těchto podkategorií obsahuje větší množství technik vizualizace a z důvodu zaměření této práce z nich budou zmíněny pouze vybraní zástupci.

Jako zástupce geometrických metod je možné zmínit korelační graf. Ten je sice běžně využíván pro prezentaci dvou nebo tří dimenzionálních dat, avšak kombinací více korelačních grafů je možné vytvořit jejich matici, která je schopná zobrazit i vícedimenzionální data. Pro zobrazení n dimenzí je potřeba vytvořit $n(n-1)/2$ korelačních grafů. Diagonála matice pak může být vyplněna například histogramem, který popisuje rozdělení hodnot dané dimenze.

Ikonografické zobrazení se snaží ulehčit čtenářům pochopení pomocí symbolů, které mění tvar na základě hodnot dat. Zástupcem je například vizualizace pomocí takzvaných *star glyphs*, ve kterých je každý objekt zobrazen jako n -úhelník, kde n je počet dimenzí. Vzdálenost vrcholů od středu n -úhelníku se odvíjí od hodnoty daného rysu. Výsledné zobrazení je pak matice těchto objektů, prezentovaných pomocí n -úhelníků.

Jako zástupce hierarchických zobrazení je možné zmínit metodu *dimensional stacking*. Ta skládá dimenze do stromových struktur pro osy XY. Venkovní dimenze je rozdělena na několik intervalů, jejichž hodnot může nabývat. Pro každý interval je vložena další dimenze, která je opět rozdělena na několik intervalů. Takto se proces opakuje, dokud nejsou přidány všechny dimenze. Výsledkem je matice, která obsahuje počet řádků a sloupců rovný počtu intervalů nejvíce zanořených dimenzí.

Výhodou těchto metod je skutečnost, že dokážou zobrazit všechny dimenze vícedimenzionálních dat. Značnou nevýhodou je ale to, že s rostoucím počtem dimenzí nemusí být možné rozeznat detaily výsledných vizualizací. Zároveň je pro pochopení prezentovaných dat potřeba jistá úroveň uživatelské zkušenosti.

Metody projekce

Metody projekce umožňují reprezentovat vícedimenzionální data pomocí menšího počtu dimenzí. Tento proces možné nalézt i pod jménem redukce dimenzí. Tyto transformace jsou prováděny takovým způsobem, aby byly co nejvíce zachovány vzájemné vzdálenosti objektů a jejich pozice v prostoru. Pro transformaci je nutné zvolit vhodnou projekci, podle které budou vytvořeny shluky hodnot. Ve většině případů se provádí transformace vícedimenzionálních dat do dvou nebo tří dimenzí. Aby došlo k co nejmenšímu zkreslení původních dat, jsou použity matematická kritéria, pomocí kterých je určena optimální projekce. Projekce je možné rozdělit podle způsobu transformace dat na lineární a nelineární.

Lineární transformace lze popsat pomocí rovnice

$$Y_i = X_i A,$$

kde X_i je původní objekt, A je transformační matice a Y_i je výsledná projekce objektu. Počet dimenzí výsledné projekce se odvíjí od počtu sloupců transformační matice.

Nelineární transformace lze popsat pomocí rovnice

$$Y = f(X),$$

kde f je nelineární transformační funkce, X je množina původních objektů a Y označuje výslednou projekci objektů. Jde o náročnější operaci oproti lineární transformaci, ale při

projekci do menšího počtu dimenzí dokáže lépe zachovat vztahy mezi objekty. Pro zachycení a následné porovnání vzájemné vzdálenosti objektů se často používá Euklidovská vzdálenost. Alternativně je možné využít například Spearmanův korelační koeficient, pomocí kterého je možné odhadnout zachování polohy objektů.

Analýza hlavních komponent

Analýza hlavních komponent (ang. *principal component analysis*) je hojně rozšířená technika pro lineární transformaci vícedimenzionálních dat do menšího počtu dimenzí. Transformované objekty musí mít nejdůležitější informace v několika prvních dimenzích, které jsou označeny jako hlavní komponenty. Při transformaci dat do n -dimenzionálního prostoru se ponechá jen prvních n dimenzí a ostatní se zahodí. Tímto se přijde pouze o méně důležité informace, které mají na polohu objektu menší vliv.

Před transformací je nutné data analyzovat a vytvořit pořadí v jakém budou dimenze seřazeny. První hlavní komponentou je ta dimenze, ve které se určí směr s nejvyšší mírou rozptylu objektů. Směr následující hlavní komponenty je na tento kolmý a obsahuje nižší míru rozptylu objektů než předešlá komponenta. Analýza se provádí pomocí korelační nebo kovarianční matice, ze které se získají vlastní vektory E a odpovídající vlastní čísla λ . Vlastní vektory E_k , $k = \{1, \dots, n\}$ jsou potom seřazeny sestupně podle vlastních čísel ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$). V tomto pořadí jsou následně přidány sloupce, ze kterých se skládá matice hlavních komponentů:

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Při transformaci do n -dimenzionálního prostoru je využito jen prvních n vlastních vektorů. Metoda je rozšířená hlavně kvůli své jednoduchosti, avšak není vhodná pro transformaci nelineárních dat.

3.3 Technologie pro vizualizaci dat

V [43] jsou uvedeny hlavní způsoby, pomocí kterých je možné provádět vizualizace na webu. Jedná se o Canvas, SVG, CSS a animace pomocí přímých změn DOM. V případě statických vizualizací a 3D transformací mohou být dostačující CSS a DOM animace. Pro složitější grafiku je vhodnější volbou Canvas nebo SVG. Tyto technologie mají své výhody a nevýhody, a jsou tak vhodné pro odlišné situace [43, 5].

Canvas

Canvas API umožňuje tvorbu rastrové grafiky na webových stránkách pomocí jazyku JavaScript a HTML elementu *canvas*. Metody Canvas API umožňují vykreslovat obsah elementu canvas pro tvorbu 2D animací, herní grafiky nebo například pro zpracování videa v reálném čase. Pro složitější grafiku je určeno WebGL API, které nejenže umožňuje vytvářet 3D reprezentace, ale také využívá hardwarové akcelerace GPU [47, 43].

WebGL API je blízce spojené s technologií OpenGL, která představuje široce využívaný standard pro práci s 3D grafikou [43]. Práce s WebGL je nízkoúrovňová, protože poskytuje pouze vykreslovací jádro, kterému je třeba dodat vlastní kód, pomocí kterého budou požadované tvary vykresleny. Aby tak bylo možné vizualizace vytvářet, je nutné vytvořit vlastní vertex a fragment shader. Vertex shader slouží k výpočtu pozic vertexů. Fragment shader

je využít při výpočtu barevných hodnot pixelů bodů, čar a trojúhelníků [64]. Celkově se jedná o složitý systém, jehož implementace vyžaduje značné úsilí. Může tak být vhodné využít některé z knihoven pro práci s WebGL, která již vlastní shadery obsahuje [43].

Jedním z důležitých rozdílů oproti ostatním elementům *Document Object Model* (DOM) je skutečnost, že je obsah canvas elementu reprezentován jako jeden celek. Oproti běžné stromové struktuře DOM tak neexistují žádné uzly pro reprezentaci dílčích tvarů a všechny události jsou zasílány pouze elementu canvas. Jediným způsobem pro interakci s částmi obsahu je zmíněné API. To pomocí metody *isPointInPath* spočítá, jestli se souřadnice kurzoru nachází v oblasti daného tvaru [43]. Určitou nevýhodou pak mohou být horší možnosti testování, což může vést na větší množství neodhalených chyb [41].

SVG

Druhou možností pro tvorbu pokročilejší 2D grafiky je *Scalable Vector Graphics*. Zde je vektorová grafika reprezentována pomocí struktury, která vychází z jazyka XML. Grafiku je možné vytvářet přidáním některého z podporovaných elementů. Ty jsou poté přímou součástí struktury DOM, takže je možné se na ně odkazovat událostmi, což je oproti Canvas značné ulehčení [43]. Mimo tuto výhodu je navíc možné upravovat existující grafiku pomocí atributů a kaskádových stylů bez jakékoliv ztráty na kvalitě obrazu. Nevýhodou však může být složitá tvorba těchto elementů. Pro ulehčení této práce je možné využít některou z knihoven pro generaci SVG nebo některého z programů určeného pro tvorbu vektorové grafiky [46]. V případech, kdy je nutné zobrazit velké množství elementů, je doporučeno zvolit Canvas [5], který je dokáže rychleji a efektivněji zobrazit [34].

3.4 Knihovny pro vizualizaci dat

Protože může být tvorba vizualizací náročná z pohledu času, potřebných znalostí a množství potřebného kódu, je vhodné využít některou z knihoven pro vizualizaci dat. Každá knihovna nabízí odlišné možnosti přizpůsobitelnosti, což je také spojeno s tím o jak nízkouúrovňový přístup tvorby vizualizací jde. Příkladem knihoven odlišné úrovně může být nízkouúrovňová knihovna D3.js a její vysokoúrovňová alternativa Observable Plot. V rámci knihovny D3 má programátor kontrolu nad všemi aspekty tvořené vizualizace, avšak za cenu velkého množství kódu. Observable plot naopak nabízí možnost vytvořit vizualizace rychle a efektivně pomocí jednotek řádků kódu, avšak za cenu určité omezenosti při tvorbě vizualizací [16, 52].

D3.js

Data-driven documents nebo zkráceně D3 je jednou z nejvíce využívaných knihoven pro vizualizaci dat na webu. Jedná se o nízkouúrovňový přístup ke tvorbě vizualizací, který je spíše alternativou k vlastní implementaci než k ostatním vizualizačním knihovnám. Pomocí D3 nejsou přímo tvořeny hotové grafy a jiné zobrazení. Naopak se jedná o skládání více jednoduchých částí dohromady, pomocí kterých se požadovaný graf vytvoří. Například pro tvorbu grafu je tak nutné data načíst, vytvořit osy grafu, vytvořit z načtených dat požadovaný graf a přidat SVG do DOM. Rozsah kódu spojeného s tvorbou grafu se tak může běžně pohybovat ve stovkách řádků.

Součástí D3 je také takzvaná dynamická vizualizace, která přidává, odebrává a upravuje zobrazené elementy na základě vstupu od uživatele. Tyto změny jsou plynulé a rychlé,

protože se upravují pouze vybrané prvky a není tak nutné provádět porovnávání celého stavu DOM před a po změnách ani znovu vykreslovat veškerý obsah. Celkově tak knihovna D3 umožňuje tvorbu vizualizací pomocí SVG, Canvas a WebGL a společně s tím nabízí velké množství nástrojů, které mohou být využity i samostatně v rámci tvorby vlastních vizualizací [16].

Three.js

Knihovna Three.js je zaměřena na tvorbu 3D obsahu na webových stránkách pomocí WebGL. Jedná se o přístup na vyšší úrovni s již implementovanými částmi, které by jinak bylo nutné WebGL dodat. Vytvářené vizualizace se skládají z několika vzájemně propojených objektů. Hlavním objektem je renderer, který vykresluje části scény na základě aktuální pozice kamery. V rámci scény se nachází 3D objekty, zdroje světla a případně i kamera, která se ale může nacházet i samostatně mimo scénu. 3D objekty jsou definovány svojí geometrií a materiálem, kterým je popsán povrch objektu. K dispozici je i množství primitiv, což jsou 3D tvary, které je možné využít pro experimentování nebo jako dočasné zástupce za tvary vytvořené v modelovacích programech [61].

Další knihovny

Mimo zmíněné populární zástupce vizualizačních knihoven existuje velké množství alternativ, které je možné využít. Výběr knihovny se může odvíjet od licence, pod kterou je poskytována, zvolených technologií vykreslování, kompatibilních frameworků, nabízených nástrojů pro tvorbu grafiky nebo například rozšířenosti v komunitě vývojářů. Jako příklady knihoven pro tvorbu vizualizací je možné zmínit následující zástupce, jejichž nemalou popularitu je možné vidět na stránkách jejich repozitářů [15]:

- **Chart.js** je aktuálně nejpopulárnější volbou pro vykreslování grafů. Jedná se o vysokoúrovňový přístup, který umožňuje rychlou a nenáročnou tvorbu grafů. Chart.js vykresluje elementy pomocí canvas, na rozdíl od ostatních knihoven postavených na základech D3. Tímto je zajištěno rychlé vykreslování i při velkém objemu dat. Nevýhodou jsou omezenější možnosti stylování, které však mohou být rozšířeny pluginy [12].
- **Leaflet** umožňuje tvorbu interaktivních map, které fungují na počítačích i mobilních zařízeních. Podkladové mapy mohou být získány z mapových serverů jako například OpenStreetMap¹, avšak je nutné uvést jejich autorství. Pro vykreslování vrstev na mapě je možné zvolit SVG nebo canvas renderer. Základní funkcionalita může být rozšířena komunitními doplňky [36].
- **Babylon.js** je knihovna určená pro tvorbu 3D vizualizací. Primární využití nachází ve vývoji her [48]. Hlavním prvkem knihovny je Babylon.js Engine, který zajišťuje komunikaci s WebGL. Základními prvky pro tvorbu vizualizace jsou scéna, kamera, zdroj světla a zobrazované objekty. Pro vývoj nabízí Babylon nástroje pro ladění a sdílení kódu. Dále umožňuje tvorbu uživatelského rozhraní, shaderů a geometrií pomocí GUI, čímž se snaží redukovat množství potřebného kódu pro tvorbu vizualizací [6].

¹<https://www.openstreetmap.org/about>

Kapitola 4

Analýza aplikace Silniční doprava a skriptu pro aktualizaci dat

Aplikace Silniční doprava je součástí mapových aplikací města Most¹, ve kterých mohou obyvatelé ORP Most dohledat jindy hůře dostupné informace. Tyto aplikace jsou důležité pro informování veřejnosti a pro práci zaměstnanců magistrátu města Most. Zaměstnanci v nich mohou rychle dohledat požadované informace a provádět na jejich základě rozhodnutí. Příkladem dostupných mapových aplikací jsou Vyhlášky města Most, Územní plány ORP Most plánu nebo Školské obvody v Mostě. Aplikace z kategorie doprava se zabývají aktuální polohou městské hromadné dopravy, popisem dopravních nehod a měřením dopravních kamer. Z nich se posledními dvěma položkami zabývá aplikace Silniční doprava.

Aplikace však v aktuálním stavu postrádá pokročilejší možnosti vizualizace a neumožňuje detailnější analýzu. Nedostatky se současně vyskytují i ve skriptu pro aktualizaci dat, který je zásadní pro fungování aplikace. Skript není schopen v plné míře pracovat s aktuálními daty a jejich získávání navíc není automatizované. Uživatelé aplikace tak mají přístup pouze k omezeným informacím, které mohou být navíc zkreslené z důvodu jejich nekompletnosti. Získané informace v této podobě tak není možné spolehlivě využít.

V následující kapitole budou popsáni uživatelé aplikace, současná podoba aplikace a skriptu, požadavky na podobu nových programů a datové sady, jejichž data jsou vhodná pro zpracování a vizualizaci. Informace pro tuto kapitolu byly získány na průběžných konzultacích a skrze emailovou komunikaci s referentem města Most Ing. Kamilem Novákem.

4.1 Uživatelé aplikace

Uživatele aplikace Silniční doprava je možné zařadit do jedné z následujících skupin:

- **Široká veřejnost**, do které se řadí zejména obyvatelé ORP Most a jeho okolí, pro které mají tyto informace hodnotu. Pro tuto skupinu jsou důležité informace o nebezpečných úsecích a statistiky provozu na silnicích. Na základě těchto informací mohou uživatelé plánovat bezpečnější cesty a pozorovat případné efekty změn zavedených magistrátem. Z těchto případů využití vyplývají následující požadavky na aplikaci:
 - Aplikace by měla být veřejně přístupná v rámci mapových aplikací města Most.

¹Mapové aplikace města Most: <https://mapy.mesto-most.cz/portal/apps/sites/#/mapovy-portal/pages/mapove-aplikace>

- Vizualizována by měla být současná i historická data dopravních nehod a rychlostních kamer.
- Úseky častých dopravních nehod by měly být odlišeny od ostatních úseků.
- **Zaměstnanci magistrátu**, které je ještě možné rozdělit do dvou podskupin. Pro první skupinu jsou klíčové informace o dopravních nehodách a statistiky provozu podobně jako pro veřejnost. Navíc jsou potřeba nástroje pro analýzu dat, aby bylo možné identifikovat problematická místa a příčiny jejich vzniku. Pro pozorování efektů úprav a omezení jsou opět klíčová historická data a trendy sledovaných hodnot. Druhá skupina, která aplikaci a skript spravuje a udržuje, má hlavní požadavky na jednoduchou udržitelnost a flexibilitu pro různé typy dat. Zaměstnanci magistrátu tak navíc k předešlým požadavkům veřejnosti potřebují zajistit následující požadavky:
 - Aplikace by měla obsahovat možnost analýzy dopravních nehod na základě jejich parametrů.
 - Aplikace by měla obsahovat vizualizace trendů hodnot v rámci dopravních komunikací a vizualizace měření rychlostních kamer.
 - Aplikace by měla umožnit vizualizovat různé druhy dopravních dat.
 - Aplikace i skript by měly být konfigurovatelné bez nutnosti úpravy obsahu programů.
 - Skript by měl umožnit zpracování požadovaných dat s možností jednoduchých úprav kódu v případě změn.

4.2 Současná podoba webové aplikace a skriptu

Stav současného řešení byl diskutován s referentem a v případě skriptu byla sdílena i jeho nejnovější verze. V případě webové aplikace byla sdílena pouze její struktura a konfigurace. Zároveň byla také poskytnuta aktuálně využívaná geodatabáze dopravních nehod a rychlostních měření. V aktuálním řešení je skript automaticky jednou měsíčně spuštěn s manuálně dodanými daty, následně do geodatabáze přidá nové záznamy a webová aplikace aktualizované záznamy zobrazí uživatelům. Detailnější popis využitých technologií a dílčích částí je popsán v následujících podkapitolách.

Využité technologie

Před popisem dílčích nástrojů, je nutné zmínit využité technologie. Město Most využívá aplikace ArcGIS² pro správu dat a pro tvorbu a publikování mapových aplikací. Tyto aplikace jsou vyvíjeny společností Esri a jsou široce využívány na celosvětové úrovni [19]. Existuje více edicí těchto programů, ale obecně se jedná o GIS, které byly popsány v sekci 2.5. Pro jejich využití je však nutné vlastnit licenci. To stejné platí pro formáty, které tyto systémy využívají pro ukládání dat a pro knihovny, které s těmito softwary umožňují práci v různých programovacích jazycích.

Pro práci s daty je využita knihovna ArcPy³, která pomocí jazyka Python nabízí možnosti správy databáze a zpracování dat. Zpracovaná data jsou uchována v geodatabázi, což

²Stránky Esri ArcGIS: <https://www.esri.com/en-us/arcgis/geospatial-platform/overview>

³Knihovna ArcPy: <https://pro.arcgis.com/en/pro-app/latest/arcpy/get-started/what-is-arcpy-.htm>

je proprietární formát společnosti Esri pro ukládání a správu prostorových dat. V rámci geodatabáze jsou data uložena podle jejich povahy:

- Data bez prostorové složky jsou uchována v podobě tabulek.
- Data s prostorovou složkou jsou uchována jako třída prvků (ang. *feature class*⁴). Jedná se o homogenní kolekci, ve které jsou všechny záznamy reprezentovány stejným typem – bodem, křivkou nebo polygonem.

Prostorová data mohou být navíc provázána s neprostorovými daty pomocí třídy relací (ang. *relationship class*⁵). Ta zajišťuje provázání dat více různých tabulek pomocí vybraných klíčů. Mezi třídami prvků a tabulkami může být jeden z následujících vztahů:

- *Peer-to-peer*, kde mohou obě strany existovat samy o sobě.
- *Composite*, kde je určen hlavní záznam a pokud dojde k jeho smazání, tak jsou smazány i záznamy závislé.

Třídy prvků a třídy relací mohou být navíc součástí datové sady prvků (ang. *feature dataset*⁶). Tímto způsobem jsou v aktuálním řešení seskupeny záznamy rychlostních kamer a dopravních nehod.

Skript pro aktualizaci dopravních dat

Skript v jazyce Python extrahuje data z poskytnutých souborů dopravních nehod, které pochází ze stránek Policie ČR a nahrává je do ArcGIS geodatabáze. Záznamy, které v databázi již existují jsou ignorovány. Po přidání všech záznamů je proveden průnik nad polygonem ORP Most a souřadnicemi vložených záznamů. Ponechány jsou pouze záznamy, u kterých je průnik neprázdný. Z každého běhu programu jsou ukládány logy do zvoleného souboru. Existují dvě verze tohoto skriptu. První verze zpracovává data z nehod do konce roku 2022 a druhá pak od počátku roku 2023. Novější verze má v aktuální podobě několik omezení a nedostatků:

- Soubory dopravních nehod musí být manuálně přidány při každé jejich aktualizaci.
- Importována je pouze část všech dostupných dat. Ve starší verzi chybí pouze data chodců, v nové verzi se jedná o data chodců, následků a zúčastněných vozidel.
- Součástí skriptu není import dat rychlostních kamer města Most a jejich měření.
- V rámci skriptu se nachází velké množství pevně nastavených hodnot, které zásadně ovlivňují běh programu a pro jejichž změnu by se muselo zasáhnout do kódu programu.
- Pomalá rychlost běhu skriptu, zejména provádění operací v pořadí, které není z pohledu rychlosti efektivní.

⁴Popis třídy Feature class: <https://pro.arcgis.com/en/pro-app/latest/help/data/feature-classes/feature-classes.htm>

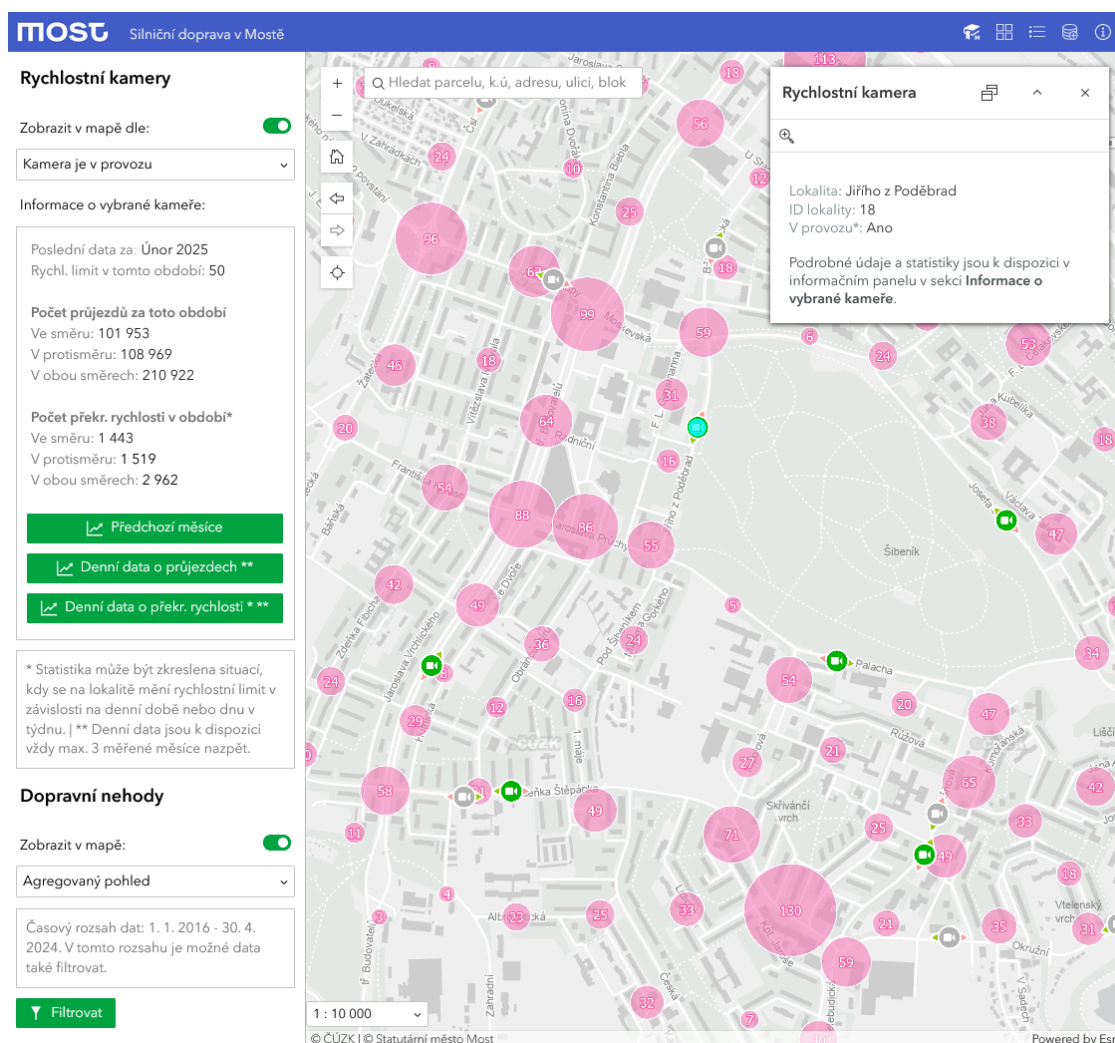
⁵Popis třídy Relationship class: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/data-management/create-relationship-class.htm>

⁶Popis kolekce feature dataset: <https://pro.arcgis.com/en/pro-app/latest/help/data/feature-datasets/feature-datasets-in-arcgis-pro.htm>

Webová aplikace Silniční doprava

Webová aplikace zobrazuje uložená data na podkladové mapě ORP Most s dopravními komunikacemi. Nad nehodami je možné provádět filtrování podle data, druhu, příčiny a přítomnosti alkoholu v krvi viníka. Podle ostatních atributů není možné filtrovat. Záznamy dopravních nehod je možné zobrazit ve dvou režimech:

- **Agregovaný pohled** seskupuje dopravní nehody dohromady podle úrovně přiblížení mapy. Tím umožňuje například zobrazit celkový počet nehod v oblasti určité křižovatky. V tomto pohledu není možné zobrazit detail jednotlivých záznamů.
- V **podrobném pohledu** je každá nehoda zobrazena jako samostatný bod. Toto zobrazení je vhodné pro větší úroveň přiblížení, protože v oddáleném pohledu není v některých případech možné rozeznat body od sebe. Oproti předešlému pohledu umožňuje zobrazit detail dané nehody ve vyskakovacím okně, ve kterém se nachází sloupec s názvy atributů a jejich hodnotami.



Obrázek 4.1: Ukázka webové aplikace Silniční doprava. Dopravní nehody jsou zobrazeny pomocí agregovaného pohledu. Údaje vybrané dopravní kamery jsou vidět v bočním panelu.

Druhým typem záznamů, které jsou na mapě zobrazovány jsou body rychlostních kamer. Po jejich rozkliknutí je v bočním panelu možné vidět statistiku denního počtu průjezdů a překročení rychlosti. Na tomto místě se dá měnit i způsob zobrazení v mapě. Rychlostní kamery je tak možné zobrazit podle:

- Stavů, zda jsou aktivní, kde zelené ikony kamer označují aktivní a zašedlé neaktivní.
- Průměrného denního počtu průjezdů, kdy je zobrazen počet průjezdů a jejich poměr v obou směrech jízdy v koláčovém grafu.
- Průměrného denního počtu překročení rychlosti, kdy je zobrazen počet překročení rychlosti a jejich poměr v obou směrech jízdy v koláčovém grafu.
- Průměrného denního počtu průjezdů dle dnů v týdnu, kdy je zobrazen sloupcový graf se sloupci pro každý den v týdnu. Pod grafem se nachází minimální a maximální hodnoty průměrného denního počtu průjezdů. Zobrazení je možné omezit podle směru jízdy.
- Průměrného denního počtu překročení rychlosti dle dnů v týdnu, kdy je zobrazen sloupcový graf se sloupci pro každý den v týdnu. Pod grafem se nachází minimální a maximální hodnoty průměrného denního počtu překročení rychlosti. Všechny grafy se pravděpodobně řídí podle nejvyšší hodnoty, a tak je obtížné rozeznat sloupce grafů u kamer s nízkými hodnotami.

V bočním panelu je možné zobrazit spojnicové grafy, které popisují vývoj hodnot měření ve dnech, kdy byla kamera aktivní. Vyskakovací okno každé kamery obsahuje informace o její lokalitě a údaj, zda je aktuálně v provozu. V panelu je spolu s ostatními hodnotami zobrazené i období poslední aktivity. Na tomto místě je také možné skrýt záznamy dopravních nehod a rychlostních kamer z mapy. Ukázku aplikace je možné vidět na obrázku 4.1.

4.3 Požadavky na nové programy

Předešlé sekce popsaly skupiny uživatel, stav současného řešení a jeho nedostatky. Z těchto poznatků a požadavků města Most, vyplývají následující požadavky na nový skript pro automatickou aktualizaci dat a novou webovou aplikaci, které nahradí stávající řešení. Vybrané technologie pro řešení tohoto zadání vychází požadavků města Most.

Požadavky na skript pro aktualizaci dat

Město Most požaduje vytvoření nového skriptu pro automatickou aktualizaci dat v jazyce Python. Ten by měl být schopný automaticky stáhnout data dopravních nehod z webových stránek Policie ČR a data měření rychlostních kamer z API otevřených dat města Most. Výsledkem zpracování dat dopravních nehod by měly být pouze záznamy, které se nachází v polygonu ORP Most, ostatní záznamy by měly být ignorovány. Vyfiltrované záznamy by měly obsahovat pouze užitečné atributy, aby se snížily časové a prostorové nároky na jejich zpracování a uložení.

Data by měla být uložena do Esri geodatabáze pomocí tříd prvků a tabulek. Třídy prvků by měly být s tabulkami relačně provázány pomocí tříd relací a composite vztahu. Při ukládání dat by se mělo kontrolovat, zda je geodatabáze využívána více uživateli současně, a je tak v takzvaném *versioned* stavu. Na základě toho by se měly spouštět úpravy

ve víceuživatelském módu pomocí třídy *Editor*⁷, která zajišťuje správné provádění změn v případě více připojených uživatel v rámci geodatabáze.

Při opakovaném spuštění skriptu by se v geodatabázi neměly vyskytnout duplicitní záznamy a vkládána by měla být pouze nová data. V případě již existujících záznamů by mělo dojít pouze k jejich aktualizaci. Výpisy z běhu skriptu by měly být uchovány ve specifikovaném souboru. Skript by měl být udržovatelný a měl by být připraven na případné změny struktury dat Policie ČR. Pro zpracování a ukládání dat by měla být využita knihovna ArcPy.

Požadavky na webovou aplikaci pro vizualizaci dat

Webová aplikace by měla být vytvořena pomocí frameworku React a knihoven z ArcGIS Maps SDK for JavaScript⁸. Data by měla být načítána z takzvaných feature služeb, pomocí kterých je možné publikovat obsah geodatabáze. Záznamy by měly být vhodně vizualizovány, aby bylo možné jednoduše identifikovat kritická místa v dopravní infrastruktuře. Nad daty dopravních nehod by mělo být možné provádět filtrování na základě hodnot jejich atributů. Záznamy by měly umožnit zobrazit detailnější popis svých hodnot po rozkliknutí zobrazovaných prvků na mapě. Pro zobrazované vrstvy by měl být k dispozici jeden nebo více druhů zobrazení včetně případné možnosti skrytí. Celkově by se mělo jednat o rozšíření a vylepšení funkcionality, kterou aktuálně nabízí aplikace Silniční doprava.

4.4 Sady dopravních dat

V této sekci jsou popsány vhodné zdroje dopravních dat, jejich způsoby získání a parametry. Podle požadavků města Most se jedná o data dopravních nehod Policie ČR a otevřená data z rychlostních kamer. Navíc jsou zde popsány i data Waze a partnerský program Waze for Cities, pomocí kterého jsou poskytována. V počátečních fázích řešení se o datech Waze uvažovalo jako o možném zdroji, nakonec však tuto datovou sadu město Most nemá zájem využívat, a tak nebyla zahrnuta.

Data dopravních nehod Policie ČR

Webové stránky Policie ČR nabízí velké množství databází, ty však ve většině případů slouží pouze pro vyhledávání nebo zobrazení souhrnných informací. Výjimkou jsou záznamy o dopravních nehodách⁹, které jsou zveřejňovány od roku 2007. Pro vlastní zpracování jsou však využitelné až od roku 2016. Do tohoto roku obsahují soubory pouze souhrnné grafy a statistiky, což neumožňuje detailnější analýzu podle vlastních parametrů. Data nehod jsou publikována pro každý měsíc v samostatném archivu. Archivy obsahují data aktuálního a všech předešlých měsíců daného roku. Pokud je cílem získat všechna data pro daný rok, stačí stáhnout pouze archiv posledního měsíce.

Poskytovaná data jsou sice strukturovaná, avšak jejich struktura se již jednou změnila na konci roku 2022. Z tohoto důvodu je vhodné být připravený na možné budoucí změny. Archivy z období od roku 2016 do konce 2022 obsahují záznamy nehod rozdělené do souborů CSV podle číselníku krajů. Navíc je zde obsažen soubor CHODCI, který obsahuje doplňující data o pěších účastnících dopravních nehod. Každý archiv tak celkově obsahuje

⁷Popis třídy Editor: <https://pro.arcgis.com/en/pro-app/latest/arcpy/data-access/editor.htm>

⁸ArcGIS Maps SDK for JavaScript: <https://developers.arcgis.com/javascript/latest/>

⁹Data dopravních nehod: <https://policie.gov.cz/clanek/statistika-nehodovosti-900835.aspx>

patnáct souborů s daty a šest prázdných souborů. Na webových stránkách policie se nachází doplňující soubor, který popisuje strukturu poskytovaných dat. Podle něj je možné určit příslušnost souborů ke krajům a zjistit jaká data jsou součástí jeho sloupců. Zároveň je možné si všimnout, že některé atributy těchto záznamů nejsou nikde popsány, a není tak možné určit o jaké hodnoty se jedná.

Aktuální struktura od počátku roku 2023 se s tou předešlou částečně shoduje. Namísto rozdělení dat do souborů podle krajů jsou však rozděleny do souborů podle typu údajů, které obsahují. Celkem se jedná o pět souborů XLS, které uchovávají data v HTML tabulkách. Jejich přibližný obsah může být popsán následovně:

- Ichodci.xls – Záznamy o chodcích, kteří byli účastníky dopravní nehody.
- Inasledky.xls – Záznamy o účastnících nehody a způsobených následcích.
- Inehody.xls – Záznamy o nehodě a popis situace, při které k ní došlo.
- IntGPS.xls – Souřadnice nehody XY v souřadnicovém systému JTSK.
- IVozidla.xls – Záznamy o zúčastněných vozidlech a řidičích.

Oproti předešlé verzi obsahují tyto soubory i větší počet sloupců s podrobnějšími nebo kompletně novými údaji. V předešlé verzi se například nevyskytoval údaj o přítomnosti alkoholu nebo jiných drog u chodců. Zároveň přibýly nové záznamy spojené s následky po nehodě, které se zde dříve vůbec nevyskytovaly. Mimo nově přidané sloupce obsahuje tabulka opět několik nadbytečných sloupců, které podle dostupných informací ze stránek policie není možné identifikovat.

V obou verzích je nutné vzájemně propojit záznamy z více souborů pomocí sloupce *p1*, který obsahuje identifikátor nehody. V některých souborech se stejný identifikátor může vyskytovat i vícekrát, například pokud se nehody účastnilo několik aut nebo bylo větší množství zraněných.

Otevřená data dopravních kamer města Most

Otevřená data města Most¹⁰ se dělí do několika kategorií jako například bezpečnost, pozemky a majetek nebo životní prostředí. Na základě předešlých požadavků jsou pro tuto práci relevantní pouze data rychlostních kamer¹¹ a lokalit měření¹², které se nachází v kategorii doprava. Data je možné stáhnout manuálně nebo přes API. V obou případech jsou data dostupná ve formátu GeoJSON. Záznamy lokalit kamer obsahují následující atributy:

- Identifikátor lokality,
- název lokality,
- status, zda je kamera aktivní,
- směr v jakém provádí měření,
- GPS pozici lokality.

¹⁰Otevřená data města Most: <https://opendata.mesto-most.cz/>

¹¹Data měření rychlostních kamer: <https://opendata.mesto-most.cz/datasets/mestomost::rychlostn%C3%AD-kamery-data-z-m%C4%9B%C5%99en%C3%AD/about>

¹²Data lokalit měření: <https://opendata.mesto-most.cz/datasets/mestomost::rychlostn%C3%AD-kamery-lokality-m%C4%9B%C5%99en%C3%AD/about>

Záznamy měření obsahují následující atributy:

- Identifikátor lokality,
- období, ve kterém měření probíhalo,
- datum měření,
- rychlostní limit,
- celkový počet průjezdů a dílčí počty průjezdů ve směru a proti směru jízdy,
- celkový počet překročení rychlosti a dílčí počty průjezdů ve směru a proti směru jízdy,
- celkový počet průjezdů v různých rychlostních intervalech a dílčí počty průjezdů ve směru a proti směru jízdy.

Data měření je možné propojit s lokalitami kamer pomocí identifikátoru lokality.

Data z programu Waze for Cities

Partnerský program Waze for Cities¹³ byl krátce zmíněn i v sekci 2.2. Waze poskytuje partnerovi informace o aktuálním dopravním stavu a hlášených dopravních událostech v oblasti, kterou spravuje. Každý záznam nahlášené události obsahuje následující atributy:

- Čas hlášení, GPS poloha a unikátní identifikátor,
- směr trajektorie, ve které se uživatel pohyboval,
- typ, podtyp a popis události,
- název ulice, jméno města a zkratka státu,
- typ silniční komunikace,
- unikátní identifikátor kolony, v případě že je s ní hlášení spojeno,
- úroveň uživatele, spolehlivost a důvěra v uživatele na základě předešlých hlášení a hodnocení,
- informace, zda událost zadal přímo partner,
- počet pozitivních reakcí od ostatních uživatelů.

Data se získávají přes API, které je poskytuje ve formátech JSON a XML. Aktualizace poskytovaných dat probíhají periodicky každé dvě minuty.

¹³Program Waze for Cities: <https://support.google.com/waze/partners/answer/10618477?hl=en>

Kapitola 5

Návrh webové aplikace a skriptu pro aktualizaci dat

Tato kapitola popisuje návrh nových částí systému na základě požadavků z předešlé kapitoly. Nejdříve je popsána architektura systému a navržený datový model, který integruje data vybraných datových sad. Dále je navržen skript pro automatickou aktualizaci dat, který umožňuje plně využít vybrané zdroje dat a umožňuje lepší udržitelnost při případných změnách. Poslední část je věnována návrhu webové aplikace a způsobům vizualizace, pomocí kterých je možné data analyzovat.

5.1 Architektura systému

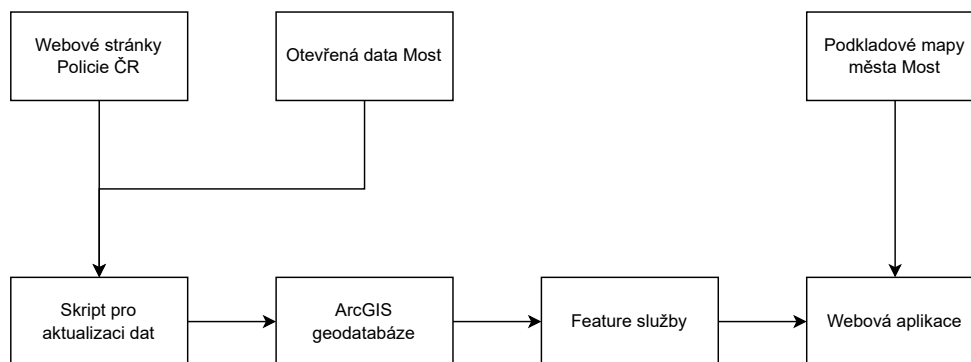
Navržený systém se podle požadavků skládá ze dvou hlavních součástí – skriptu pro automatickou aktualizaci dat a webové aplikace pro jejich vizualizaci. Dílčí části je možné popsat následovně:

- Skript bude na serveru města Most periodicky v měsíčním intervalu volán plánovačem úloh. Data pro zpracování bude automaticky získávat z webových stránek Policie ČR a API otevřených dat města Most. Získaná data zpracuje a pomocí nich pak bude aktualizován obsah geodatabáze. Z aktualizované geodatabáze budou publikovány feature služby, které k záznamům umožní vnější přístup.
- Webová aplikace bude využívat vybraných feature služeb pro získání dat. Nad získanými daty pak bude provádět mapové vizualizace. Změny zobrazení a nastavení jeho parametrů budou probíhat přímo na straně klientské aplikace. Pro případy složitějších operací bude možné odeslat dotaz na danou feature službu. Podkladové mapy budou získány z galerie podkladových map města Most.

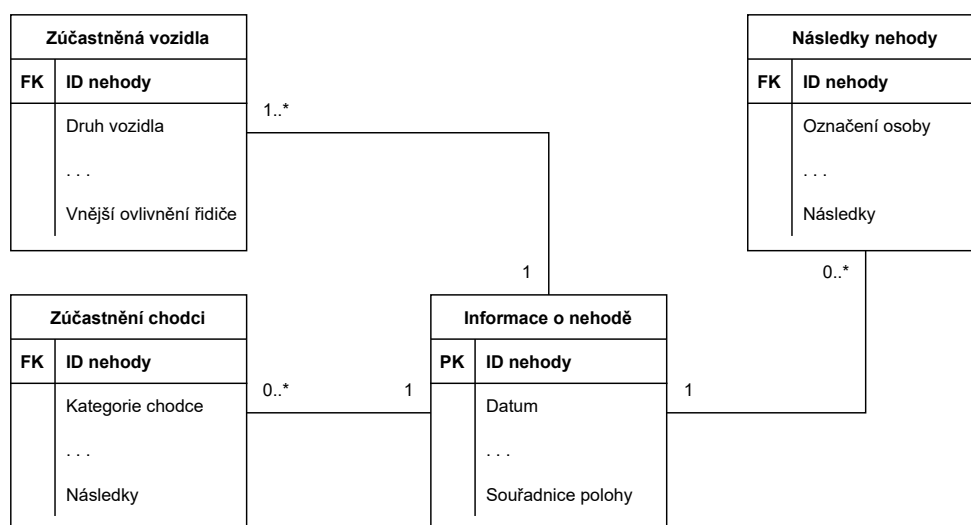
Schéma navrženého řešení je vidět na obrázku 5.1.

5.2 Datový model

Navržený systém bude podle požadavků pracovat s datovou sadou dopravních nehod Policie ČR a otevřenými daty města Most z měření rychlostních kamer, které byly popsány v sekci 4.4. Tyto dvě sady nejsou mezi sebou přímo propojené, ale obsahují záznamy, které při společné vizualizaci mohou umožnit lepší zasazení do kontextu. Toto může například



Obrázek 5.1: Schéma navrženého řešení.



Obrázek 5.2: Relační diagram uložení dat dopravních nehod.

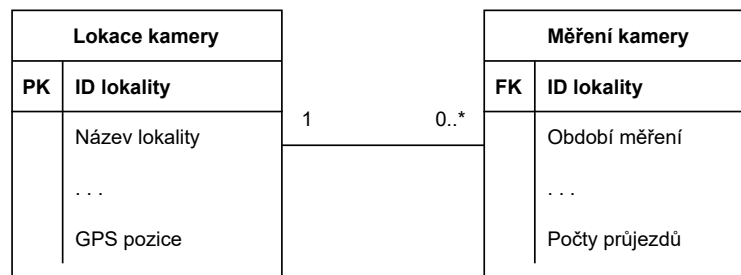
pomoci při odhalování spojitostí míst s velkým množstvím překročení povolené rychlosti a místem častých dopravních nehod.

Data dopravních nehod

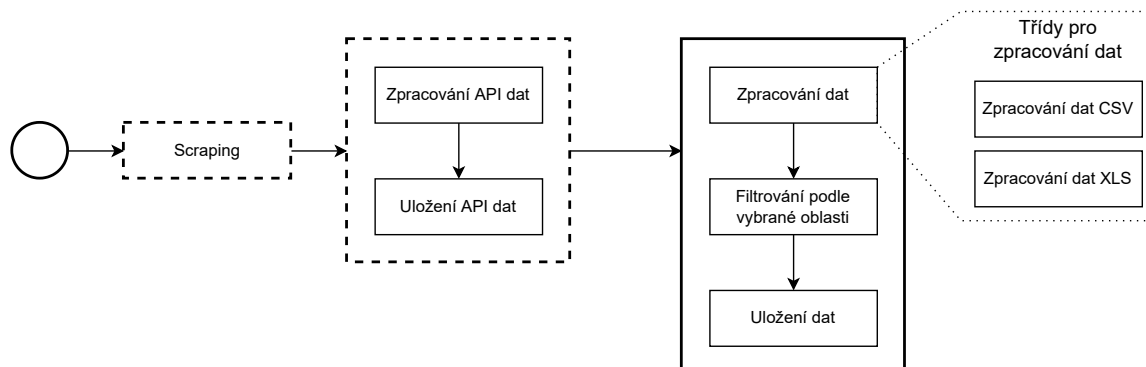
Data dopravních nehod budou v geodatabázi uložena pomocí tříd prvků, které byly společně s dalšími termíny popsány v sekci 4.2 a tabulek. Hlavní údaje o dopravní nehodě včetně její polohy budou uloženy pomocí třídy prvků. Doplňující údaje jako například záznamy o chodcích nebo zúčastněných vozidlech budou uloženy v tabulkách. Jejich záznamy pak budou propojeny s hlavními přes unikátní identifikátor nehody pomocí tříd relací. Mezi takto propojenými záznamy bude vytvořen composite vztah, aby po odstranění hlavního byly odstraněny i závislé záznamy. Pro data starší struktury půjde pouze o propojení záznamů nehod se záznamy chodců. Relační diagram je možné vidět na obrázku 5.2.

Data rychlostních kamer

Data rychlostních kamer budou v geodatabázi uložena pomocí stejných typů tabulek jako v případě dat dopravních nehod. Lokality kamer budou uloženy pomocí tříd prvků a jejich



Obrázek 5.3: Relační diagram uložení dat rychlostních kamer.



Obrázek 5.4: Schéma programu rozděleného do modulů. Na sebe navazující moduly se nachází v ohraničených oblastech. Průchod moduly v čárkovaném ohraničení je volitelný a odvíjí se od aktuální konfigurace. Pro modul zpracování dat jsou zobrazeny příklady tříd, které by mohly být podle obsahu konfigurace vloženy na toto místo při běhu programu.

měření pak pomocí tabulek. Měření s daty budou propojena pomocí unikátního identifikátoru lokality kamery pomocí třídy relací. Mezi těmito záznamy bude composite vztah, který zajistí při odstranění lokality kamery i odstranění záznamů, které jsou s ní spojené. Relační diagram je možné vidět na obrázku 5.3.

5.3 Skript pro automatickou aktualizaci dat

Skript v jazyce Python umožní automaticky stahovat dopravní data, provádět jejich zpracování a ukládat je do souborové geodatabáze. Celý program bude rozdělen na několik modulů, jejichž funkcionalitu bude možné jednoduše upravovat nebo nahrazovat jako celek. Rozdělení programu na moduly je možné vidět na obrázku 5.4. Zároveň je cílem vytvořit co nejuniverzálnější program, aby ho bylo možné využít i pro polygony odlišných oblastí pro případ, že by tento skript chtěly využít i jiné územní celky.

Moduly programu se budou vždy skládat z abstraktní základní třídy a alespoň jedné další, která z ní bude dědit a bude implementovat její metody. Takových tříd může být i větší množství, což umožní práci s různými typy dat, využívání odlišných knihoven nebo odlišné způsoby ukládání dat. Vybrané třídy pak budou vloženy na příslušná místa v programu.

Aby bylo možné upravovat funkcionalitu programu bude vytvořen konfigurační soubor. Ten bude obsahovat nastavení pro výběr zmíněných modulů a pro celý běh programu.

Budou zde definovány webové stránky a API, ze kterých mají být získána data, parametry pro zpracování dat, parametry pro načtení polygonu oblasti a parametry pro uložení dat.

V rámci některých modulů bude podle požadavků nutné využít funkce licencované knihovny ArcPy. Cílem je však vytvořit program, který bude možné využít pro jakýkoliv polygon území. Z tohoto důvodu je nutné vytvořit v takovýchto situacích alternativní třídy, které budou využívat volně dostupné alternativy. Výsledná data bude také nutné uložit odlišným způsobem, protože souborová geodatabáze je proprietárním formátem Esri. Licence ArcPy bude vyžadována pouze v případě, že budou v konfiguračním souboru nastaveny třídy, které s jejími metodami pracují, jinak bude možné program využívat bez její nutnosti.

Na základě požadavků bude program vypisovat záznamy o jeho běhu. Ty se budou nacházet na důležitých místech v programu – například před a po vstupu do některého z modulů. Přesněji se bude jednat o následující druhy výpisů:

- **Informativní výpisy** budou sloužit pro hlášení stavu běhu, například ve které části programu se aktuálně nachází.
- **Ladící výpisy** budou obsahovat informace o načtených modulech a zpracovávaných datech. Bude se jednat o detailnější popis, který usnadní hledání případných chyb.
- **Varovné a chybové výpisy** budou informovat o neočekávaných chybách a přeskočených operacích.

Způsob uložení těchto výpisů a jejich úroveň závažnosti bude možné nastavit v konfiguraci.

Získávání dat

Získávání dat bude v programu prováděno na dvou různých místech. Prvním bude modul pro *scraping*, který bude zajišťovat získávání dat dopravních nehod z webových stránek Policie ČR. Získaná data budou uložena na místo určené konfigurací, aby při opakovaném spuštění nemusela být znovu získávána. V případě dat dopravních nehod bude navíc nutné extrahovat soubory z archivů. Extrahované soubory pak budou uloženy do jedné složky, čímž bude zachováno jejich propojení.

Pokud se cílová adresa pro scraping změní, bude možné její upravení v konfiguračním souboru. Pokud se změní struktura stránek, bude pravděpodobně nutná úprava stávající třídy nebo vytvoření nové. V případě, že se počet cílových stránek zvýší, bude možné tento seznam zadat do konfiguračního souboru a pro každou stránku definovat vlastní scraper. Pokud nebude cílem provádět scraping žádných webových stránek, program umožní tuto část přeskočit.

Druhým místem získávání dat bude modul pro zpracování dat z API definovaného v konfiguračním souboru. Modul bude obsahovat třídu pro získání a zpracování otevřených dat z rychlostních kamer města Most. Stejně jako v předešlém případě bude možné upravit cílové adresy a jejich počet, aby bylo možné získat i jiná data.

Zpracování dat

Moduly zpracování budou pracovat s daty z API a lokálně uložených souborů. V případě dat z API není nutné provádět operace předzpracování, protože otevřená data dopravních kamer jsou již zpracována městem Most. Navíc není nutné provádět filtrování pro polygon města Most, protože se v něm rychlostní kamery jistě nachází. V případě dat z lokálních

souborů dopravních nehod je však předzpracování a filtrování nutné. Data je nejdříve nutné načíst a následně zpracovat podle obsažených hodnot.

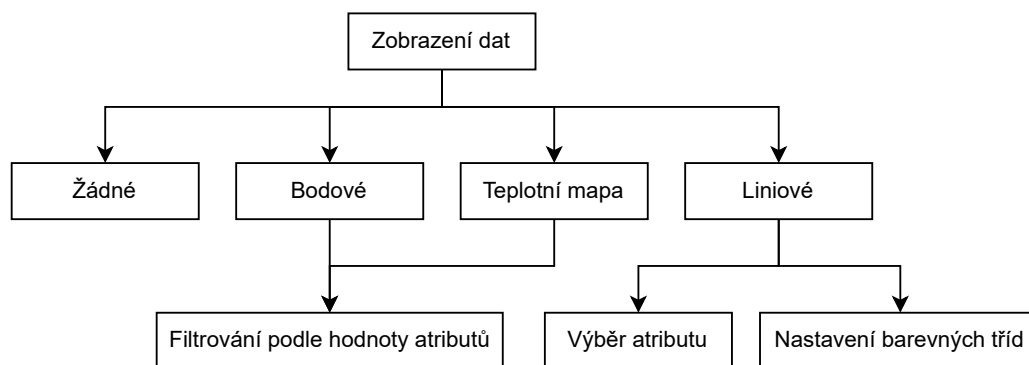
V případě dat, která obsahují prostorové záznamy, bude kontrolována jejich existence v databázi. Pokud se zde již budou vyskytovat záznamy se stejným identifikátorem nehody, budou přeskočeny. Pokud půjde o nové záznamy, bude nutné načíst polygon oblasti a následně záznamy vyfiltrovat, podle toho, jestli do oblasti náleží. Nakonec budou zbylá data transformována a budou odstraněny zbytečné atributy. Pokud data nebudou obsahovat prostorové údaje, budou filtrována podle identifikátorů, které již v databázi existují nebo byly dříve zpracovány. Záznamy, jejichž identifikátor nebude nalezen, budou patřit k nehodě mimo specifikovaný polygon. Filtrování podle polygonu u těchto dat tak nebude potřeba provádět. Následně bude data opět nutné transformovat a odstranit zbytečné atributy.

Uložení dat

Nová data budou vložena do ArcGIS geodatabáze pomocí metod knihovny ArcPy. Již existující záznamy budou pouze aktualizovány novými daty. Způsob uložení a případné vztahy mezi tabulkami bude možné nastavit v konfiguračním souboru. Všechny změny budou provedeny atomicky a případně i v režimu více připojených uživatelů. V případě, že nebude možné využít licencovaných metod, bude možné uložení dat do odlišných formátů. Nastavení výstupního formátu a způsobu uložení bude v tomto případě opět možné v rámci konfiguračního souboru.

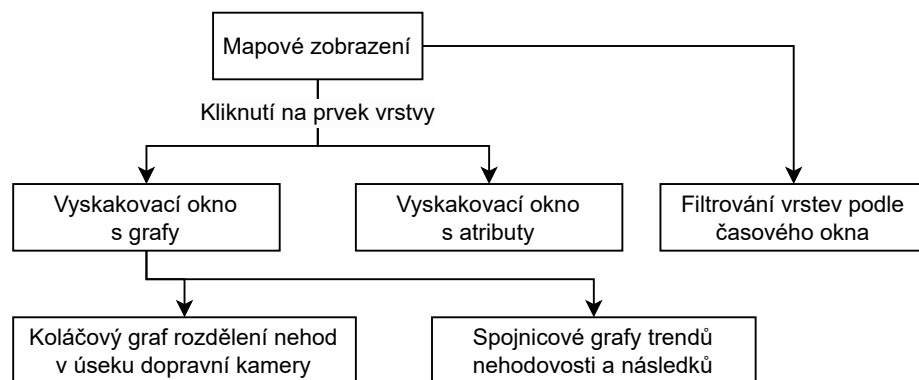
5.4 Webová aplikace pro vizualizaci dat

Webová aplikace bude na základě požadavků vytvořena pomocí frameworku React a knihoven z ArcGIS Maps SDK. Základem vizualizace bude podkladová mapa, která bude získána z galerie map města Most. Na tuto mapu budou zobrazována uložená data, která budou načtena z feature služeb publikovaných městem Most. Seznam těchto služeb bude možné definovat v konfiguračním souboru, aby je bylo možné přidávat nebo odebírat podle aktuální situace bez nutnosti zásahů do kódu.



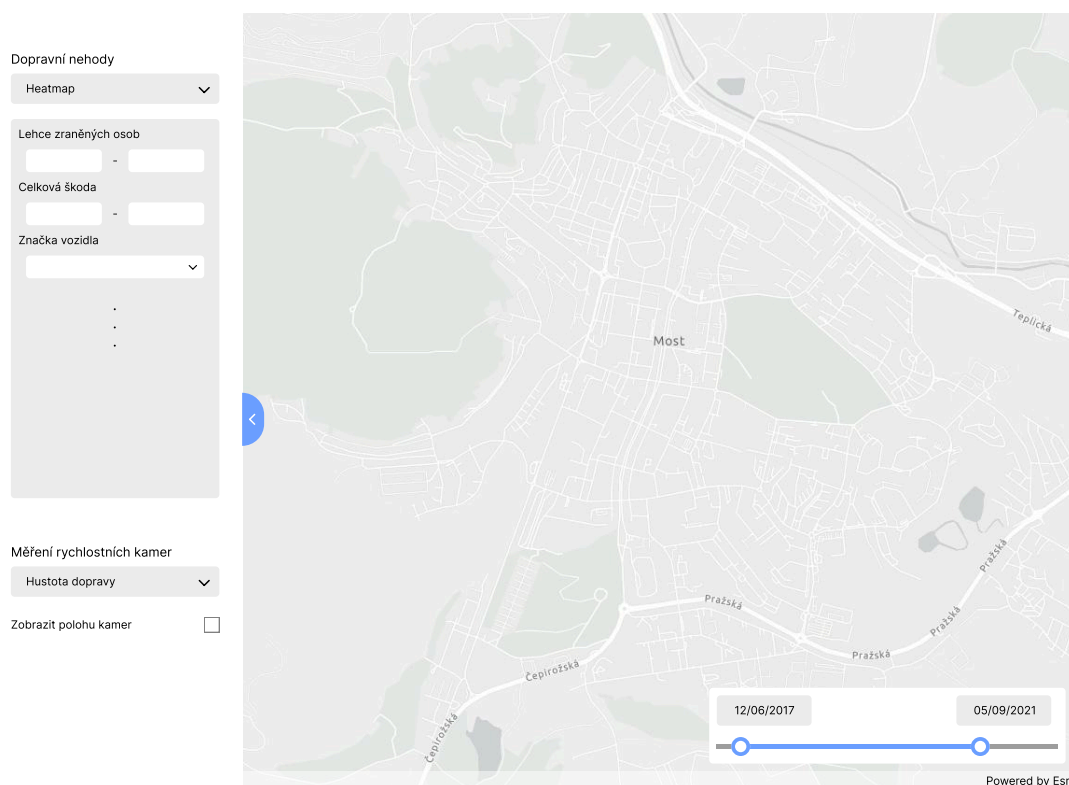
Obrázek 5.5: Dostupné typy zobrazení a možné akce pro jejich přizpůsobení.

Aplikace bude umožňovat zobrazení pomocí bodů, linií a teplotní mapy. Bodové zobrazení a teplotní mapa budou určeny pro vizualizaci míst dopravních nehod. Liniové zobrazení bude využito pro vizualizaci hodnot spojených s dopravními komunikacemi. Data dané vrstvy bude možné i skrýt, aby nepřekážely vizualizacím ostatních vrstev. Dostupné typy



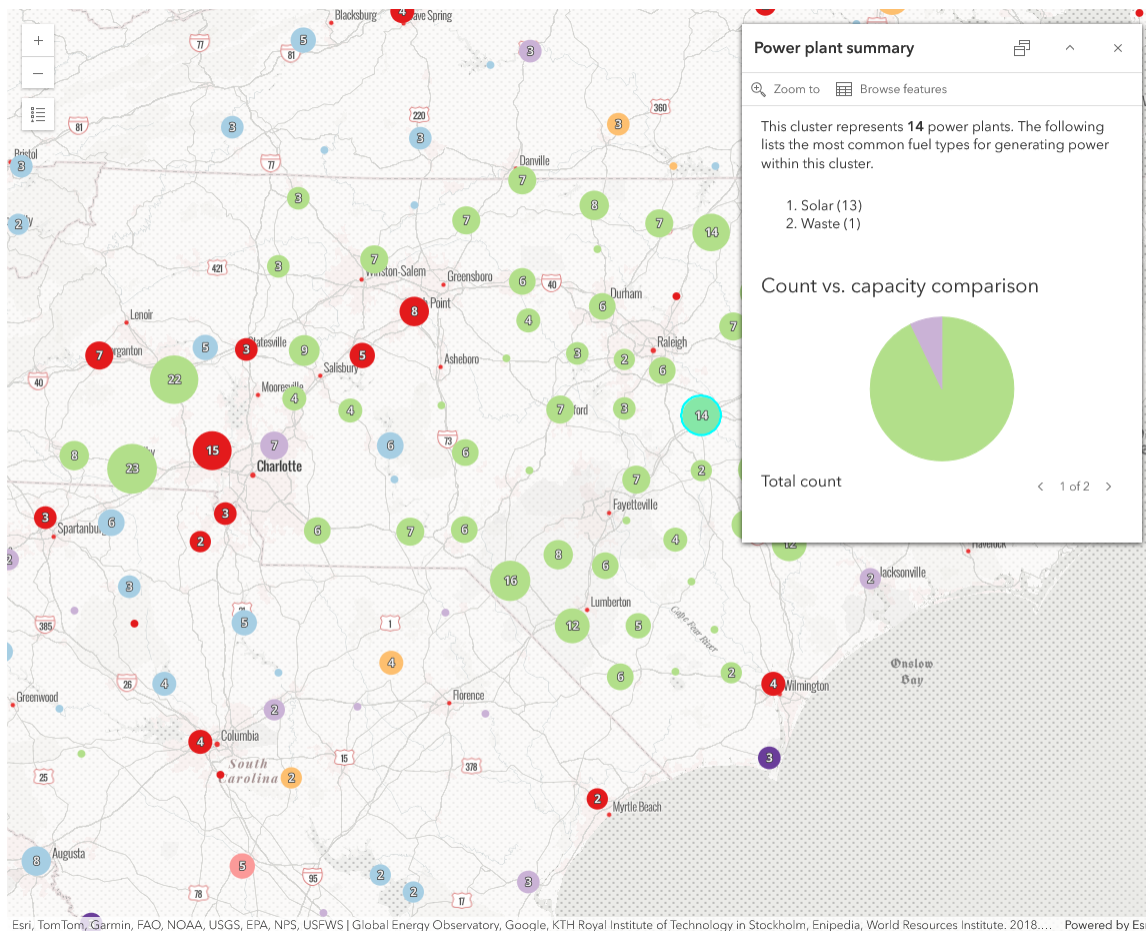
Obrázek 5.6: Dostupné akce nad datovými vrstvami v rámci mapového zobrazení.

zobrazení datových vrstev a jejich parametry bude možné upravovat přes konfigurační soubor. Záznamy dopravních nehod bude možné filtrovat pro vybrané období pomocí časového posuvníku. Společně s tím je bude možné filtrovat i podle definovaných atributů. Silnice budou rozděleny do barevných tříd na základě hodnoty vybraného atributu. Každá třída bude definovaná intervalem, jehož rozsah bude možné upravovat. Intervaly tříd společně pokryjí všechny možné hodnoty vybraného atributu. Schéma typů zobrazení a možné akce pro jejich úpravu je vidět na obrázku 5.5.



Obrázek 5.7: Návrh podoby uživatelského rozhraní webové aplikace.

Po kliknutí na zobrazovaný prvek bude otevřeno vyskakovací okno se specifickým obsahem podle toho, ke které datové vrstvě bude náležet. Pro záznamy dopravních nehod se



Obrázek 5.8: Ukázka zobrazení grafu ve vyskakovacím okně pomocí Arcade expressions. Mezi více grafy je možné přepínat pomocí tlačítek v pravém dolním rohu. Obrázek získán z [21].

bude jednat o výčet atributů a jejich hodnot. V něm budou namapovány číselné hodnoty atributů na jim odpovídající text, aby tyto informace nebylo nutné dohledávat na stránkách Policie ČR. V případě dopravních kamer bude zobrazen graf počtu nehod před umístěním, v době umístění a po odstranění kamery. Dále budou zobrazeny průměrné počty průjezdů a počty překročení rychlosti. Vyskakovací okna silnic budou obsahovat spojnicové trendy nehodovosti a následků v posledních pěti letech. Schéma dostupných akcí nad prvky datových vrstev je vidět na obrázku 5.6. Návrh rozložení uživatelského rozhraní na základě popsaného návrhu je možné vidět na obrázku 5.7. Návrh způsobu zobrazení grafů v rámci vyskakovacích oken pomocí Arcade expressions¹ je vidět na obrázku 5.8.

¹<https://developers.arcgis.com/javascript/latest/api-reference/esri-popup-content-ExpressionContent.html>

Kapitola 6

Implementace nástrojů

Tato kapitola se věnuje popisu implementace navržených nástrojů z předešlé kapitoly. Popis je rozdělen na sekci skriptu pro automatickou aktualizaci dat a sekci webové aplikace. V těchto částech jsou popsány vybrané technologie, dílčí části nástrojů a vytvořené konfigurační soubory. Sekce věnující se skriptu zahrnuje i popis ArcGIS Arcade¹ skriptů, které jsou využity pro výpočet hodnot v rámci ArcGIS geodatabáze.

6.1 Skript pro aktualizaci dat

Vytvořený Python skript *traffic_accidents.py* automaticky získává data ze stránek Policie ČR a API otevřených dat města Most, provádí jejich zpracování a ukládá je do ArcGIS geodatabáze nebo do určených souborů. Pro zpracování dat je v programu využito knihoven ArcPy a GeoPandas. Program je rozdělen do několika modulů, jejichž funkcionality se při běhu programu může měnit podle specifikované konfigurace. Struktura modulů obsahuje základní abstraktní třídu, ze které jsou odvozeny třídy pro specifické situace. Každý modul obsahuje třídy k požadavkům města Most na získávání, zpracování a ukládání dat. Z tříd modulů je vybíráno pomocí funkce *getattr*. Logování běhu skriptu je zajištěno pomocí standardní knihovny logging a jeho parametry je možné měnit v konfiguračním souboru. Vzhledem k rozsáhlé struktuře konfiguračního souboru je k dispozici i soubor *schema.json*, který popisuje obsah konfigurace a na základě kterého je její správnost při spuštění skriptu kontrolována. Tento popis je vytvořen podle JSON schema², což je standard pro popis struktury souborů JSON. Celý proces aktualizace dat včetně rozdělení skriptu do modulů a Arcade skriptů pro výpočet dodatečných hodnot je možné vidět na obrázku 6.1.

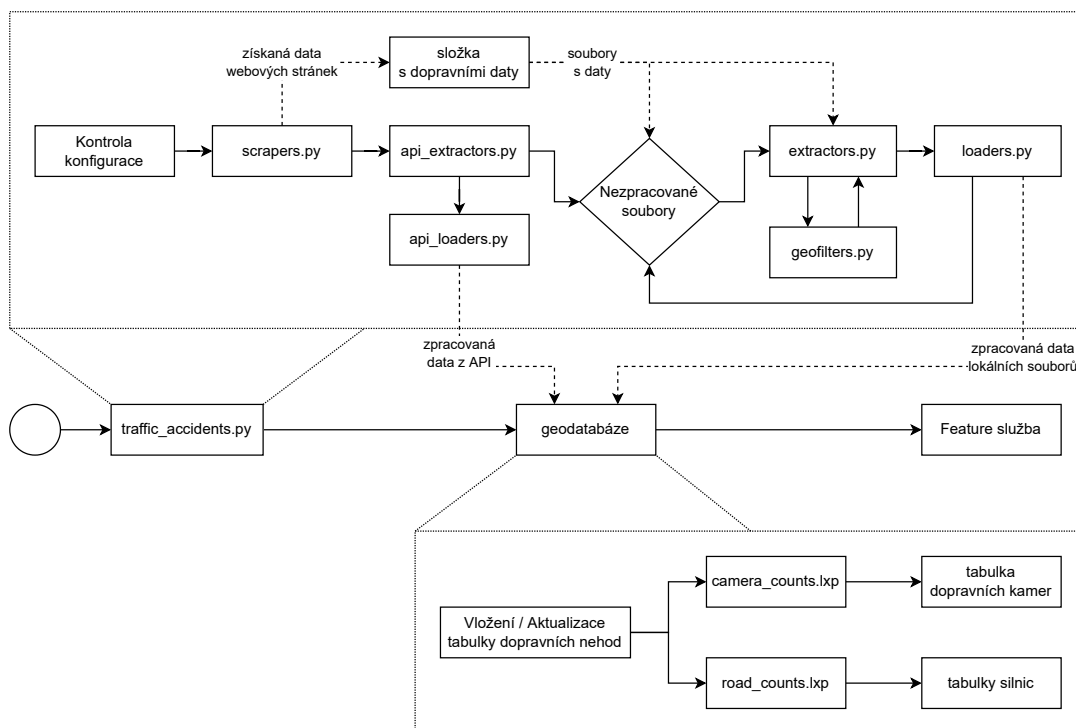
Získávání dat ze stránek Policie ČR

Po kontrole správnosti konfiguračního souboru je prvním krokem získání dat pro zpracování. Pro tento účel byl vytvořen scraper v souboru *scrapers.py*, který z webových stránek Policie ČR získá data dopravních nehod za uplynulé roky a data nejaktuálnějšího měsíce současného roku. Takto jsou získána data od počátku roku 2016 do současnosti. Starší data byla publikována pouze ve formátu PDF, a nemají tak pro zpracování význam. Scraping je zajištěn pomocí standardních knihoven jazyka Python a knihovny BeautifulSoup³. Parametry jako adresa cílové stránky nebo využitý parser jsou konfigurovatelné, ale mimo ně

¹ArcGIS Arcade: <https://developers.arcgis.com/arcade/>

²JSON schema: <https://json-schema.org/overview/what-is-jsonschema>

³BeautifulSoup: <https://beautiful-soup-4.readthedocs.io/en/latest/>



Obrázek 6.1: Schéma procesu aktualizace dat zobrazující modulární rozdělení skriptu a aktualizací skriptu Arcade. Čárkované šipky v rámci Python skriptu označují operace, které čtou nebo zapisují data.

je scraper vytvořen na míru pro stránky Policie ČR. V případě změn jejich struktury bude nutné upravit stávající implementaci nebo vytvořit novou třídu. To stejné platí i v případě přidání dalších cílů s odlišnou strukturou.

Ze stránek jsou získány odkazy pro stažení datových archivů a ty jsou následně staženy do cílové složky. Archivy jsou poté extrahovány pomocí knihovny Patool a soubory archivu jsou uloženy do vlastní složky, aby je bylo možné při zpracování provázat. Pokud není cílem provádět scraping, je možné nahradit aktuální seznam cílů prázdným seznamem a tento krok bude přeskočen. Při přeskočení tohoto kroku je nutné soubory pro zpracování dodat do určené datové složky manuálně.

Zpracování lokálních dat

Po získání dat následuje zpracování všech souborů, které se nachází v rámci určené datové složky. Soubory, které nejsou definované v konfiguraci, jsou automaticky přeskočeny. Pokud se ve složce nenachází žádné definované soubory, je přeskočena celá složka. Celá složka je přeskočena i v případě, že žádné její soubory neobsahují prostorová data, protože bez této informace není možné filtrovat pro určenou oblast. V případě složek, které splňují stanovené požadavky jsou nejdříve zpracovány soubory, které obsahují prostorová data. Toto pořadí je nutné, aby bylo možné filtrovat podle oblasti polygonu i navazující neprostorová data. Daný soubor se následně zpracuje vybranou třídou ze souboru *extractors.py*.

Zpracování probíhá v několika krocích, prvním je načtení dat z uložených souborů. Pokud jsou data načítána poprvé, je použita metoda pro získání dat z CSV nebo XLS.

Následně jsou uložena na stejné místo ve formátu pickle⁴ a originální soubor je smazán. Toto má významný vliv na celkovou rychlost skriptu, protože rychlost čtení dat je několikanásobně vyšší než při čtení původních formátů. Dodatečným benefitem je snížení celkové velikosti uložených dat. Při opakovaném spuštění skriptu jsou tak načítána přímo data ze souborů pickle. Načtená data jsou uchována v DataFrame, což je struktura knihovny Pandas, která se skládá z řádků a sloupců. Sloupce definují atributy dat a řádky pak obsahují samotné záznamy. Tato struktura a další metody knihovny Pandas umožňují efektivní práci s velkým množstvím dat, a jsou tak využívány v této a všech dalších částech programu.

V dalším kroku následuje načtení již dříve zpracovaných dat. Toto načtení probíhá z databáze v případě knihovny ArcGIS a z určeného souboru v případě knihovny GeoPandas. Po tomto kroku následuje filtrování, při kterém jsou data s prostorovými údaji filtrována podle identifikátorů již zpracovaných dat a ponechány zůstanou pouze nové záznamy. Data jsou následně filtrována pomocí vybrané třídy z modulu *geofilters.py*. Tento modul zajišťuje načtení filtrovacího polygonu, detekci a opravu chybných souřadnic záznamů a samotné filtrování záznamů pro daný polygon. Načtení polygonu a očekávané formáty se liší podle toho, zda je využita knihovna ArcPy nebo GeoPandas. Detekce a oprava souřadnic řeší situace, při kterých jsou souřadnice X a Y zaměněny. Tento problém se vyskytuje v části dat dopravních nehod. Řešením je maska, která kontroluje, zda je souřadnice X větší než souřadnice Y, což by pro oblast České republiky mělo vždy platit. Filtrování záznamů pro vybraný polygon je provedeno maskou, která kontroluje průnik polygonu s body nehod. Operace průniku jsou vektorizované, a tak je dosaženo vysoké rychlosti. Výsledkem je DataFrame záznamů, které se nachází v určeném polygonu. Data bez prostorových údajů jsou filtrována na základě identifikátorů již uložených dat a identifikátorů dříve zpracovaných dat s prostorovými údaji. Tímto se zajistí, že jsou zpracovávány pouze záznamy, které patří do požadované oblasti. V posledním kroku je upraven formát sloupců s datem a z DataFrame jsou odstraněny prázdné sloupce a sloupce definované v konfiguraci.

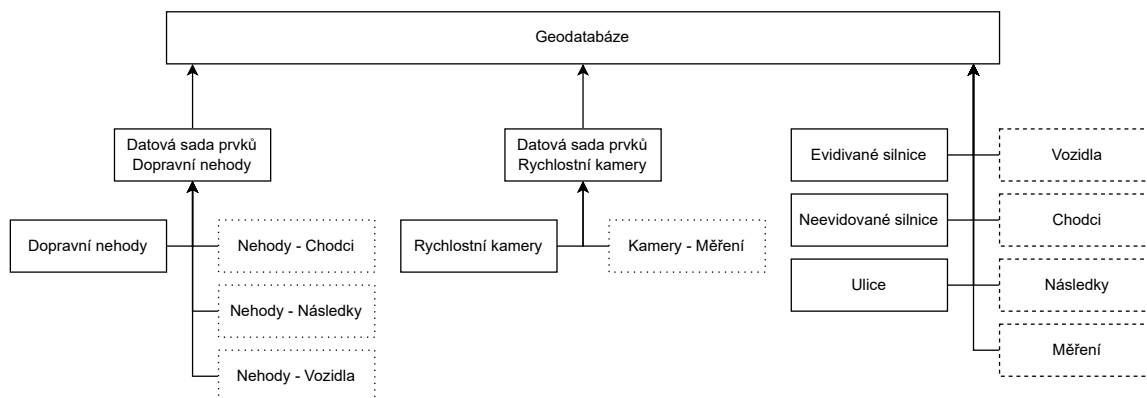
Uložení zpracovaných dat

Ukládání dat zajišťuje modul *loaders.py*, ve kterém jsou vytvořeny oddělené třídy pro ukládání dat pomocí knihoven ArcPy a GeoPandas. Třída pro knihovnu ArcPy začne kontrolou přítomnosti specifikované geodatabáze a případně ji vytvoří. Zpracovaná data jsou podle konfigurace uložena na určené místo v geodatabázi jako třída prvků v případě prostorových dat nebo jako tabulka v případě neprostorových dat. Třídy prvků se mohou navíc nacházet v rámci datové sady prvků, jejíž existenci je nutné zkontrolovat a případně ji vytvořit.

V případě, že daná třída prvků nebo tabulka neexistuje, je nutné ji vytvořit. Do vytvořených tříd prvků a tabulek jsou přidány sloupce určených typů s případným aliasem. Názvy sloupců jsou získány z DataFrame zpracovaných dat. Jejich typ je možné určit automaticky na základě typů sloupců nebo manuálně pomocí konfigurace. Případné aliasy sloupců, což jsou namapované názvy atributů na zpravidla pochopitelnější textové reprezentace, se opět řídí podle obsahu konfigurace. Do vytvořené tabulky jsou zpracovaná data vkládána pomocí třídy Editor, která zajišťuje atomické změny, které fungují správně i ve víceuživatelském módu a třídy InsertCursor, která samotné vkládání provádí.

V případě, že daná tabulka již existuje, jsou zpracovaná data maskována podle toho, zda se již v geodatabázi nachází záznamy se stejným identifikátorem. Záznamy jsou rozděleny na nové, které jsou vkládány a aktualizující, které pouze aktualizují hodnoty již existujících záznamů. Pokud se v uložených záznamech vyskytují duplicitní identifikátory, je porovná-

⁴Python Pickle: <https://docs.python.org/3/library/pickle.html>



Obrázek 6.2: Schéma vytvořené geodatabáze podle požadavků města Most. Plné čáry ohraňují třídy prvků, čárkované ohraňují tabulky a tečkované ohraňují třídy relací.

váno jejich pořadí. Pokud tak například dva uložené a tři zpracované záznamy sdílí stejný identifikátor, pak jsou první dva zpracované záznamy využity pro aktualizaci a třetí je vložen jako nový. Sloupce uložených a aktualizujících záznamů se mohou lišit, a tak je nutné tyto DataFrame sloučit, aby záznamy po aktualizaci obsahovaly všechny sloupce. Před aktualizací tabulky jsou kontrolovány její sloupce a chybějící jsou přidány stejným způsobem jako při vytváření nové tabulky. Vkládání nových dat pomocí tříd Editor a InsertCursor je také stejné, pro aktualizaci dat je však využita třída UpdateCursor.

Ke všem záznamům je před vložením nebo aktualizací přidán sloupec se značkou, která určuje, zda jde o poslední modifikaci databáze. Tato značka je klíčová pro proces aktualizace geodatabáze a její role je popsána v podsekcí 6.1, která popisuje skripty Arcade. Po aktualizaci dat jsou mezi tabulkami vytvořené vztahy pomocí tříd relací. Vytváření vztahů se odvíjí od obsahu konfiguračního souboru a je prováděno po aktualizaci tabulek, aby byla zajištěna existence tříd prvků, jejichž data jsou zpracována přednostně. Schéma vytvořené geodatabáze pro město Most se nachází na obrázku 6.2.

Druhá třída zajišťuje uložení dat pomocí metod knihovny GeoPandas. Zpracované záznamy jsou podle přítomnosti identifikátorů v uložených záznamech opět rozděleny na nové a aktualizující. Sloupce sdílené mezi uloženými a aktualizujícími daty jsou upraveny novými hodnotami a sloupce nové jsou k ukládaným datům přidány. Z nových a aktualizovaných dat je vytvořena GeoDataFrame, která je následně uložena do specifikovaného souboru.

Zpracování dat z API

Moduly pro zpracování a uložení dat z API byly záměrně ponechány nakonec, protože sdílejí velkou část funkcionality s předešlými moduly. Zejména modul pro ukládání dat *api_loaders.py* se od předešlého popisu ukládání liší pouze v jednodušším uložení dat pomocí metod knihovny GeoPandas. Získání a zpracování dat modulem *api_extractors.py* je pro knihovny ArcPy a GeoPandas podobné. Protože jsou poskytována data již ve formátu GeoJSON, není nutné provádět žádné významnější zpracování. Nejpodstatnější částí je postupná tvorba a odesílání požadavků na určenou adresu, kterou je nutné opakovat, dokud není značka *exceededTransferLimit* v odpovědi nastavena na false. Postupné posílání požadavků je zásadní, aby bylo zajištěno, že jsou získaná data kompletní, protože se data nemusí vlézt do jedné odpovědi.

Výpočet hodnot pomocí skriptů Arcade

Protože jsou v rámci webové aplikace zobrazovány hodnoty, které v uložených datech nejsou přítomny, je nutné je dopočítat. Příkladem mohou být počty nehod v úsecích dopravních kamer před jejich instalací, v rámci jejich instalace a po jejich odstranění nebo trendy nehodovosti na silnicích. Pro tento účel byly vytvořeny dva Arcade skripty, které provedou výpočty těchto hodnot po aktualizaci třídy prvků dopravních nehod. Skript *camera_counts.lxp* aktualizuje hodnoty třídy prvků dopravních kamer a *road_counts.lxp* aktualizuje hodnoty tříd prvků dopravních komunikací. Vrstvy dopravních komunikací byly získány z mapového serveru Českého úřadu zeměměřického a katastrálního⁵ a následně byly upraveny v programu ArcGIS Pro, aby obsahovaly pouze prvky pro oblast ORP Most. Pro správnou funkci skriptů je nutné zajistit přítomnost všech potřebných atributů v třídách prvků dopravních kamer a silnic.

Skript pro výpočet hodnot dopravních kamer postupně prochází záznamy dopravních kamer a vyhledává nejbližší silnice pomocí funkcí Buffer a Intersects. Poté je ze silnic vybrána ta s nejkratší vzdáleností od bodu kamery a pomocí funkcí Buffer a Intersects jsou získány počty dopravních nehod v úseku dané kamery. Podle záznamů měření, které jsou s dopravními kamerami propojené, je zjištěno období, ve kterém dopravní kamera prováděla měření. Na základě toho jsou zjištěny počty nehod před instalací, v rámci instalace a po odstranění kamery. Mimo dopravní nehody jsou počítány i průměrné hodnoty průjezdů a překročení rychlosti pro dny v týdnu. Tato data jsou opět získána z propojené tabulky měření.

Výpočty trendů počtů dopravních nehod, zranění a úmrtnosti na silnicích zajišťuje druhý skript. Obdobně jako v předešlém skriptu jsou procházeny nehody, je vytvořen Buffer, pomocí Intersects jsou nalezeny silnice a z nich je vybrána podle vzdálenosti ta nejbližší, pokud se nějaká v bufferu nachází. K silnicím jsou tak postupně přičítány nehody a další atributy podle povahy nehody. Běh obou skriptů je podmíněn výskytem dříve zmíněné značky, čímž je zajištěno, že se skript spustí pouze jednou při vložení nebo aktualizaci posledního záznamu. Tímto se předchází významnému zpomalení, které by jinak dělalo skripty nepoužitelnými.

6.2 Webová aplikace pro vizualizaci a analýzu dat

Webová aplikace vizualizuje data získaná pomocí skriptu pro aktualizaci dat a umožňuje jejich analýzu. Data jsou získávána přes feature službu publikovanou městem Most. Tento přístup tak případně umožňuje vizualizovat i jiné zdroje dat, které jsou přístupné přes odlišné feature služby. Aplikace je vytvořena v jazyce TypeScript ve frameworku React. Pro vizualizaci dat a přístup k mapovým komponentům je využita sada knihoven z ArcGIS Maps SDK for JavaScript⁶. Mapové komponenty jsou vhodnou volbou z důvodu plánovaného ukončení podpory dříve využívaných widgetů v následujících letech⁷. Doplnující komponenty pochází z knihovny Calcite components⁸, které jsou použity při výběru a úpravě parametrů zobrazení dat. Zvolené komponenty a ostatní části aplikace jsou responzivní, a tak je možné ji využít i na mobilních zařízeních nebo tabletech.

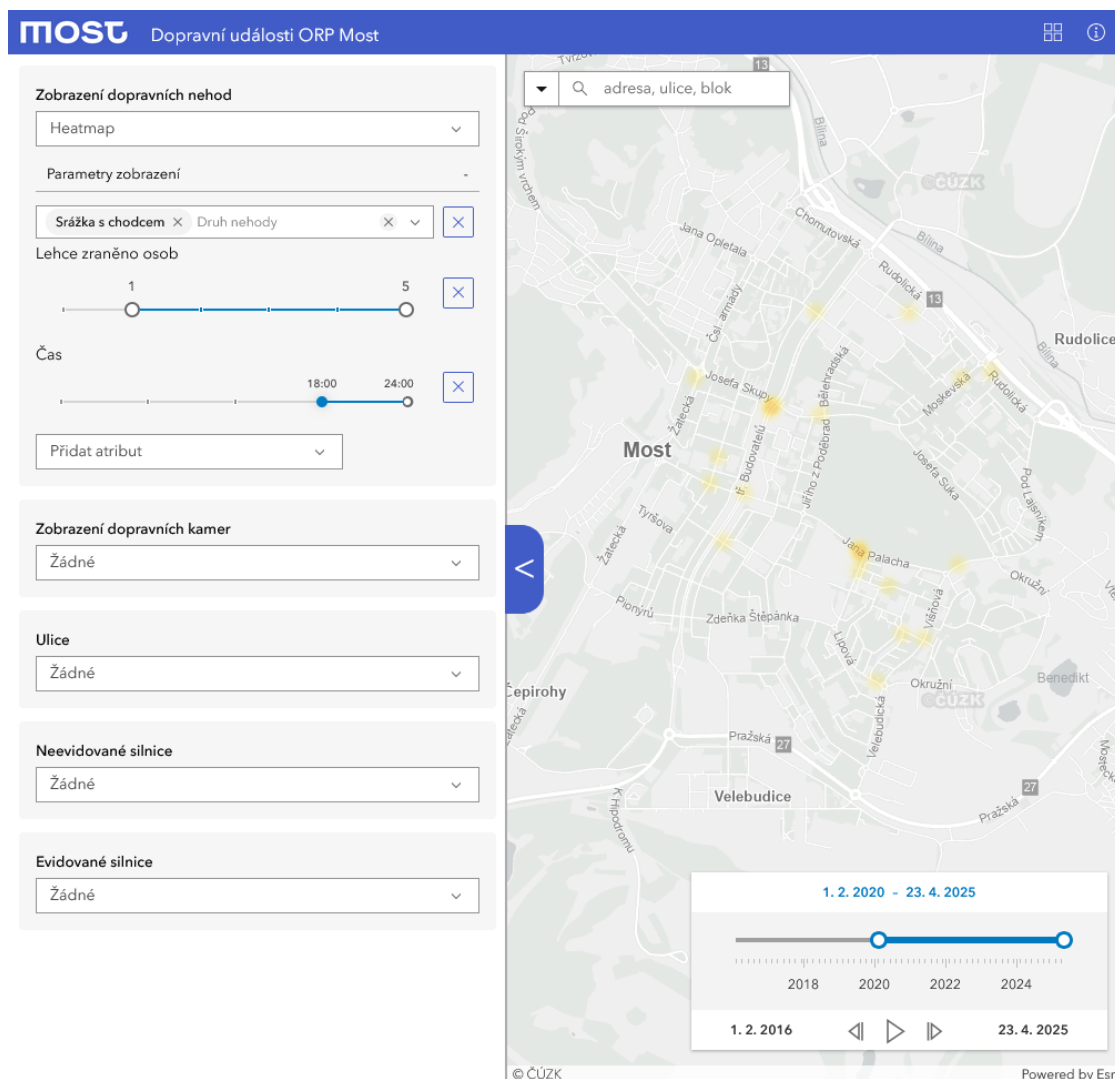
⁵MapServer ČÚZK: https://ags.cuzk.gov.cz/arcgis/rest/services/ZABAGED_POLOHOPIS/MapServer

⁶ArcGIS Maps SDK for JavaScript: <https://developers.arcgis.com/javascript/latest/>

⁷<https://developers.arcgis.com/javascript/latest/components-transition-plan/>

⁸Calcite components: <https://developers.arcgis.com/calcite-design-system/components/>

Obsah a chování aplikace je možné měnit pomocí konfiguračního souboru, který je stejně jako v případě skriptu doprovázen popisem v souboru schema.json, který popisuje strukturu a očekávané hodnoty konfigurace. Datové vrstvy, jejich zobrazení, dostupné atributy pro filtrování, obsah vyskakovacích oken nebo výchozí vybrané hodnoty při otevření aplikace je možné měnit pouhou změnou konfigurace. Protože je konfigurační soubor automaticky zkopírován do cílové složky sestavené aplikace, stačí provést změny pouze na tomto místě a aplikaci není nutné znovu sestavovat a nasazovat.

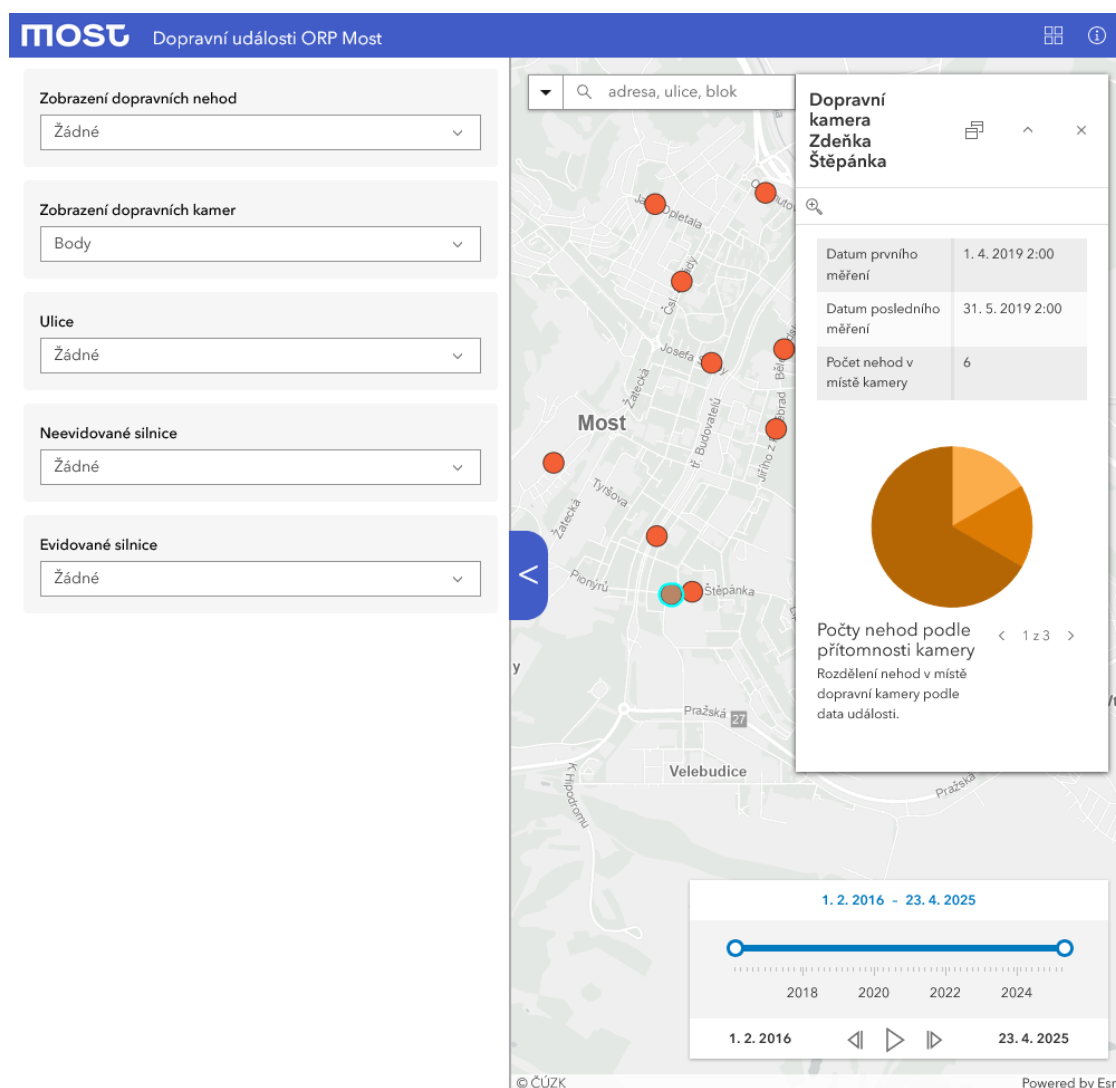


Obrázek 6.3: Ukázka zobrazení dopravních nehod v oblasti města Most v posledních pěti letech pomocí teplotní mapy. Nehody jsou omezeny na srážky s chodci ve večerních hodinách, při kterých došlo k lehkým zraněním.

Dopravní nehody

Vrstvu dopravních nehod je možné zobrazit pomocí samostatných bodů nebo teplotní mapy. Zobrazená data je možné filtrovat na základě atributů v sekci parametry zobrazení. Přítomny jsou tři typy atributů – numerické, kategorické a časové. Numerické atributy jsou

filtrovány výběrem intervalu, jehož maximální hodnota je omezená maximální hodnotou nalezenou v datech. Filtrování kategoričnými atributy probíhá v režimu AND při výběru jedné hodnoty nebo OR při výběru více hodnot. Čas je v datech dopravních nehod uchován jako celé číslo, a tak je nutné provádět úpravy mezi jeho hodnotou a zobrazovanou reprezentací. Filtrování je možné provádět pomocí jednoho nebo více atributů s vybranými hodnotami. Atributy je možné volně přidávat, odebírat a upravovat, což ihned ovlivňuje zobrazená data. Vrstvu je možné filtrovat i podle data, což je zajištěno mapovou komponentou TimeSlider. Maximální a vybrané hodnoty numerických atributů se přizpůsobí novému maximu pro daný atribut.



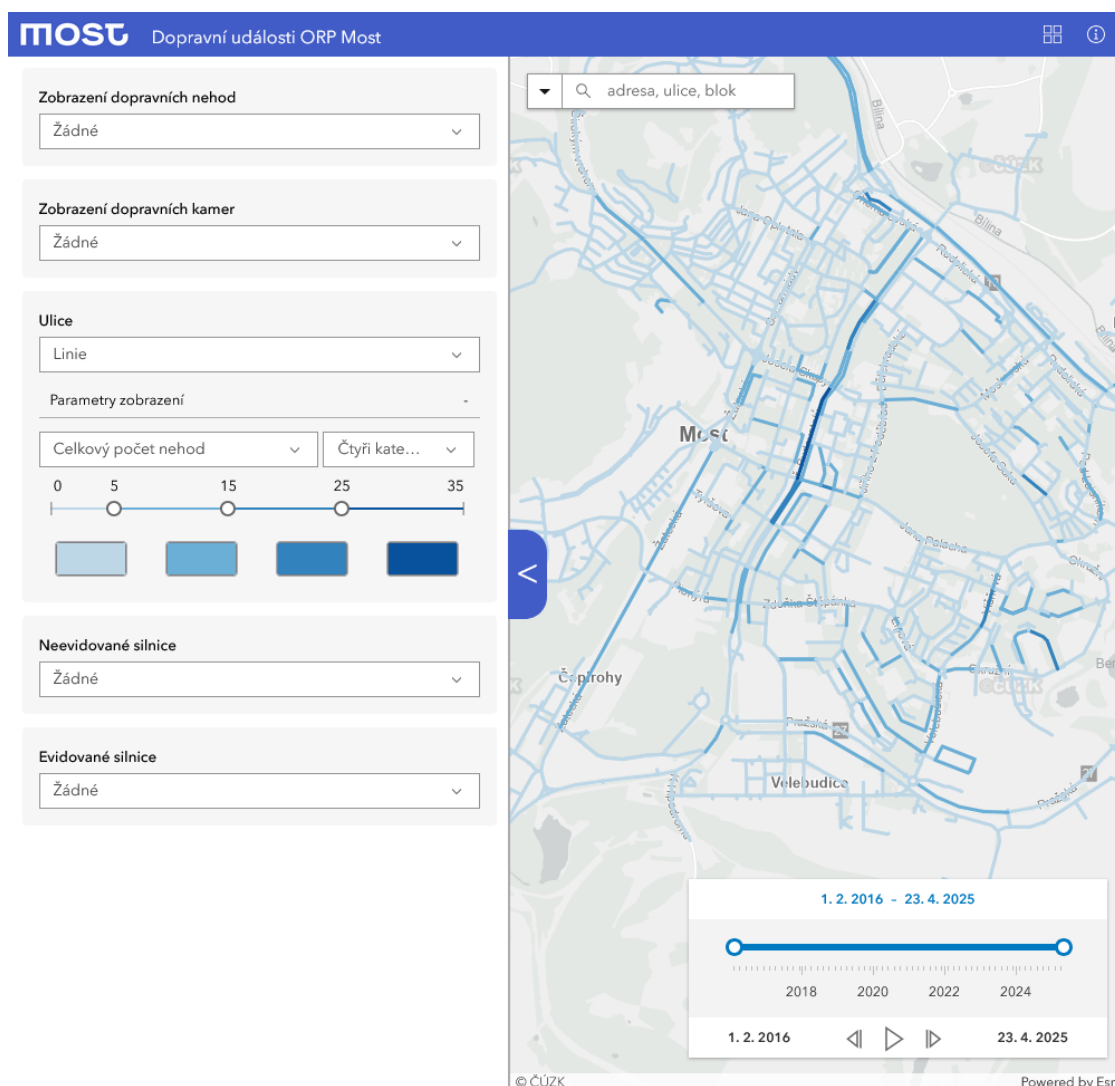
Obrázek 6.4: Ukázka bodového zobrazení vrstvy dopravních kamer města Most. Vyskakovací okna bodů kamer obsahují grafy nehodovosti, počtů průjezdů a počtů dopravních přestupků.

Po kliknutí na záznam dopravní nehody v mapě se objeví vyskakovací okno s detaily. Aby byl zobrazený text pochopitelný bez nutnosti hledání významu různých čísel a kódů, je prováděno jejich mapování na odpovídající textové reprezentace. Toto je zajištěno Arcade

výrazy definovanými v konfiguračním souboru. Ukázku zobrazení dopravních nehod pomocí teplotní mapy je možné vidět na obrázku 6.3.

Dopravní kamery

Vrstvu dopravních kamer je možné zobrazit pomocí bodů na mapě. Vyskakovací okno obsahuje informace o rozmezí, ve kterém byla měření prováděna a o celkovém počtu nehod. Počty nehod jsou zobrazeny v koláčovém grafu, ve kterém jsou kategorizovány na nehody před instalací kamery, v období instalované kamery a po odstranění kamery. Mimo graf nehod jsou zde zobrazeny i sloupcové grafy průměrného počtu průjezdů a průměrného počtu dopravních přestupků pro dny v týdnu. Ukázku zobrazení dopravních kamer je možné vidět na obrázku 6.4.



Obrázek 6.5: Ukázka liniového zobrazení vrstvy ulic města Most. Ulice jsou rozděleny do čtyř barevně odlišených tříd podle celkového počtu dopravních nehod, které se na nich odehrály.

Silniční komunikace

Vrstvu silničních komunikací je možné zobrazit pomocí liniového zobrazení. Zde jsou dílčí úseky silnic obarveny podle třídy, do které spadají. Třída úseku je určena podle hodnoty vybraného atributu. Mimo vybraného atributu je možné měnit i počet tříd, do kterých se silnice dělí, jejich intervaly a jejich barevnou reprezentaci na mapě. Maximální povolená hodnota je opět vázána k maximální hodnotě daného atributu v záznamech silnic. Ve vyskakovacích oknech jsou vizualizovány trendy nehodovosti a následků za posledních pět let pomocí spojnicových grafů. Ukázkou liniového zobrazení silnic je možné vidět na obrázku 6.5.

Kapitola 7

Testování vytvořených nástrojů

Následující kapitola se věnuje poskytnuté zpětné vazbě z testování vytvořených nástrojů a jejich možným úpravám a rozšířením. Testování proběhlo v rámci tří uživatelských skupin. První dvě skupiny zahrnují zaměstnance města Most z pohledu využití webové aplikace a z pohledu správy nástrojů a práce s daty. Poslední skupinou jsou uživatelé z veřejnosti, kteří s používáním podobných webových aplikací mají menší zkušenosti, a dokážou tak poskytnout odlišný pohled. Hlavními položkami zpětné vazby jsou využitelnost nových nástrojů pro dosažení požadovaných cílů, porovnání nových nástrojů s jejich předchůdci a případné náměty na rozšíření funkcionality.

7.1 Uživatelské testování

Uživatelské testování obou vytvořených nástrojů bylo provedeno interně mezi zaměstnanci magistrátu města Most a jeho výsledky a získané poznatky byly komunikovány přes Ing. Kamila Nováka. První skupinou byli zaměstnanci z oddělení GIS z odboru informačních systémů, kteří oba nástroje testovali z pohledu zasazení do systému, konfigurovatelnosti a údržby. Druhá skupina zaměstnanců z odboru správních činností se soustředila pouze na testování webové aplikace pro analýzu dopravních dat. Obě skupiny tak testovaly pouze ty nástroje, které budou součástí jejich pracovní náplně. S několika členy veřejnosti bylo navíc provedeno doplňující testování webové aplikace pro získání jiného pohledu.

Testování nástrojů zaměstnanci magistrátu města Most

Skript byl průběžně testován panem Novákem v rámci celého řešení práce a na základě zpětné vazby byly iterativně prováděny úpravy. Splněny tak byly všechny požadavky na kompatibilitu se systémy ArcGIS, flexibilitu skriptu pomocí konfigurace, logování běhu programu, a navíc bylo umožněno i automatické získávání dat. K vytvořenému skriptu tak nejsou žádné výhrady ze strany města Most. Funkčnost skriptu byla otestována i pro jiné územní celky a výsledná data byla prezentována účastníkům konference Excel@FIT [68].

S ovládáním webové aplikace a jejími možnostmi byli uživatelé seznámeni emailem i na osobním setkání s panem Novákem. Uživatelé postupně celou aplikaci prošli a jejich relevantní připomínky byly shrnuty v rámci zpětné vazby, kterou po dokončení testování pan Novák odeslal. Ve výsledcích byly navíc zahrnuty i komentáře k možným budoucím rozšířením a úpravám. Hlavní body zpětné vazby byly následující:

- Část uživatelů hodnotila pozitivně možnost filtrování nehod podle libovolného množství atributů. Druhá část uživatelů by však preferovala přednastavené scénáře jako například nehody, v nichž hrál roli alkohol a jedním z účastníků byl autobus.
- Uživatelům často dělal problém filtrování nehod pomocí atributů, kdy vybrali atributy, ale netušili, že mají navíc specifikovat hodnotu vybraných atributů. Některým uživatelům dělal problém orientovat se ve větším množství filtrů.
- Někteří uživatelé by preferovali, kdyby se při změně typu zobrazení zachovaly nastavené filtry.
- Některým uživatelům chyběly jednotky u filtrování pomocí číselných atributů.
- Filtrování pomocí komponentu TimeSlider nedělalo uživatelům problém a bylo hodnoceno pozitivně. Někteří uživatelé by uvítali možnost přizpůsobit časové rozmezí pomocí výběru přesného data v kalendáři.
- Někteří uživatelé by uvítali i filtrování grafů ve vyskakovacích oknech pomocí komponentu TimeSlider.
- Některým uživatelům chyběly grafy denních a měsíčních průjezdů u dopravních kamer.
- Některým uživatelům se nelíbilo rozdělení liniových dat na ulice, evidované silnice a neevidované silnice. Toto rozdělení je pouze dočasné a bude upraveno, jakmile bude nalezen vhodnější zdroj dat silnic.
- Většina uživatelů byla spokojená s vizualizací nehod pomocí teplotní mapy. Zobrazení by ještě doplnili legendou a rozšířili by barevnou škálu, aby bylo možné odlišit místa s velkým počtem nehod i při malém měřítku.
- Někteří uživatelé hodnotí pozitivně podrobné zobrazení všech informací ve vyskakovacím okně nehody, někteří by raději dynamicky zobrazovali jen relevantní informace pro danou nehodu.

Zmíněné body obsahují hlavně požadavky na úpravu konfigurace a rozšíření aktuálního řešení. Celkově byly uživatelé s výslednou aplikací spokojeni a po objasnění otázek se v aplikaci rychle zorientovali. K dispozici mají mimo grafy průjezdů a překročení rychlosti všechna potřebná data pro jejich práci. Oddělení GIS má v plánu začlenit aplikaci do agend uživatelů, provést jejich školení a představit možnosti dalšího rozvoje. Následně bude začleněnou aplikaci upravovat podle požadavků uživatel a případně na ní bude stavět nová řešení.

Testování webové aplikace uživateli z veřejnosti

Z veřejnosti bylo vybráno deset uživatel, kteří s vytvořenou aplikací ani jejím předchůdcem nemají žádné zkušenosti. Každý z uživatel měl za úkol splnit seznam pěti úkolů, které postupně využily všechny části aplikace. Seznam úkolů byl následující:

1. Zobrazit nehody se špatnými rozhledovými poměry v čase mezi 20:00 a 24:00 v posledních pěti letech.
2. Zobrazit teplotní mapu dopravních nehod s chodci.

3. Vyhledat počet průměrných průjezdů v pátek pro libovolnou kameru.
4. Zobrazit počty nehod na ulicích města Most v posledních dvanácti měsících rozdělených na tři intervaly a upravit barvu nejvyššího intervalu.
5. Změnit podkladovou mapu a vyhledat ulici Zahrádkářská ve městě Most.

Při plnění těchto úkolů bylo sledováno chování uživatelů a problematické situace, ve kterých si uživatel nebyl jistý svým postupem nebo mu dosažení chtěného výsledku dělalo značné problémy. Jako nejvíce problematické se ukázaly úkoly tři a čtyři. Uživatelé zde strávili výrazně vyšší množství času než při ostatních úkolech a v některých případech potřebovali nasměrovat správným směrem. Zobrazení dopravních kamer včetně jejich detailu nepředstavovalo problém, avšak tlačítko pro listování ve vyskakovacím okně se na první pokus podařilo najít jen minimu účastníků. V úkolu čtyři působilo zmatení filtrování nehod pro posledních dvanáct měsíců, při kterém všichni zúčastnění začali nastavovat interval v TimeSlider komponentu v pravém dolním rohu. I přes tyto obtíže byly všechny úkoly úspěšně splněny v rozmezí šesti až devíti minut.

7.2 Možné úpravy a rozšíření aktuální implementace

Skript pro automatickou aktualizaci dat je navržen takovým způsobem, aby umožňoval rozšíření novými třídami. Přímo se tak nabízí možnost přidat podporu zpracování pro další typy dat mimo CSV a XLS. Dalším rozšířením mohou být i nové scraper třídy pro získání nových dopravních dat. V rámci získávání dat z API může být nutné přidat nové třídy pro jejich zpracování, v případě že budou data poskytována v odlišném formátu. Pro filtrování za využití jiných tvarů, než je aktuálně očekávaný polygon – například pomocí multipolygon geometrie, je nutné přidat dodatečnou funkcionalitu. Rozšířit je možné i počet Arcade skriptů pro výpočet nových hodnot nebo se tyto skripty pokusit kompletně nahradit požadavky z webové aplikace na feature služby.

Webovou aplikaci je možné rozšířit o některé body zpětné vazby uživatelů z magistrátu města Most zmíněné v podsekcí 7.1. Obecně by bylo vhodné vylepšit uživatelskou zkušenost, aby uživatelé jednodušeji našli požadované informace. Místem pro zlepšení je například TimeSlider, jehož efekt by bylo vhodné buď rozšířit na všechny zobrazené vrstvy nebo automaticky zobrazit jen při vrstvě dopravních nehod. Dostupná zobrazení bodů a linií mohou být rozšířena například o zobrazení polygonů pro vizualizaci záznamů určité oblasti.

Celkově byly oba nástroje vytvořeny takovým způsobem, aby umožňovali využití různými územními celky. Toto znamená že je sice možné provést úpravy výchozí funkcionality, avšak specifické požadavky těchto celků se budou pravděpodobně lišit a je na nich, aby tento základ rozšířili podle svých potřeb.

Kapitola 8

Závěr

Cílem práce bylo vytvořit nové nástroje pro město Most umožňující automatickou aktualizaci, vizualizaci a analýzu dat dopravních nehod a dat měření dopravních kamer. Tyto nástroje byly úspěšně vytvořeny a jejich podoba byla průběžně konzultována s referentem města Most Ing. Kamilem Novákem. Výsledné nástroje byly otestovány zaměstnanci magistrátu města Most a zhodnoceny jako vhodné pro zapojení do agend uživatelů.

Nejdříve byly prostudovány dopravní data, jejich zdroje, metody pro jejich zpracování, způsoby vizualizace a jejich využití v rámci měst za pomoci GIS. Následně byly popsány další způsoby vizualizace včetně možností pro zobrazení multidimenzionálních dat. Probrány byly i technologie pro tvorbu vizualizací na webu společně s populárními zástupci knihoven. Dále byl analyzován stav využívaného skriptu pro aktualizaci dat a webové aplikace Silniční doprava, která zajišťovala vizualizaci dat dopravních nehod a dat z dopravních kamer. Byly popsány nedostatky těchto nástrojů a zjištěny požadavky na jejich náhradu včetně analýzy dostupných zdrojů dat. Na základě požadavků byl vytvořen návrh datového modelu, architektury systému, skriptu pro aktualizaci dat a webové aplikace. Nové nástroje byly implementovány podle tohoto návrhu a současně umožňují využití jakýmkoliv dalšími územními celky. Vytvořené nástroje byly průběžně testovány Ing. Kamilem Novákem a na závěr byly otestovány i ostatními zaměstnanci magistrátu města Most. Nástroje byly na základě zpětné vazby vyhodnoceny jako vhodná náhrada za jejich předchůdce a vznikl plán jejich integrace do pracovní činnosti zaměstnanců. Současně byly diskutovány možné oblasti jejich úpravy a rozšíření. S vybranými členy veřejnosti bylo provedeno uživatelské testování pro doplnění pohledu nezkušených uživatel. Výsledná webová aplikace je veřejně dostupná na stránkách města Most. Výsledky práce byly prezentovány na konferenci Excel@FIT [68].

Celkově byly vytvořeny nástroje, které umožní zaměstnancům magistrátu města Most pracovat s aktuálními daty dopravních nehod v kompletní podobě pomocí veřejně dostupné webové aplikace, která nabízí rozšířené možnosti pro analýzu dopravních dat oproti svému předchůdci. Modularita a konfigurovatelnost nástrojů umožňuje upravit nebo rozšířit aktuální řešení při budoucích změnách zdrojů a zpracování dat. Nástroje je možné využít i pro jiné územní celky. Skript pro automatickou aktualizaci dat je navíc možné využít i bez nutnosti vlastnit licenci ArcGIS. Na těchto základech může město Most nebo kdokoli jiný stavět vlastní řešení podle svých požadavků. V práci by bylo možné pokračovat rozšířením těchto základů, což by zahrnovalo přidání podpory pro další zdroje a typy dat a vylepšení uživatelské zkušenosti v rámci webové aplikace. Alternativně může být aktuální řešení rozšířeno přímo podle požadavků územního celku, který by o tyto nástroje projevil zájem.

Literatura

- [1] ADOBE. *What are TIFF files and how do you open them?* 2023. Dostupné z: <https://www.adobe.com/creativecloud/file-types/image/raster/tiff-file.html>. Accessed: 2025-01-12.
- [2] AGHASI, N. H. M. Application of GIS for Urban Traffic Accidents: A Critical Review. *Journal of Geographic Information System*, 2019. Dostupné z: <https://api.semanticscholar.org/CorpusID:133878898>.
- [3] ALBRIGHT, D. HISTORY OF ESTIMATING AND EVALUATING ANNUAL TRAFFIC VOLUME STATISTICS. *Transportation Research Record*, 1991. Dostupné z: <https://api.semanticscholar.org/CorpusID:108281503>.
- [4] ALOMAR, A.; ALRASHED, N.; ALTURAIKI, I. a ALTWAIJRY, H. How Visual Analytics Unlock Insights into Traffic Incidents in Urban Areas. In: MEISELWITZ, G., ed. *Social Computing and Social Media. Applications and Analytics*. Cham: Springer International Publishing, 2017, s. 378–393. ISBN 978-3-319-58562-8.
- [5] APACHE. *Render with SVG or Canvas*. 2023. Dostupné z: <https://apache.github.io/echarts-handbook/en/best-practices/canvas-vs-svg/>. Accessed: 2025-04-15.
- [6] BABYLON.JS. *Babylon.js Documentation*. 2025. Dostupné z: <https://doc.babylonjs.com/>. Accessed: 2025-04-17.
- [7] BELLINGER, G.; CASTRO, D. a MILLS, A. *Data, Information, Knowledge, and Wisdom*. 2004. Dostupné z: <https://homepages.dcc.ufmg.br/~amendes/SistemasInformacaoTP/TextosBasicos/Data-Information-Knowledge.pdf>. Accessed: 2025-01-16.
- [8] BIKAKIS, N. *Big Data Visualization Tools*. 2023. Dostupné z: <https://arxiv.org/abs/1801.08336>.
- [9] BROTOISWORO, P. a SUNIO, V. Exploring the Utilization of Crash Data from Waze and Twitter for Official National Crash Database. In: Září 2019.
- [10] BUTLER, H.; DALY, M.; DOYLE, A.; GILLIES, S.; SCHAUB, T. et al. *The GeoJSON Format RFC 7946*. RFC Editor, srpen 2016. Dostupné z: <https://doi.org/10.17487/RFC7946>.
- [11] CHANG, F.-R.; HUANG, H.-L.; SCHWEBEL, D. C.; CHAN, A. H. a HU, G.-Q. Global road traffic injury statistics: Challenges, mechanisms and solutions. *Chinese Journal of Traumatology*, 2020, sv. 23, č. 4, s. 216–218. ISSN 1008-1275. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S1008127520301437>.

- [12] CHART.JS. *Chart.js*. 2025. Dostupné z: <https://www.chartjs.org/docs/latest/>. Accessed: 2025-04-17.
- [13] CHEN, H. M. *Information Visualization*. American Library Association, April 2017. Library Technology Reports. Dostupné z: <https://journals.ala.org/index.php/ltr/issue/view/633>.
- [14] CHEN, W.; GUO, F. a WANG, F.-Y. A Survey of Traffic Data Visualization. *IEEE Transactions on Intelligent Transportation Systems*, 2015, sv. 16, č. 6, s. 2970–2984.
- [15] CUBE. *Front-end data tools for application developers*. 2025. Dostupné z: <https://awesome.cube.dev/>. Accessed: 2025-04-17.
- [16] D3.JS. *What is D3?* 2024. Dostupné z: <https://d3js.org/what-is-d3>. Accessed: 2025-04-16.
- [17] DJAHEL, S.; DOOLAN, R.; MUNTEAN, G.-M. a MURPHY, J. A Communications-Oriented Perspective on Traffic Management Systems for Smart Cities: Challenges and Innovative Approaches. *IEEE Communications Surveys Tutorials*, 2015, sv. 17, č. 1, s. 125–151.
- [18] DZEMYDA, G. a KURASOVA, O. *Multidimensional Data Visualization: Methods and Applications*. Květen 2015. ISBN 978-1-4419-0235-1.
- [19] ESRI. *ArcGIS*. 2025. Dostupné z: <https://www.esri.com/en-us/arcgis/geospatial-platform/overview>. Accessed: 2025-05-10.
- [20] ESRI. *Overview of georeferencing*. 2025. Dostupné z: <https://pro.arcgis.com/en/pro-app/latest/help/data/imagery/overview-of-georeferencing.htm>. Accessed: 2025-01-14.
- [21] ESRI. *Popup charts for clusters*. 2025. Dostupné z: <https://developers.arcgis.com/javascript/latest/sample-code/featurereduction-cluster-popup-chart/>. Accessed: 2025-05-12.
- [22] ESRI. *What is geocoding?* 2025. Dostupné z: <https://desktop.arcgis.com/en/arcmap/latest/manage-data/geocoding/what-is-geocoding.htm>. Accessed: 2025-01-14.
- [23] EUROPEAN COMMISSION. *Road safety statistics 2023 in more detail*. 2025. Dostupné z: https://transport.ec.europa.eu/background/road-safety-statistics-2023_en. Accessed: 2025-01-02.
- [24] FOLLA, K. a (NTUA), K. K. *Country Overview 2024: Czechia*. Brussels: European Commission, Directorate General for Transport, feb 2024. Dostupné z: https://road-safety.transport.ec.europa.eu/document/download/c005a276-1d9d-4907-a0a4-51360d593daa_en?filename=erso-country-overview-2024-czechia.pdf. Prepared within the EC Service Contract MOVE/C2/SER/2022-55/SI2.888215 with National Technical University of Athens (NTUA), SWOV Institute for Road Safety Research, and Kuratorium für Verkehrssicherheit (KFV).

- [25] FRANKE, L. a HAEHN, D. Modern Scientific Visualizations on the Web. *Informatics*, 2020, sv. 7, č. 4. ISSN 2227-9709. Dostupné z: <https://www.mdpi.com/2227-9709/7/4/37>.
- [26] FUJDIK, R.; MASEK, P.; MLYNEK, P.; MISUREC, J. a MUTHANNA, A. Advanced optimization method for improving the urban traffic management. In: *2016 18th Conference of Open Innovations Association and Seminar on Information Security and Protection of Information Technology (FRUCT-ISPIT)*. 2016, s. 48–53.
- [27] GARCÍA, S.; RAMÍREZ GALLEGÓ, S.; LUENGO, J.; BENÍTEZ, J. M. a HERRERA, F. Big data preprocessing: methods and prospects. *Big Data Analytics*, 2016, sv. 1, č. 1, s. 9. ISSN 2058-6345. Dostupné z: <https://doi.org/10.1186/s41044-016-0014-0>.
- [28] GOOGLE. *Waze Data Feed specifications*. 2023. Dostupné z: <https://support.google.com/waze/partners/answer/13458165?hl=en#zippy=%2Ctraffic-alerts%2Creliability-and-confidence-scores>. Accessed: 2025-04-11.
- [29] GOOGLE. *Apply for Waze for Cities*. 2025. Dostupné z: <https://support.google.com/waze/partners/answer/10453062>. Accessed: 2025-01-06.
- [30] GOOGLE. *Get traffic data with the Waze Data Feed*. 2025. Dostupné z: <https://support.google.com/waze/partners/answer/10618035?hl=en#zippy=%2Creported-incidents-included-in-the-waze-data-feed>. Accessed: 2025-01-06.
- [31] GUTIERREZ OSORIO, C. a PEDRAZA, C. Modern data sources and techniques for analysis and forecast of road accidents: A review. *Journal of Traffic and Transportation Engineering (English Edition)*, 2020, sv. 7, č. 4, s. 432–446. ISSN 2095-7564. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S209575642030101X>.
- [32] HANDSCOMBE, J. a YU, H. Q. Low-Cost and Data Anonymised City Traffic Flow Data Collection to Support Intelligent Traffic System. *Sensors*, 2019, sv. 19, č. 2. ISSN 1424-8220. Dostupné z: <https://www.mdpi.com/1424-8220/19/2/347>.
- [33] JAROSLAV ZENDULKA, R. L. a. I. R. *Získávání znalostí z databází: Studijní opora*. Fakulta informačních technologií, 2009.
- [34] KHAN, H. *HTML5 Canvas Vs SVG — Which Design Technology to Choose For Your Applications?* 15. September 2020. Dostupné z: <https://medium.com/illumination/html5-canvas-vs-svg-which-design-technology-to-choose-for-your-applications-633c5735efe>. Accessed: 2025-04-15.
- [35] KOVÁČOVÁ, L. *Stupně otevřenosti otevřených dat a česká legislativa*. 20. November 2024. Dostupné z: <https://opendata.gov.cz/informace:stupn%C4%9B-otev%C5%99enosti-datov%C3%BDch-sad>. Accessed: 2025-01-05.
- [36] LEAFLET. *Leaflet API reference*. 2025. Dostupné z: <https://leafletjs.com/reference.html>. Accessed: 2025-04-17.
- [37] LEDUC, G. *Road Traffic Data: Collection Methods and Applications*, Leden 2008.

- [38] LEMONDE, C.; ARSENIO, E. a HENRIQUES, R. Integrative analysis of multimodal traffic data: addressing open challenges using big data analytics in the city of Lisbon. *European Transport Research Review*, 2021, sv. 13, č. 1, s. 64. ISSN 1866-8887. Dostupné z: <https://doi.org/10.1186/s12544-021-00520-3>.
- [39] LENKEI, Z. *Crowdsourced traffic information in traffic management: Evaluation of traffic information from Waze*. 2018. Diplomová práce. KTH Royal Institute of Technology. Dostupné z: <https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1264077>. Accessed: 2025-01-05.
- [40] LI, Q. Overview of Data Visualization. In: *Embodying Data: Chinese Aesthetics, Interactive Visualization and Gaming Technologies*. Singapore: Springer Singapore, 2020, s. 17–47. ISBN 978-981-15-5069-0. Dostupné z: https://doi.org/10.1007/978-981-15-5069-0_2.
- [41] MACKLON, F.; VIGGIATO, M.; ROMANOVA, N.; BUZON, C.; PAAS, D. et al. A Taxonomy of Testable HTML5 Canvas Issues. *IEEE Transactions on Software Engineering*, 2023, sv. 49, č. 6, s. 3647–3659.
- [42] MAI TAN, H.; PHAM NGUYEN, H.-N.; LONG, N. X. a MINH, Q. T. Mining Urban Traffic Condition from Crowd-Sourced Data. *SN Computer Science*, 2020, sv. 1, č. 4, s. 225. ISSN 2661-8907. Dostupné z: <https://doi.org/10.1007/s42979-020-00244-6>.
- [43] MARINACCI, J. *HTML Canvas Deep Dive*. Josh On Design.com, 2012. Dostupné z: <https://freecomputerbooks.com/HTML-Canvas-Deep-Dive.html>. Accessed: 2025-01-17.
- [44] MCSHANE, C. The Origins and Globalization of Traffic Control Signals. *Journal of Urban History*, 1999, sv. 25, č. 3, s. 379–404. Dostupné z: <https://doi.org/10.1177/009614429902500304>.
- [45] MDN. *Canvas tutorial*. 2024. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/API/Canvas_API/Tutorial. Accessed: 2025-04-15.
- [46] MDN. *Introducing SVG from scratch*. 2024. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/SVG/Tutorials/SVG_from_scratch. Accessed: 2025-04-15.
- [47] MDN. *WebGL tutorial*. 2024. Dostupné z: https://developer.mozilla.org/en-US/docs/Web/API/WebGL_API/Tutorial. Accessed: 2025-04-15.
- [48] MDN. *Building up a basic demo with Babylon.js*. 2025. Dostupné z: https://developer.mozilla.org/en-US/docs/Games/Techniques/3D_on_the_web/Building_up_a_basic_demo_with_Babylon.js. Accessed: 2025-04-17.
- [49] MINISTERSTVO DOPRAVY. *Centrální registr vozidel*. 2025. Dostupné z: <https://md.gov.cz/Statistiky/Silnicni-doprava/Centralni-registr-vozidel>. Accessed: 2025-01-15.

- [50] MISEK, J. *Informace o otevřených datech*. 2022. Dostupné z: <https://opendata.gov.cz/informace:start>. Last updated: 2022-09-28, Accessed: 2025-01-05.
- [51] NAYAK, A. a GOYAL, K. Traffic modeling and accidental data analysis using GIS: A Review. *IOP Conference Series: Earth and Environmental Science*, 2024, sv. 1327, s. 012028. 1st International Conference on Sustainable Experimentation and Modelling in Civil Engineering 2023, 10/08/2023 - 11/08/2023, Phagwara, India.
- [52] OBSERVABLE. *Why Plot?* 2024. Dostupné z: <https://observablehq.com/plot/why-plot#plot-builds-on-d3>. Accessed: 2025-04-16.
- [53] QGIS. *QGIS Documentation*. 2025. Dostupné z: <https://docs.qgis.org/3.34/en/docs/index.html>. Accessed: 2025-01-11.
- [54] RITTER, N. a RUTH, M. The GeoTiff data interchange standard for raster geographic images. *International Journal of Remote Sensing*. Taylor & Francis, 1997, sv. 18, č. 7, s. 1637–1647. Dostupné z: <https://doi.org/10.1080/014311697218340>.
- [55] RUSSELL, E. *9 things to know about Google's maps data: Beyond the Map*. 30. September 2019. Dostupné z: <https://mapsplatform.google.com/resources/blog/9-things-know-about-googles-maps-data-beyond-map/>. Accessed: 2025-01-05.
- [56] SHAHZAD, F.; SHELAMI, T. R.; SHAKSHUKI, E. M. a SHAIKH, O. A Review of Latest Web Tools and Libraries for State-of-the-art Visualization. *Procedia Computer Science*, 2016, sv. 98, s. 100–106. ISSN 1877-0509. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S187705091632138X>. The 7th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2016)/The 6th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH-2016)/Affiliated Workshops.
- [57] SHAITO, M.; ELMASRI, R. a LEVINE, D. Comparison of Map Visualization Techniques Used for Spatial and Spatio-Temporal Data: An Analytical Survey Applied to COVID-19 Data. *Medical Research Archives*, 2022, sv. 10, č. 9. ISSN 2375-1924. Dostupné z: <https://esmed.org/MRA/mra/article/view/3072>.
- [58] SIDI, F.; SHARIAT PANAHY, P. H.; AFFENDEY, L. S.; JABAR, M. A.; IBRAHIM, H. et al. Data quality: A survey of data quality dimensions. In: *2012 International Conference on Information Retrieval Knowledge Management*. 2012, s. 300–304.
- [59] SIUHI, S. a MWAKALONGE, J. Opportunities and challenges of smart mobile applications in transportation. *Journal of Traffic and Transportation Engineering (English Edition)*, 2016, sv. 3, č. 6, s. 582–592. ISSN 2095-7564. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S2095756416302690>.
- [60] SOUZA, A. M. de; BRENNAND, C. A.; YOKOYAMA, R. S.; DONATO, E. A.; MADEIRA, E. R. et al. Traffic management systems: A classification, review, challenges, and future perspectives. *International Journal of Distributed Sensor Networks*, 2017, sv. 13, č. 4, s. 1550147716683612. Dostupné z: <https://doi.org/10.1177/1550147716683612>.

- [61] THREE.JS. *Three.js Fundamentals*. 2024. Dostupné z: <https://threejs.org/manual/#en/fundamentals>. Accessed: 2025-04-16.
- [62] UNITED STATES GEOLOGICAL SURVEY. *What is a Geographic Information System (GIS)?* 2024. Dostupné z: <https://www.usgs.gov/faqs/what-a-geographic-information-system-gis>. Updated: November 27, 2024. Accessed: 2025-01-10.
- [63] VACA, J. Martin Jeřábek (Mapy.cz): Dávno už nejsme jen aplikací pro turisty. *Lupa.cz*, November 2023. Dostupné z: <https://www.lupa.cz/clanky/martin-jerabek-mapy-cz-davno-uz-nejsme-jen-aplikaci-pro-turisty/>. Accessed: 2025-01-05.
- [64] WebGL FUNDAMENTALS. *WebGL Fundamentals*. 2024. Dostupné z: <https://webglfundamentals.org/webgl/lessons/webgl-fundamentals.html>. Accessed: 2025-04-16.
- [65] WORLD HEALTH ORGANIZATION. *Road Safety*. 2025. Dostupné z: <https://www.who.int/health-topics/road-safety>. Accessed: 2025-01-15.
- [66] YUAN, H. a LI, G. A Survey of Traffic Prediction: from Spatio-Temporal Data to Intelligent Transportation. *Data Science and Engineering*, 2021, sv. 6, č. 1, s. 63–85. ISSN 2364-1541. Dostupné z: <https://doi.org/10.1007/s41019-020-00151-z>.
- [67] ČERNÝ, P. Mrtví na cestách. *Statistika a My*, 2024. Dostupné z: <https://statistikaamy.csu.gov.cz/mrtvi-na-cestach>. Accessed: 2024-12-13.
- [68] ŠTĚPÁNEK, M. Application for Traffic Event Analysis in the City of Most, 2025. Dostupné z: <https://excel.fit.vutbr.cz/submissions/2025/040/40.pdf>.