

Univerzita Pardubice
Fakulta ekonomicko-správní

Big data a ochrana dat ve strategiích a projektech Smart Cities

Diplomová práce

2024

Bc. Petr Hervert

Univerzita Pardubice
Fakulta ekonomicko-správní
Akademický rok: 2023/2024

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Bc. Petr Hervert**
Osobní číslo: **E22452**
Studijní program: **N0688A140007 Informatika a systémové inženýrství**
Specializace: **Informační a bezpečnostní systémy**
Téma práce: **Big data a ochrana dat ve strategiích a projektech Smart Cities**
Zadávající katedra: **Ústav systémového inženýrství a informatiky**

Zásady pro vypracování

Cílem práce je analyzovat problematiku big dat, práci s těmito daty a jejich ochranu ve strategiích a projektech Smart Cities ve vybraných světových městech a na základě těchto zjištění navrhnout modely a doporučení pro tuto problematiku.

Osnova:

- Základní pojmy a definice.
- Identifikace a porovnání přístupů a kritérií v dané oblasti.
- Analýza a vyhodnocení získaných dat.
- Návrh doporučení a opatření pro danou problematiku.
- Ověření a diskuse k dané problematice.

Rozsah pracovní zprávy: **cca 50 stran**
Rozsah grafických prací:
Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam doporučené literatury:

HENDL, Jan. *Big data: věda o datech – základy a aplikace*. Praha: Grada Publishing, 2021. Průvodce. ISBN 978-80-271-3031-3.
NEVIMA, Jan, Ingrid MAJEROVÁ, Marian LEBIEDZIK, et al. *SMART technologie pro zvyšování kvality života v městech a regionech*. [Průhonice]: Professional Publishing, 2023. ISBN 978-80-88260-62-2.
PELTON, Joseph N.; SINGH, Indu B. *Smart cities of today and tomorrow: Better technology, infrastructure and security*. Cham: Springer International Publishing, 2019. ISBN 978-3-319-95821-7.
PRAKASH, Kolla Bhanu, et al. (Eds.). *Big data analytics and intelligent techniques for smart cities*. Boca Raton: CRC Press, 2021. ISBN 978-0367753559.
SLAVÍK, Jakub. *Smart city v praxi: jak pomocí moderních technologií vytvořit město příjemné k životu a přátelské k podnikání*. Praha: Profi Press, 2017. ISBN 978-80-86726-80-9.

Vedoucí diplomové práce: **Ing. et Ing. Martin Lněnička, PhD.**
Ústav systémového inženýrství a informatiky

Datum zadání diplomové práce: **1. září 2023**
Termín odevzdání diplomové práce: **30. dubna 2024**

L.S.

prof. Ing. Jan Stejskal, Ph.D.
děkan

prof. Ing. Jitka Komárková, Ph.D.
garant studijního programu

PROHLÁŠENÍ AUTORA

Prohlašuji:

Práci s názvem *Big data a ochrana dat ve strategiích a projektech Smart Cities* jsem vypracoval samostatně. Veškeré literární prameny a informace, které jsem v práci využil, jsou uvedeny v seznamu použité literatury.

Byl jsem seznámen s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

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

V Pardubicích dne 30.04.2024

Bc. Petr Hervert v. r.

PODĚKOVÁNÍ

Chtěl bych poděkovat především vedoucímu mé práce panu Ing. et Ing. Martinu Lněničkovi, Ph.D. za odborné vedení a poskytnutí cenných rad a informací při zpracování mé diplomové práce. Dále bych chtěl poděkovat i ostatním vyučujícím fakulty ekonomicko-správní Univerzity Pardubice za získané znalosti, které dopomohli k vypracování této práce. Na závěr chci poděkovat rodině za podporu po celé době studia.

ANOTACE

Diplomová práce se zabývá problematiku big dat, prací s těmito daty a jejich ochranou ve strategiích a projektech Smart Cities ve vybraných světových městech. Práce obsahuje postup pro řešení této problematiky, který se skládá z analýzy a vyhodnocení získaných dat. Výstupem práce jsou modely a seznam ověřených doporučení.

KLÍČOVÁ SLOVA

big data, ochrana dat, Smart City, strategie, projekt, Python, metoda Delphi

TITLE

Big data and data protection in strategies and projects of Smart Cities

ANNOTATION

The diploma thesis deals with the topic of big data, work with these data and their protection in Smart Cities strategies and projects in selected world cities. The thesis contains a procedure for solving this issue, which consists of the analysis and evaluation of the obtained data. The output of the thesis are models and a list of validated recommendations.

KEYWORDS

big data, data protection, Smart City, strategy, project, Python, Delphi method

OBSAH

Úvod	13
1 Vymezení základních pojmů	14
1.1 Big data	14
1.1.1 Vymezení a definice	14
1.1.2 Ukládání a archivace big dat.....	15
1.1.3 Big data analytics.....	17
1.1.4 Trendy a další související termíny	17
1.2 Smart City	18
1.2.1 Vymezení a definice	18
1.2.2 Strategie a projekty Smart City.....	18
1.2.3 Měření a hodnocení Smart City	21
1.2.4 Smart City indexy	23
1.2.5 Smart City technologie	25
1.3 Ochrana dat	25
1.3.1 Ochrana osobních údajů.....	25
1.3.2 Kyberprostor a kyberbezpečnost	26
2 Identifikace a porovnání přístupů a kritérií v dané oblasti	27
2.1 Big data a Smart City	27
2.2 Požadavky na strategie a projekty Smart City	30
2.3 Přístupy a kritéria pro hodnocení strategií a projektů Smart Cities.....	32
2.4 Identifikace oblastí Smart City	33
3 Metodika a postup řešení	36
3.1 Výběr vzorku Smart Cities	36
3.2 Vyhledávání strategií a postup analýzy	37
3.3 Zápis dat do úvodní tabulky.....	38
3.4 Zápis dat do tabulky big dat.....	41
4 Analýza a vyhodnocení získaných dat	43
4.1 Porovnání strategií	43
4.2 Název dokumentu	43
4.3 Výskyt big dat.....	44
4.4 Počty z regionů	45
4.5 Rok vydání.....	46
4.6 Rok platnosti	46
4.7 Forma dokumentu	47

4.8	Použití v oblastech	48
4.9	Analýza v Pythonu	48
4.9.1	Příprava dat	48
4.9.2	Popisná statistika	49
4.9.3	Korelace	50
4.9.4	Shluková analýza	52
4.10	Projekty Smart City	57
4.11	Ochrana dat ve strategiích	58
4.11.1	Singapur	59
4.11.2	Stockholm	60
5	Návrh doporučení a opatření pro danou problematiku	61
5.1	Problematika big dat v rámci Smart Cities	61
5.2	Problematika ochrany dat v rámci Smart Cities	62
6	Ověření a diskuse k dané problematice	63
6.1	Ověření pomocí metody Delphi	63
6.2	Konečný seznam doporučení	67
6.3	Diskuse	71
Závěr	73
Použitá literatura	74
Seznam příloh	79

SEZNAM OBRÁZKŮ A TABULEK

Seznam obrázků

Obrázek 1: Kategorie a indikátory pro hodnocení Smart City. Zdroj: Sharifi (2019, str. 1273).	22
Obrázek 2: Dimenze CIMI. Zdroj: Berrone a Ricart (2022, str. 11).	24
Obrázek 3: Oblast Smart City a big dat technologií. Zdroj: Hashem (2016, s. 749).....	28
Obrázek 4: Správa/analýza dat pro aplikace v cloudu. Zdroj: Khan et al. (2013, s. 382).	28
Obrázek 5: Klasifikace případů užití big dat ve Smart Cities. Zdroj: Lim et al. (2018, str. 92).	29
Obrázek 6: Doporučení pro big data aplikace ve Smart City. Zdroj: Lim et al. (2018, str. 92).	30
Obrázek 7: Metodika a postup řešení cíle práce. Zdroj: <i>Vlastní zpracování, draw.io</i>	36
Obrázek 8: Postup vyhledávání strategií a jejich analýzy. Zdroj: <i>Vlastní zpracování, draw.io</i>	38
Obrázek 9: Porovnání strategií měst. Zdroj: <i>Vlastní zpracování, MS Excel</i>	43
Obrázek 10: Názvy dokumentů. Zdroj: <i>Vlastní zpracování, MS Excel</i>	44
Obrázek 11: Počet výskytů s pojmem „Smart“ v názvu. Zdroj: <i>Vlastní zpracování, MS Excel</i>	44
Obrázek 12: Výskyt big dat ve strategiích. Zdroj: <i>Vlastní zpracování, MS Excel</i>	45
Obrázek 13: Porovnání počtů měst z regionů (s big daty proti bez big dat). Zdroj: <i>Vlastní zpracování, MS Excel</i>	45
Obrázek 14: Roky vydání. Zdroj: <i>Vlastní zpracování, MS Excel</i>	46
Obrázek 15: Roky platnosti strategií. Zdroj: <i>Vlastní zpracování, MS Excel</i>	47
Obrázek 16: Součty použití big dat v oblastech. Zdroj: <i>Vlastní zpracování, MS Excel</i>	48
Obrázek 17: Shlukování 28 měst úvodní tabulky. Zdroj: <i>Vlastní zpracování, PyCharm</i>	53
Obrázek 18: Shlukování 14 měst tabulky big dat. Zdroj: <i>Vlastní zpracování, PyCharm</i>	55

Seznam tabulek

Tabulka 1: Technologie Smart City. Zdroj: Kitchin (2016, str. 12).	25
Tabulka 2: Vybrané požadavky na systémy big dat. Zdroj: <i>Vlastní zpracování</i>	32
Tabulka 3: Souhrn oblastí pro koncept Smart City. Zdroj: <i>Vlastní zpracování</i>	34
Tabulka 4: Úvodní tabulka. Zdroj: <i>Vlastní zpracování, MS Excel</i>	39
Tabulka 5: Přehled měst a jejich strategií. Zdroj: <i>Vlastní zpracování</i>	40
Tabulka 6: Tabulka big dat. Zdroj: <i>Vlastní zpracování, MS Excel</i>	42
Tabulka 7: Forma dokumentu. Zdroj: <i>Vlastní zpracování</i>	47
Tabulka 8: Výstupy popisných statistik z 28 měst úvodní tabulky. Zdroj: <i>Vlastní zpracování</i>	49
Tabulka 9: Výstupy popisných statistik ze 14 měst tabulky big dat. Zdroj: <i>Vlastní zpracování</i>	50
Tabulka 10: Výstup korelace 28 měst úvodní tabulky. Zdroj: <i>Vlastní zpracování</i>	51
Tabulka 11: Výstup korelace 14 měst tabulky big dat. Zdroj: <i>Vlastní zpracování</i>	52
Tabulka 12: Rozřazení do shluků (úvodní tabulka). Zdroj: <i>Vlastní zpracování</i>	54
Tabulka 13: Rozřazení do shluků (tabulka big dat). Zdroj: <i>Vlastní zpracování</i>	55
Tabulka 14: Popis Shluku 0. Zdroj: <i>Vlastní zpracování</i>	56
Tabulka 15: Popis Shluku 1. Zdroj: <i>Vlastní zpracování</i>	56
Tabulka 16: Popis Shluku 2. Zdroj: <i>Vlastní zpracování</i>	57
Tabulka 17: Popis Shluku 3. Zdroj: <i>Vlastní zpracování</i>	57
Tabulka 18: Seznam expertů a jejich odbornost. Zdroj: <i>Vlastní zpracování</i>	64
Tabulka 19: Úvodní seznam doporučení pro řešenou problematiku. Zdroj: <i>Vlastní zpracování</i>	65
Tabulka 20: Konečný seznam doporučení pro řešenou problematiku. Zdroj: <i>Vlastní zpracování</i>	68

SEZNAM ZKRATEK A ZNAČEK

AI	Artificial Intelligence
BPO	Business Process Outsourcing
CBPR	Cross-Border Privacy Rules
CII	Critical Information Infrastructure
CIMI	Cities in Motion Index
CSV	Comma-separated values (hodnoty oddělené čárkami)
ČR	Česká republika
EB	Exabyte (Jednotka kapacity paměťových médií)
EU	Evropská unie
GCI	Global Cities Index
GDPR	General Data Protection Regulation
GPCI	Global Power City Index
GPS	Global Positioning System
HDI	Index lidského rozvoje (Human Development Index)
HW	Hardware
ICT	Informační a komunikační technologie
IT	Informační technologie
IoT	Internet of Things (Internet věcí)
NIST	National Institute of Standards and Technology
NoSQL	Not only Structured Query Language
PDF	Portable Document Format (Univerzální formát souborů)
PRP	Privacy Recognition for Processors
R	Programovací jazyk pro analýzu dat
SCG	Smart City Governments
SCI	Smart City Index

SQL Standardizovaný strukturovaný dotazovací jazyk

SW Software

XLSX Formát souboru excel

ÚVOD

V současném dynamickém a propojeném světě, charakterizovaném neustálým technologickým pokrokem a rozvojem, se města stále častěji stávají centrem inovací a digitální transformace. Koncept Smart City vznikl jako odpověď na narůstající urbanizaci a potřebu efektivnějšího řízení měst a jimi nabízených veřejných služeb s cílem zlepšit kvalitu života občanů. Tento koncept spojuje Informační a komunikační technologie (ICT) s infrastrukturou měst, vytvářející tak nové možnosti a výzvy v oblasti správy velkých objemů dat a jejich ochrany.

Zároveň s rostoucími objemy, druhy, a především rychlostí generování a zpracovávání dat se Smart Cities stále častěji musí ve svých strategiích a projektech vypořádat s big daty a současně nutností ochrany těchto dat v rámci projektů Smart Cities. V kontextu Smart Cities se big data stávají nezbytným nástrojem pro získávání hlubších znalostí a informací, které jsou klíčové pro účinnější správu a plánování rozvoje města.

Cílem této práce je zkoumání problematiky big dat a ochrany dat v rámci strategií a projektů Smart Cities, konkrétně zjištění, do jaké míry se strategie této problematice věnují a na základě toho navrhnout vhodná doporučení. V případě big dat se jedná o poměrně nový pojem, koncept Smart City je o několik let starší, ačkoliv se stále jedná o termín novodobé historie související s rozvojem ICT. Práce zkoumá vybraná světová města na základě prostudování strategických dokumentů souvisejících s konceptem Smart City. Získaná data pak zpracovává v programu Microsoft Excel a dále analyzuje pomocí modelů v programovacím jazyku Pythonu s využitím vhodných knihoven.

V první části této práce bude provedena analýza literatury související s tímto tématem. Kromě toho se bude zabývat i zkoumáním žebříčků a indexů souvisejících s touto problematikou. Další část diplomové práce se bude hlouběji zabývat výběrem vhodných měst a jejich strategií pro analýzu problematiky. Důležitou částí této diplomové práce bude zhodnocení těchto vybraných existujících strategií, a to jak pomocí základních ukazatelů, tak pomocí pokročilých postupů a modelů v Pythonu. V závěrečné části této diplomové práce budou po ověření experty metodou Delphi prezentována doporučení pro tuto problematiku.

1 VYMEZENÍ ZÁKLADNÍCH POJMŮ

1.1 Big data

1.1.1 Vymezení a definice

Význam big dat v posledních několika letech narůstá zejména díky růstu množství existujících dat, ale také proto, že nově vznikající data již nejsou vázána na ustálené struktury, např. jen na relační model dat. Data přicházejí v různorodých formátech, a není vždy vhodné přetvářet je do jednotného formátu, což je dáno odlišnostmi v současných aplikacích a jejich požadavcích (Hendl, 2021). Tento stav zároveň nastoluje otázku, jak big data přesně definovat. Americká společnost Gartner definuje big data jako data, jejichž velikost (v angličtině *volume*), rychlost nárůstu (*velocity*) a různorodost (*variety*) neumožňují zpracování pomocí známých a ověřených technologií v rozumném čase. Tyto 3 vlastnosti jsou označovány jako 3V a postupně k nim přibývají i další vlastnosti, jako např. nejistá věrohodnost (*veracity*) a vysoká hodnota (*value*) pro vlastníka dat, označováno jako 5V. Mimo zatím zmíněné k nim přibývají i další 2 v podobě limitované doby platnosti (*validity*) pro jejich využití a s tím spojená přechodná doba jejich ukládání (*volatility*) (Holubová et al., 2015, str. 19-20).

Jiné zdroje uvádí i jiné vlastnosti pro big data. Například podle serveru impact.com se uvádí 7V, a to již zmíněné velikost, rychlost nárůstu, různorodost, nejistá věrohodnost a vysoká hodnota. Mimo tyto pak dále zmiňuje vizualizaci (*visualization*) a variabilitu dat (*variability*) čili nesrovnalosti v datech (Moore, 2023). Arockia et al. (2017) pak definují 17V pro big data a Hussien (2020) dokonce 56V charakteristik, které definují big data. S ohledem na tyto skutečnosti lze tedy shrnout, že existující definice big data vychází primárně z těchto charakteristik, když jsou dále doplněné např. o nutnost využívání platform, nástrojů a služeb, které umožňují s big daty pracovat v rámci celého jejich životního cyklu od jejich získávání a přenosu až po vizualizaci a publikaci výstupů (Favaretto et al., 2020; Hendl, 2021; Holubová et al., 2015).

Big data pocházejí z mnoha zdrojů – příkladem můžou být systémy pro zpracování transakcí, databáze zákazníků, dokumenty, emaily, lékařské záznamy, záznamy o aktivitách na internetu, mobilní aplikace nebo sociální sítě. Dále sem patří také data generovaná stroji, jako jsou síťové a logovací soubory a data se senzorů výrobních strojů, průmyslových zařízení nebo zařízení internetu věcí. Patří sem i zastřešující pojmy, které jsou primárně založeny na využívání big dat z různých senzorů a zařízení v reálném čase, jako jsou např. chytré domácnosti (*smart homes*) a chytrá města/vesnice (*Smart Cities/villages*) (Botelho a Bigelow, 2022; Hendl, 2021). Kromě

dat z interních systémů prostředí big dat často zahrnují i externí data o spotřebitelích, finančních trzích, počasí a dopravních podmínkách, geografické informace, vědecké výzkumy a další. Obrázky, videa a zvukové soubory jsou také formou big dat a mnoho aplikací big dat zahrnuje proudová data, která jsou zpracovávána a shromažďována průběžně (Botelho a Bigelow, 2022). V neposlední řadě je nutné zmínit big data týkající se bezpečnosti, ať už jde o důvěrné a citlivé informace nebo data mající vliv na národní bezpečnost. Pro práci s těmito daty musí existovat specifické postupy a metody, které by však měly existovat i pro další sektory (Ikegwu et al., 2022).

1.1.2 Ukládání a archivace big dat

1.1.2.1 NoSQL databáze

Donedávna nejpoužívanějším systémem pro efektivní ukládání a správu dat byly tradiční databázové systémy, ať už relační, objektové, objektově relační atd., využívající zpravidla dotazovací jazyk Structured Query Language (SQL). Pro zpracování big dat však tyto systémy již nestačí a bylo nutné vytvořit nové modely vhodné pro big data, které jsou dnes souhrnně označovány jako NoSQL databáze. Dalšími možnými způsoby ukládání a archivace big dat jsou Hybridní databáze (kombinace různých, dříve oddělených přístupů) nebo také NewSQL databáze (funkcionalita jako u relačních databází rozšířená o škálovatelnost typickou pro NoSQL databáze) (Hendl, 2021; Holubová et al., 2015).

Obecně se o NoSQL databázích dá říct, že umí zpracovat obrovské objemy rychle se měnících nestrukturovaných dat jinými způsoby než relační databáze (SQL) s řádky a tabulkami. Důležitá je u NoSQL databází škálovatelnost. U NoSQL databází se jedná o horizontální škálování dat prostřednictvím horizontálního dělení napříč servery, v případě tradičních SQL databází se jednalo o vertikální škálování dat prostřednictvím zvyšování zatížení serveru (Azure, 2024). Databáze bez SQL odkazují na vysoce výkonná nerelační úložiště dat. Vynikají v jednoduchosti použití, škálovatelnosti, odolnosti a charakteristikách dostupnosti. Místo spojování tabulek normalizovaných dat ukládá NoSQL nestrukturovaná nebo částečně strukturovaná data, často ve dvojicích klíč-hodnota nebo v dokumentech JavaScript Object Notation (JSON). Databáze bez SQL obvykle neposkytují záruky Atomicity, Consistency, Isolation, Durability (ACID) nad rámec jednoho oddílu databáze. Služby s velkým objemem, které vyžadují dobu odezvy po sekundách, upřednostňují úložiště dat NoSQL (Microsoft Learn, 2023).

1.1.2.2 Cloud computing

Cloud computing je založen na využívání ICT a související výpočetních zdrojů s cílem jejich poskytování zákazníkům prostřednictvím různých obchodních modelů. Díky cloud computingu mohou jednotlivci a organizace získat přístup na vyžádání ke sdílenému fondu spravovaných a škálovatelných výpočetních zdrojů, jako jsou servery, úložiště a aplikace (Sunyaev, 2020). Cloud computing je tedy poskytování výpočetních služeb, jako jsou servery, úložiště, databáze, sítě a software (SW), analytika a další pomocí internetu (cloudu), přičemž se využívá s cílem poskytnutí flexibilních zdrojů, rychlejších inovací a úsporám z rozsahu. Dalším benefitem cloud computingu jsou platební modely, kdy se platí jen za využívaný čas a celkově i nižší počáteční investice v porovnání s jinými řešeními (Berisha et al., 2022).

Mezi formy nasazení cloudu patří veřejný, soukromý a hybridní cloud. Hlavní rozdíl spočívá v tom, kdo všechno má do daného cloudu přístup a jak je zajištěna jeho bezpečnost. Dále lze rozlišovat 3 základní modely poskytování služeb (Holubová et al., 2015, str. 187-188):

- *Software as a Service* (SaaS) – SW jako služba, kterou obvykle využívají přímo koncoví uživatelé dané aplikace;
- *Platform as a Service* (PaaS) – platforma jako služba, která je určena vývojářům, pro něž nabízí sadu nástrojů pro vývoj aplikací, nebo obecně pro provozování vlastního softwaru;
- *Infrastructure as a Service* (IaaS) – infrastruktura jako služba, která slouží pro poskytování infrastruktury, tj. obvykle poskytování robustního či jinak ne snadno dostupného hardwaru, typicky formou vizualizace.

V současnosti převládá využívání služeb cloud computingu pro potřeby ukládání a archivace big dat, když datová centra fungující na tomto principu nabízejí nejenom díky úsporám z rozsahu nákladově efektivnější, robustnější a často bezpečnější služby než interní řešení dané organizace (Yang et al., 2017).

1.1.2.3 Datové jezero

Datovým jezerem se rozumí masivně škálovatelné úložiště, které uchovává obrovské množství nezpracovaných dat v jejich nativní podobě, tzn. tak, jak přijdou. Datová jezera jsou typicky budována tak, aby byla schopna zpracovávat velké a rychle přicházející objemy dat. Data v datovém jezeru se stávají přístupnými v momentě jejich vytvoření. Datové jezera si lze představit jako velký datový fond, do kterého se přivádějí všechna nashromážděná historická

a nová data (strukturovaná, nestrukturovaná, polostrukturovaná z různých senzorů, zařízení atd.) v téměř reálném čase do jednoho jediného místa (Miloslavskaya a Tolstoy, 2016).

Dnes jsou datová jezera klíčovou komponentou infrastruktur pro ukládání a archivaci big dat. Mezi hlavní výzvy při budování datových jezer pak patří: 1) volba vhodné architektury, které bude flexibilní, aby mohla reagovat na různé typy big dat, 2) dotazovací jazyky a obecně volba metadat a popisu uložených dat, které umožní efektivní práci s big daty, 3) algoritmy machine learning nad datovými jezery, tzn., jak efektivně zpracovávat a využívat big data zde uložená, 4) metodiky pro řízení a správu dat, a 5) zajištění bezpečnosti datových jezer (Cuzzocrea, 2021).

1.1.3 Big data analytics

Big data analytics představuje použití pokročilých analytických technik pro velmi rozsáhlé a různorodé soubory velkých dat, které zahrnují různě strukturovaná data z různých zdrojů a v různých velikostech. Big data analytics je často složitý proces zkoumání velkých objemů dat s cílem odhalit informace – například skryté vzorce, korelace, tržní trendy a preference zákazníků – které mohou organizacím pomoci přijímat informovaná obchodní rozhodnutí. Big data analytics je formou pokročilé analytiky, která zahrnuje komplexní aplikace s prvky, jako jsou prediktivní modely, statistické algoritmy a analýzy typu "co když" (Robinson, 2023). V knize *Big Data Analytics and Intelligent Techniques for Smart Cities* je big data analytika popsána jako sofistikovaná analytická metoda, kterou lze použít pro zpracování souborů big dat. Analytický základ velkého vzorku dat umožňuje odhalit a zkoumat vzory, které lze využít pro rozvoj smart oblastí města (Prakash et al., 2021, str. 3).

1.1.4 Trendy a další související termíny

Co se týče trendů, Innovature BPO zmiňuje digitální transformaci, využití big dat pro pokrok ve výzkumu změny klimatu, analýzu v reálném čase, big data v lokálních obchodech, pomáhání firmám v rozvoji za pomoci big dat, využití big dat v medicíně, využití big dat v autonomních vozidlech, simulace za pomoci big dat nebo zpracování přirozeného jazyka pomocí big dat (Innovature BPO, 2023). Další zdroj, Exploding Topics, pak zmiňuje analýzu v reálném čase, Artificial Intelligence (AI) a machine learning pro automatizaci a analýzu big dat, datová jezera a data lakehouse (kombinace datových jezer a datových skladů) a jako poslední trend zmiňuje správu big dat a jak k ní přistupovat (Howarth, 2023).

Co se týče dalších termínů relevantních k big data problematice, ty zmiňuje například Tom Potanski na webu Devs Data. Dalšími termíny související s big daty jsou podle něj například

business intelligence, biometrika, cloud computing, datový vědec, vizualizace, neuronové sítě, komparativní analytika, data mining, NoSQL, R a Python, shluková analýza a fuzzy logika (Potanski, 2023). Pro problematiku big dat je propojení s novými trendy důležité proto, že big data a jejich potenciální přínosy nejsou stále plně využívány z důvodu chybějících technologií, algoritmů, metodik atd.

1.2 Smart City

1.2.1 Vymezení a definice

Město bylo vždy spojeno s technologiemi, když města poskytovala prostředí pro rozvoj nových technologií a byla budována s využitím těchto nových technologií. Tak je tomu i dnes, pouze technologie jsou stále vyspělejší (Lisdorf, 2019, str. 19). Smart City motivuje lidi k aktivitám, setkávání, rozhovorům a shromažďování na ulicích, v obchodech a na veřejných prostranstvích. Zároveň by se zde lidé měli cítit bezpečně proto, že kyberneticko-fyzická infrastruktura města je navržena záměrně s cílem vytvořit atmosféru důvěry, společenství a sdílené odpovědnosti (Barlow a Levy-Bencheton, 2018, str. 6). Koncept Smart City by měl přispívat k řešení otázek o správě měst jako kdo a jak je řídí a jaké aspekty města lze řídit, jak lze města plánovat, aby vytvářela komunity, řídila růst a efektivně rozdělovala zdroje a služby atd. Správa a řízení města jsou komplexní úkoly, které vyžadují koordinaci mezi městskou administrativou a obyvateli. Moderní doba přináší nové možnosti, jako jsou počítačové systémy a rozhodování založené na datech (Halegoua, 2020, str. 1-2).

Definice a charakteristiky konceptu Smart City se liší v závislosti na zdroji. Společné pro většinu z nich však je spoléhání se na ICT jako na hlavní komponentu, která má zajistit zvýšení kvality života pro občany, zpravidla prostřednictvím poskytovaných veřejných služeb. Často pak jsou Smart Cities považována za prostředí či ekosystémy, kde jsou implementovány ICT prvky ke sběru dat pro podporu, monitorování a zlepšení městské infrastruktury jako například v oblasti dopravy, řízení odpadů, energetiky a řešení nouzových událostí (Halegoua, 2020, str. 4-5). Další zdroj pak dále uvádí i další oblasti využití včetně např. veřejné bezpečnosti (Harrison a Donnelly, 2011, str. 2).

1.2.2 Strategie a projekty Smart City

Strategie je zapotřebí vždy, když cesta k určitému cíli není jednoduchá nebo když je třeba posoudit potřebné zdroje, jejich efektivní použití a vhodnou posloupnost kroků vedoucích ke splnění cíle. V organizacích s definovanou organizační strukturou mohou za celkovou strategii odpovídat vedoucí pracovníci, ale existují i samostatné strategie pro marketing, lidské

zdroje atd. (Freedman, 2015). Každé strategické prohlášení (dokument) musí začínat definicí cílů, kterých má strategie dosáhnout, na což pak navazují jednotlivé kroky, přidělení zdrojů, časový plán atd. (Collis a Rukstad, 2008, str. 3).

Smart City strategie je dokument každého města nebo obce, ve kterém je stanoveno individuální chápání a vlastní přístup ke Smart City. Zahrnuje komplexní plán, jak město přijímá a využívá technologie, data a inovativní řešení ke zvýšení kvality života svých obyvatel, zlepšení městských služeb a infrastruktury a podpoře udržitelnosti. Cílem takovéto strategie je konkrétní plán o tom, jak využívat digitální technologie a rozhodování založené na datech k účinnému a efektivnímu řešení městských problémů a zároveň hospodářský rozvoj a sociální blahobyt (Isarsoft GmbH, 2023). Strategie Smart Cities lze primárně rozlišovat podle toho, zda se týkají celé země či národa, nebo jsou zaměřeny spíše na místní úroveň, ať už jde o čtvrť, obec, město, metropolitní oblast nebo region. Strategie se dají rozdělit do dvou kategorií, a to soft (měkce) a hard (tvrdě) infrastrukturně orientované. Zatímco tvrdá infrastruktura se týká např. dopravy, přívodů vody a elektřiny nebo odpadu, tak měkká infrastruktura se týká lidí, a to především s ohledem na sociální a lidský kapitál, tzn. znalosti, začlenění, sociální inovace nebo sociální spravedlnost (Angelidou, 2014, str. 2-3). Strategie Smart Cities lze dělit také podle dalších kritérií, např. s ohledem na kulturně-historické pozadí a preference daného celku, ekonomická a finanční pravidla a rozpočty, zaměření jen na konkrétní oblasti (např. bezpečnost občanů ve městě) nebo skupiny obyvatel (např. studenty). Více informací lze nalézt např. ve zdrojích Barlow a Levy-Bencheton (2018), Halegoua (2020), Lisdorf (2019) nebo Slavík (2017).

Česká republika (ČR) má vypracovanou metodiku pro Smart Cities od Ministerstva pro místní rozvoj ČR s názvem *Metodika Smart Cities: Metodika pro přípravu a realizaci konceptu Smart Cities na úrovni měst, obcí a regionů*, která obsahuje pasáž, jak by měl strategický dokument vypadat. Originální dokument vyšel v roce 2015 a od té doby byla metodika jednou aktualizovaná, kdy aktuální verze je z roku 2018. Primárním cílem strategie Smart City je zajištění kvalitního života obyvatelům, kdy jsou jako nástroj využívány moderní technologie pro ovlivňování kvality života ve městě, a následně k dosahování hospodářských a sociálních cílů města (Ministerstvo pro místní rozvoj ČR, 2018, str. 4). Podle této metodiky není nutné mít zvláštní strategický dokument a prvky Smart City mohou být zapracována do standardní rozvojové strategie daného území. Rozhodne-li se samospráva zpracovat samostatnou strategii Smart City, pak by měl tento dokument být s hlavní rozvojovou strategií koordinován (ve smyslu zabránění duplicitě informací a roztržitosti projektů a jejich kontroly) a měl by obsahovat následující základní funkce (Ministerstvo pro místní rozvoj ČR, 2018, str. 14-15):

- představení města a konceptu Smart Cities;
- popis a zhodnocení výchozí situace;
- formulace strategických cílů
- analýza připravovaných rozvojových projektů;
- návrh nových rozvojových projektů;
- finanční zdroje pro realizaci rozvojových projektů;
- realizační plán (též akční plán);
- monitoring a evaluace strategie a akčního plánu.

Švýcarsko má pro změnu dokument zvaný *Smart City – guide to the implementation of Smart City initiatives in Switzerland*, který zmiňuje jak projekty v rámci Smart City, tak zmiňuje strategie, které rozděluje do čtyř typů a pojednává o tom, co by ve strategii mělo být. Norsko má dokument *Roadmap for smart sustainable cities and communities in Norway*, který poskytuje rady, samotnou osnovu pro dokument ale neposkytuje. Jak švýcarský, tak i norský dokument zmiňuje big data, v porovnání s Českou metodikou, která s tímto pojmem nepracuje. V rámci Evropské unie (EU) existuje Smart Cities Marketplace, kdy se jedná o organizaci spojující různé zúčastněné strany v oblasti Smart Cities a poskytuje best practices. Tato iniciativa podporovaná Evropskou komisí vydala v roce 2023 digitální brožuru s názvem *Smart City Guidance Package*, která poskytuje informace k problematice plánování Smart Cities.

Smart City strategie je zpravidla vydávána, resp. jsou v ní plánovány nějaké cíle a kroky k jejich dosažení, na nějaké časové období např. 5 nebo 10 let. Součástí strategie mohou být i konkrétní projekty, nicméně s ohledem na rychle se měnící podmínky, jsou tyto projekty zmíněny spíše obecně, bez přesnějšího časového plánu. Některá města však např. každý rok zveřejňují nějaký konkrétní strategický plán pro projekty Smart Cities, kde už jsou probíhající nebo plánované projekty blíže specifikovány.

Ve smyslu projektového managementu má projekt několik základních charakteristik, kterými jsou jedinečnost, vymezenost v čase, penězích a zdrojích, složitost a komplexnost, rizikovitost a realizovatelnost týmem lidí z různých částí organizace (Doležal a Krátký, 2016, str. 16). Pamela McGhee a Peter McAliney ve své knize projekt popisují jako jedinečnou, jednorázovou práci s definovaným začátkem a koncem, jejíž cíle jsou předem definovány těmi, kteří na tom mají zájem (McGhee a McAliney, 2007, str. 3). V kontextu Smart City, se definice projektu dělí na dva směry. Např. článek *What is a Smart City project?* od Umberta Rosatiho a Sergia Contiho pojednává o projektu jako o celkovém procesu budování Smart City (Rosati a Conti,

2016). Článek *A structured method for Smart City project selection* autorů Yenchum Jim Wu a Jeng-Chung Chen naopak pojednává o projektech spíše ve smyslu charakteristik projektového managementu a koncept Smart City dělí na projekty v různých oblastech (Wu a Chen, 2021).

1.2.3 Měření a hodnocení Smart City

Jako vstup pro rozhodování o efektivnosti a plnění cílů strategií a projektů Smart Cities slouží různé kategorie a ukazatele, které hodnotí výsledky přijatých opatření v oblasti zavádění moderních technologií v jednotlivých oblastech Smart City. Výsledky opatření mohou mít různou úroveň podrobnosti a doporučujících kroků. Ukazatele odrážející výsledky opatření pro realizaci Smart City mají různý způsob vyjádření. Nejčastějšími jsou (Slavík, 2017, str. 62):

- jednoduše „ano/ne“ (například využívání/nevyužívání opatření);
- statisticky zjištěné absolutní hodnoty;
- provozní/fyzikální ukazatele vztažené k období nebo počtu obyvatel;
- jednotkové finanční ukazatelé vztažené k období;
- poměrově finanční a objemové ukazatele;
- poměrové ukazatele;
- bodovací systém na škále.

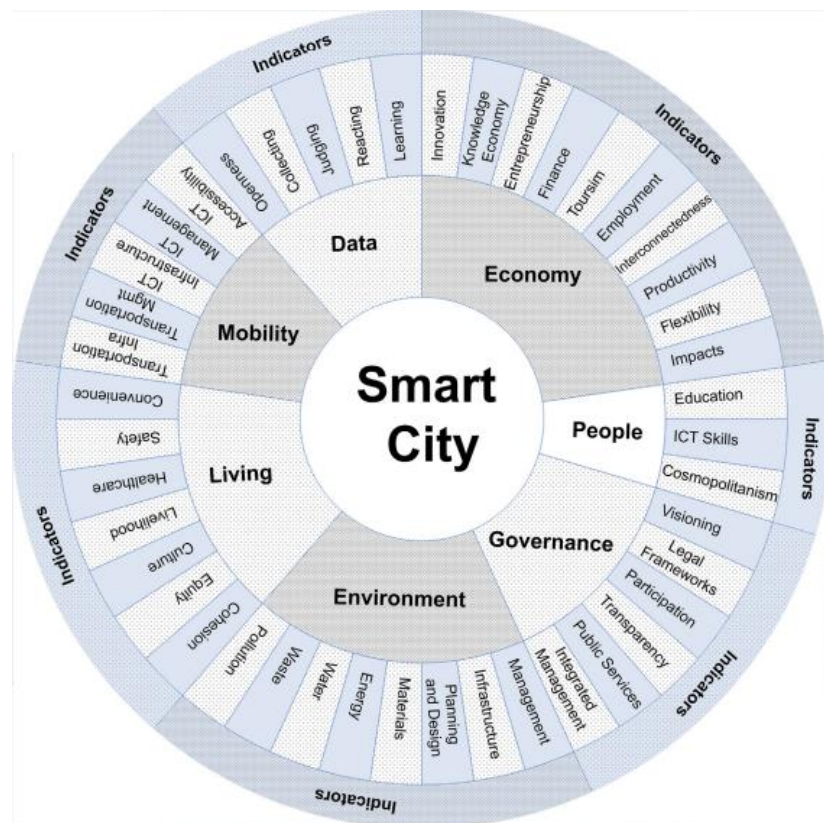
Na evropské úrovni se systémový přístup do hodnotících ukazatelů Smart City snaží vnést například projekt CITYkeys, který vychází z průzkumů u vybraných evropských měst a regionů a dále je doplňuje vlastními návrhy ukazatelů. Projekt navrhuje pro Smart City sadu 73 ukazatelů rozdělených do 4 skupin a dalších podskupin. Mimo zmíněné se samozřejmě dají projekty v rámci chráněných měst hodnotit i dalšími způsoby (Slavík, 2017, str. 65).

Cost-benefit analýza (analýza nákladů a přínosů) představuje nejčastěji používanou metodu při hodnocení společenské efektivity, jejímž základním smyslem je porovnat veškeré dopady a efekty připravovaného nebo realizovaného projektu. Nejedná se pouze o běžnou finanční kalkulaci, ale porovnávají se veškeré socio-ekonomické, společenské a externalitní vstupy a výstupy projektu. Projekt je žádoucí realizovat v případě, kdy celospolečenské náklady na zavedení a využívání těchto řešení nebudou převyšovat jejich objektivní a subjektivní přínosy pro všechny dotčené subjekty (Nevima et al., 2023, str. 42).

První hodnotící rámec, který představuje ucelenou strukturu kategorií a ukazatelů, včetně postupu pro jejich využití v praxi, byl publikován v roce 2007. Giffinger et al. (2007) zde uvádí 6 charakteristik (kategorií) a ke každé několik ukazatelů. Charakteristikami jsou tedy podle Giffingera: 1) Smart ekonomika (Konkurenceschopnost), 2) Smart lidé (Sociální a lidský

kapitál), 3) Smart samospráva (participace), 4) Smart mobilita (Doprava a ICT), 5) Smart prostředí (Přírodní zdroje) a 6) Smart život (Kvalita života). Smart ekonomika zahrnuje ukazatele, které se týkají hospodářské konkurenceschopnosti, jako jsou inovace, podnikání, ochranné známky, produktivita a flexibilita trhu práce, jakož i integrace na (mezi)národním trhu. Smart lidé nejsou charakterizováni pouze úrovní kvalifikace nebo vzdělání občanů, ale také kvalitou sociálních interakcí v oblasti integrace a veřejného života a otevřeností vůči "vnějšmu" světu. Smart samospráva zahrnuje aspekty politické participace, služeb pro občany i fungování správy. Místní a mezinárodní dostupnost jsou důležitými aspekty Smart mobility, stejně jako dostupnost ICT a moderních a udržitelných dopravních systémů. Smart prostředí je popsáno atraktivními přírodními podmínkami (klima, zeleň atd.), znečištěním, hospodařením se zdroji a také úsilím o ochranu životního prostředí. Smart život zahrnuje různé aspekty kvality života, jako je kultura, zdraví, bezpečnost, bydlení, cestovní ruch atd. (Giffinger et al., 2007, str. 11-12).

Indikátory lze vybírat pomocí jedné nebo kombinace těchto možností následujících metod: přehled literatury, odborný průzkum nebo konzultace se zúčastněnými stranami. Na základě jejich podobnosti se vybrané ukazatele klasifikují do kategorií. Příklad běžně používaných kategorií prezentuje následující Obrázek 1 (Sharifi, 2019, str. 1273).



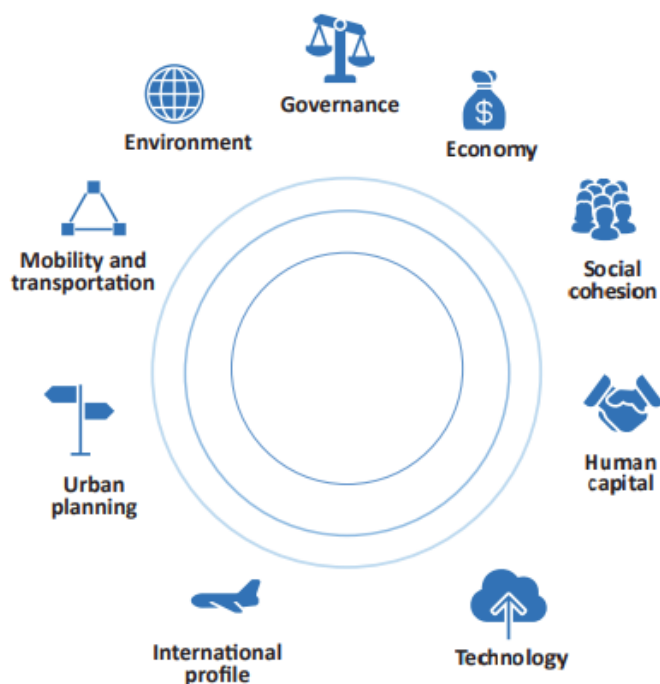
Obrázek 1: Kategorie a indikátory pro hodnocení Smart City. Zdroj: Sharifi (2019, str. 1273).

1.2.4 Smart City indexy

Smart Cities se dají porovnávat z různých hledisek a existuje několik rámců a indexů pro tuto problematiku. Ty jsou zpravidla publikovány jako reporty (zprávy), které porovnávají různá města podle různé sady kategorií a indikátorů, tzn. rámců, jejichž součástí je také metodický postup, jak tento rámec aplikovat v praxi, a získat tak pořadí měst. Tyto zprávy jsou vydávány zpravidla každý rok nebo jednou za dva roky. Ve snaze porovnat existující rámce a indexy, případně navrhnout nějaké změny, které by umožnily hodnotit Smart Cities komplexněji, např. s ohledem na nové technologie, vznikají různé výzkumné studie. Toto zkoumala např. studie *Measuring progress of Smart Cities: Indexing the Smart City indices*, kde na základě srovnání indexů vytyčily 9 základních kategorií, které by měl komplexní index zahrnovat. Součástí toho je 6 všeobecně uznaných dimenzí – ekonomie, prostředí, samospráva, život, mobilita, lidé a 3 prvky z témat stávajících indexů Smart Cities, a to vnímání, soukromí a kybernetická bezpečnost (Lai a Cole, 2023, str. 52).

Jednou z prvních zpráv, která obsahovala Global Cities Index (GCI), byla v roce 2008 Global Cities Report od společnosti Kearney. V průběhu let se měnil hodnotící rámec i počet měst, když GCI 2023 hodnotí, do jaké míry jsou města schopna přilákat, udržet a vytvářet globální toky kapitálu, lidí a myšlenek. Index měří 29 indikátorů napříč 5 dimenzemi, kterými jsou: business aktivity, lidský kapitál, předání informací, kulturní zkušenost a politická angažovanost (Kearney, 2023). Od roku 2014 publikuje IESE Business School index nazývaný Cities in Motion Index (CIMI), který je počítán na základě 9 dimenzí, jak ukazuje model viz Obrázek 2. Jeho cílem je hodnotit vybraná města z hlediska devíti klíčových dimenzí: ekonomiky, lidského kapitálu, technologie, životního prostředí, mezinárodního profilu, sociální soudržnosti, mobility a dopravy, správy veřejných věcí a územní plánování (Berrone a Ricard, 2022, str. 5). Poslední verze indexu je z roku 2022 a hodnotí 183 měst.

Global Power City Index (GPCI) je publikován Mori Memorial Foundation od roku 2015. GPCI hodnotí v poslední zprávě z roku 2023 celkem 48 města na základě jejich schopnosti přilákat lidi, kapitál a podniky. Zpráva Smart City Governments (SCG) hodnotí rozvoj 50 Smart Cities z pohledu městské samosprávy. Identifikuje soubor 10 klíčových nástrojů, které starostové považují za účinné při rozvoji Smart Cities. Žebříček pokrývá 50 měst v 6 regionech: Asie-Pacifik, Afrika, Evropa, Střední východ, Severní Amerika a Jižní Amerika, když z každého regionu pochází 5 měst. Poslední publikace žebříčku je z roku 2020/2021. K dispozici je za 50 dolarů i prémiová, obsáhlejší verze (Eden Strategy Institute, 2021).



Obrázek 2: Dimenze CIMI. Zdroj: Berrone a Ricart (2022, str. 11).

IMD Smart City Index (SCI) byl poprvé publikován v roce 2019, kdy hodnotil 102 měst oproti 141 v roce 2023. Jak IMD uvádí ve své aktuální zprávě z roku 2023, Smart Cities se stále mění, a proto je důležité upravovat i způsoby výpočtu. Ten byl dosud veden na základě národního indexu lidského rozvoje (Human Development Index, HDI), nyní se však do úvahy bere HDI městské. Zároveň byl tento index doplněn o další města (nárůst o 20 % oproti předchozím letům), což vede k větší přesnosti. Index zkoumá dva základní pilíře (pilíř struktury týkající se infrastruktury měst a pilíř technologie popisující technologické opatření a služby dostupné obyvatelům) a ty jsou dále rozděleny do pěti klíčových oblastí, a to zdravotnictví a bezpečnost, pohyblivost, aktivity, příležitosti a vládu (IMD Business School, 2023, str. 4, 9, 34).

Zpráva společnosti Corporate Knights s názvem Sustainable Cities Index Report je zaměřena na výsledky, přičemž 11 z 12 ukazatelů zahrnuje fyzická měření znečištění ovzduší částicemi, přístup k pitné vodě a její spotřebu, produkci odpadů, závislost na automobilové dopravě a hustotu silniční sítě, podíl tranzitního a aktivního přepravního režimu, volný prostor, místní emise skleníkových plynů a emise skleníkových plynů z dopravy a odolnost vůči dopadům změny klimatu. Dvanáctý ukazatel je zaměřen na politiku a zahrnuje závazky města v oblasti obnovitelných zdrojů energie, snižování emisí skleníkových plynů a čisté dopravy (Torrie a Morson, 2023, str. 3).

1.2.5 Smart City technologie

Jak bylo zmíněno na začátku kapitoly, Smart City se týkají různých oblastí. Celkový přehled oblastí a příkladů technologií Smart Cities těchto oblastí je k vidění v následující Tabulka 1.

Tabulka 1: Technologie Smart City. Zdroj: Kitchin (2016, str. 12).

Oblast	Příklad technologií
Vláda	Systémy elektronické veřejné správy; online transakce; systémy pro provoz města; systémy řízení výkonnosti; městské informační tabule
Bezpečnost a záchranné složky	Centralizované řídicí místnosti; digitální dohled; prediktivní policejní dohled; koordinované reakce na mimořádné události
Doprava	Chytré dopravní systémy; integrované jízdenky; chytré cestovní karty; sdílení jízdních kol; informace pro cestující v reálném čase; chytré parkování; řízení logistiky; dopravní aplikace
Energetika	Chytré sítě; chytré měřiče; aplikace pro spotřebu energie; chytré osvětlení
Odpad	Lisovací popelnice a dynamické směřování/sběr
Životní prostředí	Senzorové sítě (např. znečištění, hluk, počasí; pohyb po zemi; řízení povodní)
Budovy	Systémy pro správu budov; senzorové sítě
Domácnosti	Chytré měřiče, chytré spotřebiče ovládané aplikacemi

1.3 Ochrana dat

1.3.1 Ochrana osobních údajů

V ČR je ochrana osobních údajů legislativně ošetřena zákonem č. 110/2019 Sb. o zpracování osobních údajů (Česko, 2019). Vychází z nařízení Evropského parlamentu a Rady 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů nazývaném General Data Protection Regulation (GDPR). Zákon č. 110/2019 Sb. stanovuje pravidla pro zpracování osobních údajů, včetně práv subjektu údajů (fyzické osoby).

Osobní údaje jsou jakékoli informace, které se týkají identifikované nebo identifikovatelné žijící osoby. K osobním údajům patří i různé jednotlivé informace, které společně jako celek mohou vést k identifikaci určité osoby. Osobní údaje, které sice byly zbaveny informací umožňujících identifikaci, zašifrovány nebo pseudonymizovány, ale lze je použít ke zpětné identifikaci osoby, zůstávají osobními údaji a GDPR se na ně i nadále vztahuje. Osobní údaje, které byly anonymizovány takovým způsobem, že příslušná osoba již není identifikovatelná,

již nejsou považovány za osobní údaje. Údaje se pokládají za skutečně anonymizované, pokud je anonymizace nezvratná (Česko, 2019).

1.3.2 Kyberprostor a kyberbezpečnost

Pojem kyberprostor pochází z řeckého slova kyber (navigovat). Termín byl poprvé použit v roce 1984 Williamem Gibsonem v novele Neuromancer. Tento zdroj popisuje kyberprostor jako digitální prostor propojených počítačů přístupný pomocí počítačových konzolí: vizuální, barevný, elektronický. Od té doby byl pojem několikrát adaptován různými způsoby, přesto pokaždé směřuje k počítačově zprostředkované komunikaci. (Dodge a Kitchin, 2003, str. 1). V dnešní době se dá o kyberprostoru mluvit jako o vzájemně závislé síti infrastruktur ICT, která zahrnuje internet, telekomunikační sítě, počítačové systémy a vestavěné procesory a řídicí jednotky v kritických průmyslových oblastech (Ross et al., 2021, str. 78).

Digitalizace a udržování spojení hrají zásadní roli. Zabezpečení už proto není jen požadavkem správců zabezpečení v IT organizaci, ale je to nyní požadavek všech subjektů, které jsou tak či onak spojeny s jakýmkoli typem dat (Ozkaya, 2019, str. 5). Ochrana systémů ICT a jejich obsahu se nazývá kybernetická bezpečnost. Kyberbezpečnost je široký pojem, který se obvykle vztahuje na jednu nebo více ze tří věcí (Fischer, 2014, str. 4):

- Soubor činností a opatření, jejichž cílem je chránit – před útokem, narušením a jinými hrozbami – počítače, počítačové sítě, hardware (HW) a další zařízení, SW a informace, které obsahují a komunikují, včetně SW a dat, ale i dalších prvků kyberprostoru.
- Stav nebo kvalita ochrany před těmito hrozbami.
- Široká oblast úsilí zaměřená na provádění a zlepšování těchto činností a kvality.

V Česku pro kybernetickou bezpečnost existuje zákon Zákon č. 181/2014 Sb., o kybernetické bezpečnosti a o změně souvisejících zákonů (zákon o kybernetické bezpečnosti). Tento zákon upravuje práva a povinnosti osob a působnost a pravomoci veřejné moci v oblasti kybernetické bezpečnosti a zajišťování bezpečnosti sítí elektronických komunikací a informačních systémů (Česko, 2014).

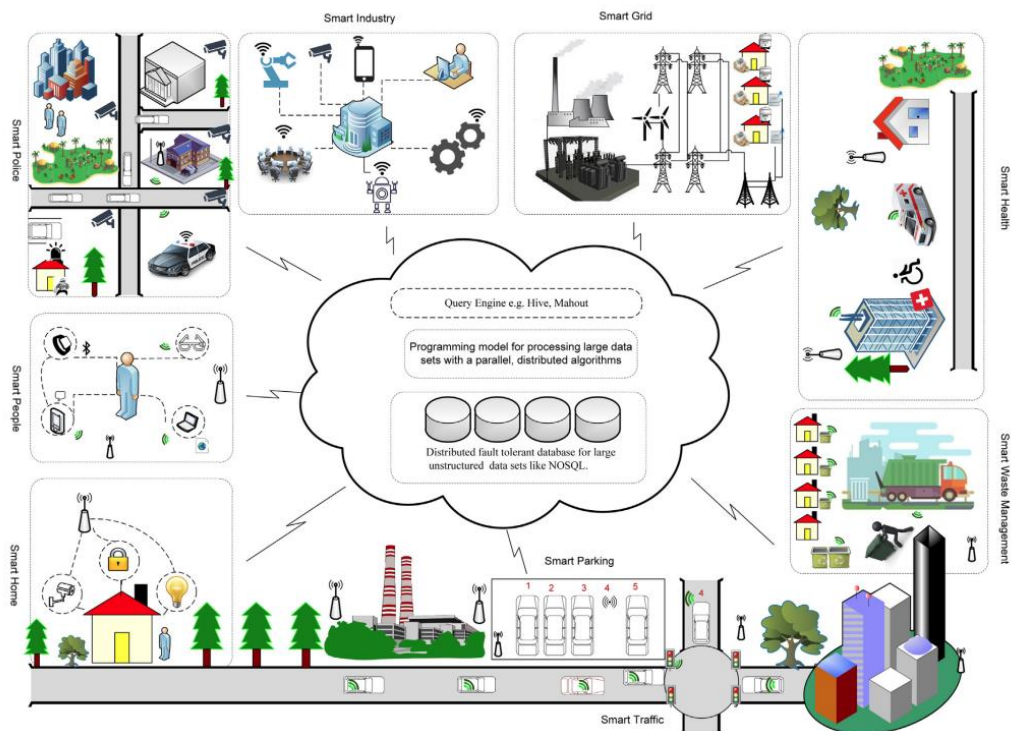
2 IDENTIFIKACE A POROVNÁNÍ PŘÍSTUPŮ A KRITÉRIÍ V DANÉ OBLASTI

2.1 Big data a Smart City

Dvě z nejpoužívanějších slov současnosti jsou data mining a big data. Data mining zkoumá rozsáhlé databáze a prostřednictvím datové analytiky hledá skryté významy a odhaluje klíčové trendy. Pomocí těchto dat, někdy nazývaná jako smart data, lze zjistit, jak různé typy a míry změn ovlivňují městskou komunitu. Lze například sestavit kauzální model komunity založený na klíčových faktorech, jako jsou požadavky vzdělávacího a zdravotnického systému, energie, životního prostředí, dopravy, komunikace, územního plánování, bydlení a výstavby atd., a pak sledovat, jak tyto různé faktory městských změn reagují, když se mění míra růstu populace, stárnutí občanů a pracovní síly atd. v závislosti na tyto různých složky. Jak se město zároveň rozrůstá a diverzifikuje svou ekonomiku, musí mít infrastrukturu, která se může měnit tak, aby vyhovovala změněným požadavkům. Tento typ analýzy smart dat je proto klíčem k rozvíjejícím se potřebám Smart City (Pelton a Singh, 2018, str. 23).

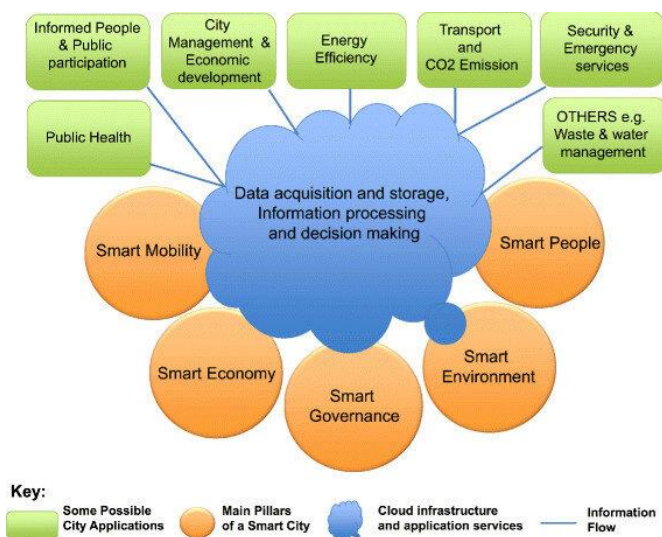
Smart Cities ze své podstaty generují značné množství dat ze svých každodenních operací. Trendy jako internet věcí (Internet of Things, IoT) a open data vedou města ke shromažďování a zpřístupňování dalších významných objemů dat, kdy některé z nich jsou statické ale stále větší množství dat je v reálném čase. Tato data mohou, jsou-li dobře spravována a analyzována, nabídnout poznatky a přinést ekonomickou hodnotu, kterou mohou města a zúčastněné strany využít ke zlepšení efektivity. Tyto poznatky pak mohou vést k inovacím nových služeb zlepšujících životy občanů (Nevima et al., 2023, str. 26). Kitchin dochází k závěru, že vládám a úředníkům big data nabízejí účinnější a efektivnější možnosti pro správu a regulaci měst (Kitchin, 2014, str. 12).

Přehledový obrázek k tomuto tématu se objevil v článku *The role of big data in Smart City*, viz Obrázek 3. Ten zachycuje propojení smart technologií s big daty a cloud computingem, v níž si různé smart aplikace vyměňují informace pomocí vestavěných sensorových zařízení a dalších zařízení integrovaných s cloudovou výpočetní infrastrukturou, čímž zároveň vytvářejí velká množství nestrukturovaných dat. Tato data jsou shromažďována a ukládána v cloudu nebo datovém centru (Hashem, 2016). Obrázek rozděluje části města do smart kategorií jako je smart domácnost, smart lidi, smart policie, smart průmysl, smart síť, smart zdravotnictví, smart řízení odpadu, smart doprava a smart parkování.



Obrázek 3: Oblast Smart City a big dat technologií. Zdroj: Hashem (2016, s. 749).

Podobný obrázek se vyskytuje i v publikaci *Cloud Based Big Data Analytics for Smart Future Cities*, viz Obrázek 4. Střed obrázku znázorňuje infrastrukturu, tedy cloud, kolem nějž jsou v oranžových kruzích pilíře Smart City: smart mobilita, smart ekonomie, smart samospráva, smart prostředí a smart lidi. V zelených obdélnících jsou uvedeny příklady možných aplikací pro Smart City, např. smart zdravotnictví, informování lidí a participace veřejnosti, řízení města a ekonomický rozvoj, efektivní energie, doprava a emise, bezpečnostní a pohotovostní služby a v neposlední řadě také řízení vody a odpadu.

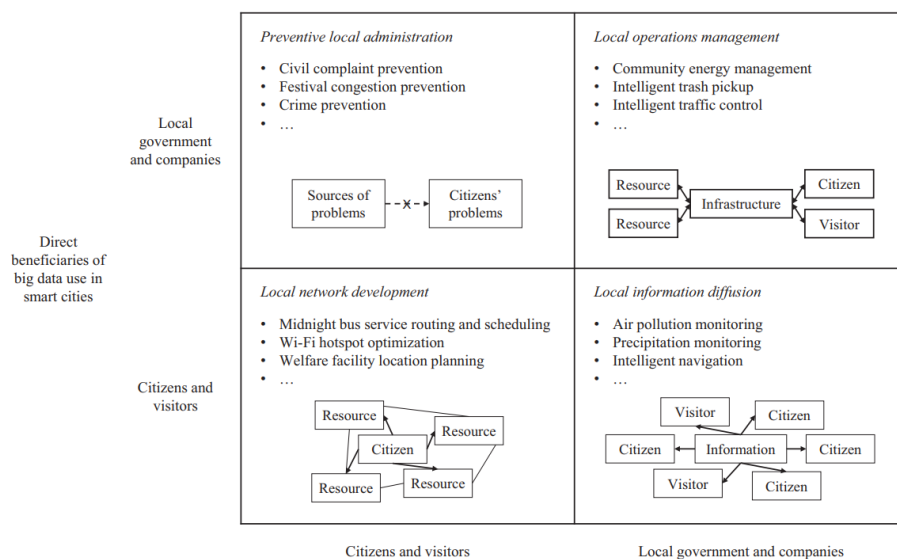


Obrázek 4: Správa/analýza dat pro aplikace v cloudu. Zdroj: Khan et al. (2013, s. 382).

Smart City propojuje město s digitálními technologiemi prostřednictvím všudypřítomného IoT. Ten shromažďuje obrovské množství dat pomocí velkého množství senzorů. Big data musí projít ukládáním, zpracováním, následnými dotazy (query) a analýzou, aby mohla být využita pro všechny typy chytrých aplikací a služeb. Navíc roste potřeba provádět ukládání, zpracování, dotazování a analýzu big dat v reálném čase, což přináší řadu problémů a výzev (Li et al., 2015):

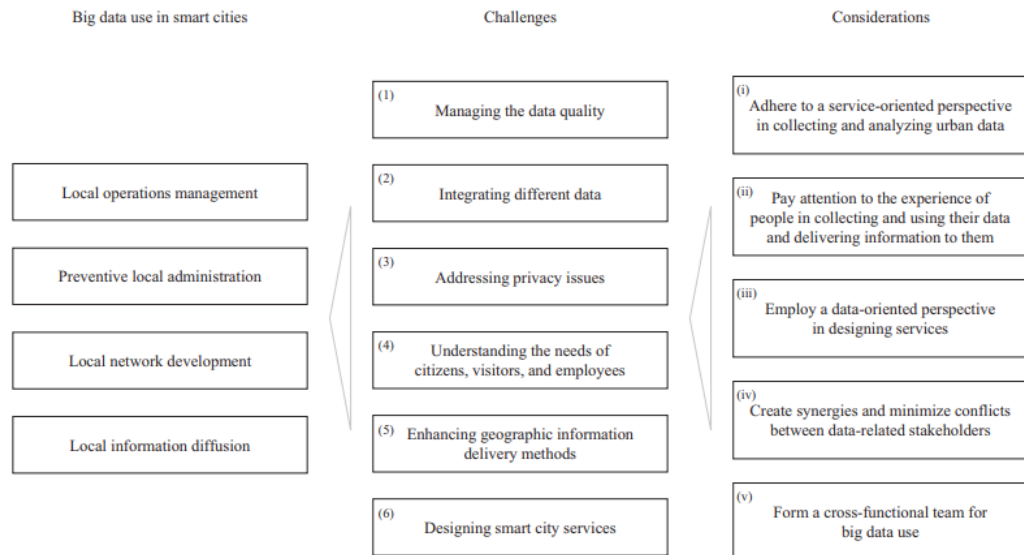
- Vysoké náklady na ukládání big dat. Například bezpečnostní systém města Tchien-ťin si vyžádá rozpočet 50 miliard jüanů na vybudování skladovací kapacity 4,6 EB.
- Nízký stupeň automatizace rychlých dotazů a vyhledávání big dat. Tradiční informační systémy provádějí jen jednoduchý sběr a ukládání dat a postrádají efektivní automatické vyhledávání a analýzu klíčových sémantických informací, jako je například chování.
- Získávání znalostí z big dat je velmi obtížné. Big data obsahují nejen data a informace ale také bohatství pravidel a znalostí. Tato pravidla a znalosti nejsou dány přímo. Místo toho je třeba je třeba provést hloubkovou analýzu a dolování dat.

Další článek *Smart Cities with big data: Reference models, challenges, and considerations* představuje klasifikaci případů užití big dat ve Smart Cities do čtyř kategorií. Dvě osy představují přímé příjemce využívání big dat (osa y) a hlavní zdroje big dat (osa x) a každá z nich je rozdělena na dvě úrovně, a to na místní poskytovatele služeb (tj. místní samosprávu a podniky) a jednotlivé zákazníky (tj. občany a návštěvníky). Každý kvadrant, jak ukazuje Obrázek 5, obsahuje název, reprezentativní případy a koncepční znázornění kategorie. Každá ze čtyř kategorií se zaměřuje na určitou oblast interakce dat a informací a navrhuje obecný referenční model pro Smart Cities využívající data, který je nezávislý na oblasti odvětví (Lim et al., 2018).



Obrázek 5: Klasifikace případů užití big dat ve Smart Cities. Zdroj: Lim et al. (2018, str. 92).

Stejný článek obsahuje i Obrázek 6, který navazuje na ten předchozí. Jedná se především o to, jaké výzvy a jaká doporučení by mělo město vzít v úvahu při vývoji či implementaci aplikací využívající big data.



Obrázek 6: Doporučení pro big data aplikace ve Smart City. Zdroj: Lim et al. (2018, str. 92).

2.2 Požadavky na strategie a projekty Smart City

Budování Smart City zaměřeného na data vyžaduje podporu technologií a postupů pro dolování dat a integraci různých datových formátů. Technologie vizualizace dat pak mohou prezentovat složitá data o městě ve stručné a přehledné podobě, což vyplňuje mezeru v porozumění mezi technologií a jejími uživateli (Wu et al., 2018). Ze všech systémů Smart City bude pochopitelně generováno obrovské množství dat. Pro zpracování různých typů dat s různou rychlostí je zapotřebí účinný systém správy big dat. Tento systém musí být spolehlivý a škálovatelný bez výpadků. Průběžné generování, sběr, zpracování a ukládání velkého množství heterogenních dat ze senzorů Smart City pak přináší další výzvy. Nicméně big data shromážděná v rámci celého města jsou nesmírně užitečná a jsou nezbytná pro dosažení cílů Smart City. Například snímače Global Positioning System (GPS) instalované na vozidlech mohou poskytnout cenné informace o dopravním toku, ale zároveň generují velmi rychle velké množství dat (Bawany a Shamsi, 2015).

Při zvažování aplikací pro Smart Cities založených na big datech je nutné řešit několik požadavků, které vyplývají ze zvláštních potřeb Smart Cities a big dat. Tyto požadavky jsou identifikovány na základě aplikací pro zpracování big dat a výzev spojených s implementací

těchto aplikací do Smart Cities. V článku *Applications of big data to Smart Cities* se uvádějí následující obecné nefunkční požadavky (Al Nuaimi et al., 2015, str. 9-12):

- **Big data management:** Aplikace Smart Cities generují data v různých formátech a z různých sektorů. Vzhledem k tomuto je zapotřebí pokročilých funkcí pro správu dat, které povedou k rozpoznávání různých formátů a zdrojů dat, jejich strukturování, správu, klasifikaci a kontrolu těchto typů a struktur.
- **Platformy pro zpracování big dat:** Big data aplikace pro Smart Cities potřebují provádět analýzu dat, která obvykle vyžaduje velkou kapacitu pro zpracování. Proto je potřeba škálovatelných a spolehlivých softwarových a hardwarových platform. Tyto platformy by měly nabízet vysoký výpočetní výkon, měly by být stabilní pro různé aplikace náročné na data, podporovat proudové zpracování a poskytovat vysokou odolnost proti chybám.
- **Chytrá síťová infrastruktura:** Většina big data aplikací pro Smart Cities vyžaduje inteligentní sítě propojující jejich komponenty, včetně zařízení obyvatel, jako jsou auta, chytré domácnosti a telefony. Tato síť by měla být schopna efektivně přenášet shromážděná data z jejich zdrojů do míst, kde se big data shromažďují, ukládají a zpracovávají, a přenášet odpovědi zpět různým subjektům.
- **Pokročilé algoritmy:** Standardní algoritmy z běžných aplikací nemusí být efektivní pro big data aplikace z hlediska jejich jedinečných požadavků.
- **Otevřená standardizovaná technologie:** Vzhledem k tomu, že aplikace big dat pro Smart Cities zahrnují rozsáhlé heterogenní systémy a data, doporučuje se proto řídit se otevřeným standardem pro navrhování a implementaci takových řešení.
- **Bezpečnost a soukromí:** Jelikož většina shromažďovaných a zpracovávaných údajů v aplikacích Smart Cities bude obsahovat určitou formu citlivých nebo soukromých informací, je důležité zajistit, aby všechny technologické a aplikační komponenty zahrnovaly a udržovaly přijatelnou úroveň mechanismů zabezpečení a ochrany soukromí.
- **Informovanost občanů:** Občané musí být informováni o tom, že IT řešení pro Smart Cities jsou provedeny správně a bezpečně.
- **Role vlády:** Odpovědné subjekty Smart Cities musí stanovit hlavní zásady otevřenosti, transparentnosti, účasti a spolupráce, aby měly výměnu a tok big dat pod kontrolou.

Požadavků lze identifikovat velké množství. Tabulka 2 ukazuje souhrn těch vybraných a ve své podstatě klíčových požadavků na systémy big dat v souvislosti se Smart Cities.

Tabulka 2: Vybrané požadavky na systémy big dat. Zdroj: *Vlastní zpracování*.

Požadavek	Popis	Zdroj
Škálovatelná a odolná datová infrastruktura	Smart Cities by měla budovat takové sítě, které dokáží pojmout rostoucí objemy dat a zároveň vydrží různé výkyvy v objemu toku dat.	Al Nuaimi et al. (2015), Bawany a Shamsi (2015), Hashem (2016) Khan et al. (2013)
Správa dat	Smart City shromažďují velké množství dat z různých druhů zdrojů a určitou výzvou je tyto zdroje ve výsledku získat, zpracovat a spojit.	Al Nuaimi et al. (2015), Bawany a Shamsi (2015), Khan et al. (2013), Li et al. (2015)
Ukládání dat	Otázka ukládání dat je jednou z výzev Smart City a existuje k ní několik přístupů. Tento požadavek by se dal zahrnout i do správy dat, ale významově přišlo vhodné tento požadavek oddělit.	Bawany a Shamsi (2015), Hashem (2016), Li et al. (2015)
Zpracování dat	Nejdůležitější oblastí pro využití big dat je big data analytics, která se někdy provádí i v reálném čase (například v dopravě).	Hashem (2016), Li et al. (2015), Pelton a Singh (2018)
Ochrana dat	Shromažďovaná a zpracovávaná data mohou částečně či úplně obsahovat citlivé či soukromé údaje. Proto je důležité, aby všechny části systémů ve Smart City byly bezpečné.	Al Nuaimi et al. (2015), Lim et al. (2018)

2.3 Přístupy a kritéria pro hodnocení strategií a projektů Smart Cities

Na základě studia a analýzy literatury a souvisejících zdrojů v předchozích kapitolách lze tvrdit, že přístupy a kritéria používané pro hodnocení úspěšnosti strategií a projektů Smart Cities se liší. V globálním pohledu se nejčastěji používají různé rámce a indexy, které jsou zpravidla pravidelně aktualizovány s ohledem na moderní trendy v ICT nebo požadavky na udržitelnost řešení, tzn. především snižování uhlíkové stopy a propojování ICT se zelenými technologiemi. Přestože jsou zprávy a související indexy jako CIMI, GCI, GPCI atd. zpravidla dostupné široké veřejnosti a poskytují širokou škálu doporučení, jak budovat Smart City, a které přístupy nebo komponenty se v praxi nejvíce osvědčily, tak tyto indexy porovnávají jen vybraná světová města, často hlavní města, která mají k dispozici jak finanční, tak výpočetní i lidské zdroje.

Nicméně i z těchto zpráv je možné získat obecná doporučení pro tvorbu Smart City strategií a projektů, tzn. především 1) definovat si cíle a dimenze, které jsou pro Smart City klíčové, 2) analyzovat současnou situaci z s ohledem na tyto dimenze, např. jako silné a slabé stránky,

a 3) definovat konkrétní projekty a postupy jejich řešení s ohledem na cíle, a to včetně nějakých měřitelných indikátorů, pomocí kterých lze postup monitorovat. S ohledem na požadavky big dat a obecně ochranu dat lze i v těchto globálních zprávách nalézt určitá doporučení, která nejčastěji vycházejí z rozdělení na *funkční požadavky* pro konkrétní projekty a odpovídající dimenze, např. ověření přístupu k big datům pro různé skupiny uživatelů nebo získávání dat pouze od občanů daného města nebo zaměření jen na turisty, a nefunkční požadavky, např. škálovatelnost a odolnost infrastruktury musí pojímat jak data od občanů, tak od turistů nebo dalších osob pohybujících se na území města.

Jelikož má ale každé Smart City svoje specifika, často vycházející z ekonomické, sociální nebo kulturní situace v dané zemi, tak by měly přístupy a kritéria pro hodnocení projektů a strategií Smart Cities vycházet primárně z toho, jaké finanční, lidské a výpočetní zdroje má dané město k dispozici. Obzvláště pokud jde o specifické oblasti jako jsou big data. Zároveň je zde nutné zmínit, že zprávy a související indexy jako CIMI, GCI, GPCI atd. nemohou pokrýt celou hloubku analýzy toho, jak dané Smart City funguje a jaké jsou jeho strategie a projekty. Proto budou tyto strategie podrobně analyzovány v dalších kapitolách.

2.4 Identifikace oblastí Smart City

Různé zdroje i indexy se snaží vymezit koncept Smart City rozdělením do určitých oblastí nebo dimenzí. Tabulka 3 tak slouží jako obecný přehled toho, jak k tomuto rozdělení různé zdroje přistupují. Tyto informace mohou být klíčové např. pro to, jaké oblasti jsou ve Smart Cities aktuálně preferovány, a lze se takto dostat k ověřeným v praxi fungujícím řešením, resp. best practices. Zároveň porovnáním těchto informací v čase, což se týká především indexů, které jsou publikovány v nějakém časovém intervalu a zjištění v nich obsažená pokrývají určité časové rozpětí, se lze zpětně poučit z chyb nebo špatně zvolených strategií a projektů.

Závěrem je ještě nutné dodat, že v tabulce jsou znázorněny pouze ty oblasti, které byly podpořeny více zdroji. Jiní autoři přidávají i další možné oblasti, popřípadě stejnou oblast popisují jiným názvem, v krajním případě pak některé oblasti slučují dohromady. V tabulce je zaznamenáno devět oblastí, které se v nalezených zdrojích nejčastěji vyskytovaly. Konkrétně se jedná o oblasti Smart lidé, Smart život, Smart průmysl, Smart síť, Smart zdravotnictví, Smart doprava, Smart samospráva, Smart prostředí a Smart bezpečnost. Tyto oblasti se nesnaží vytyčít pouze autoři v jejich pracích, ale i samotné indexy hodnotící města, především pak Smart Cities.

Tabulka 3: Souhrn oblastí pro koncept Smart City. Zdroj: *Vlastní zpracování.*

Oblast	Zdroj
Smart lidé	Giffinger et al. (2007), Hashem (2016), Khan et al. (2013), Lai a Cole (2023), Sharifi (2019)
Smart život	Giffinger et al. (2007), Lai a Cole (2023), Sharifi (2019)
Smart průmysl	Giffinger et al. (2007), Hashem (2016), Khan et al. (2013), Lai a Cole (2023), Sharifi (2019)
Smart síť	Hashem (2016), Ministerstvo pro místní rozvoj ČR (2018)
Smart zdravotnictví	Hashem (2016), IMD Business School (2023)
Smart doprava	Berrone a Ricard (2022), Giffinger et al. (2007), Hashem (2016), IMD Business School (2023), Khan et al. (2013), Sharifi (2019)
Smart samospráva	Berrone a Ricard (2022), Giffinger et al. (2007), IMD Business School (2023), Khan et al. (2013), Lai a Cole (2023), Sharifi (2019), Ministerstvo pro místní rozvoj ČR (2018)
Smart prostředí	Berrone a Ricard (2022), Giffinger et al. (2007), Khan et al. (2013), Lai a Cole (2023), Sharifi (2019)
Smart bezpečnost	Hashem (2016), IMD Business School (2023), Lai a Cole (2023)

Oblast Smart lidé se věnuje především propojení lidí s městem, či propojení lidí navzájem. Může se to týkat podoblastí jako vzdělání, veřejný život nebo sociální interakce. Rozdělení do oblastí podle Hashem (2016) naznačuje propojení chytrých zařízeních občana, jako je např. notebook či počítač, mobilní telefon, chytré hodinky a další chytrá zařízení, s městem. Smart domácnost by také mohla být zařazena do této oblasti, nicméně zde bude docházet k jistému propojení s následující oblastí, kterou je Smart život zahrnující podoblasti týkající se kvality života ve městě, tzn. bydlení, kultury, turismu a dalších.

Smart průmysl nezahrnuje pouze průmysl, ale i další oblasti týkající se ekonomiky a podnikání. Tato oblast může buď vymezovat to, jak jsou jednotlivé průmyslové organizace provázány se Smart City, nebo jaký je jejich technologický pokrok. Smart síť je zde vedena ve smyslu propojení elektřiny, tepla a dalších zdrojů, jako je např. rozvod vody.

Smart zdravotnictví se týká především fungování nemocnic a soukromých zdravotnických jednotek v rámci města. Nicméně daná oblast může být rozšířena i o Smart zdraví, která je jednou z oblastí SCI. Tento index pro tuto oblast sleduje např. ukazatele toho, zda jsou zdravotní služby uspokojivé či jak je složité nebo naopak jednoduché objednat lékařskou prohlídku online.

Smart doprava bývá často mezi autory označována jako mobilita. Do této oblasti se dá zahrnout i městská hromadná doprava. Smart parkování také může být součástí této oblasti. Další oblastí jsou i sdílené služby v rámci dopravy, ať už v rámci automobilů, či půjčování kol či koloběžek. Řízení vytíženosti dopravy či poskytování informací o vytíženosti dopravy také patří do této oblasti, včetně chytrých křižovatek, přechodů atd.

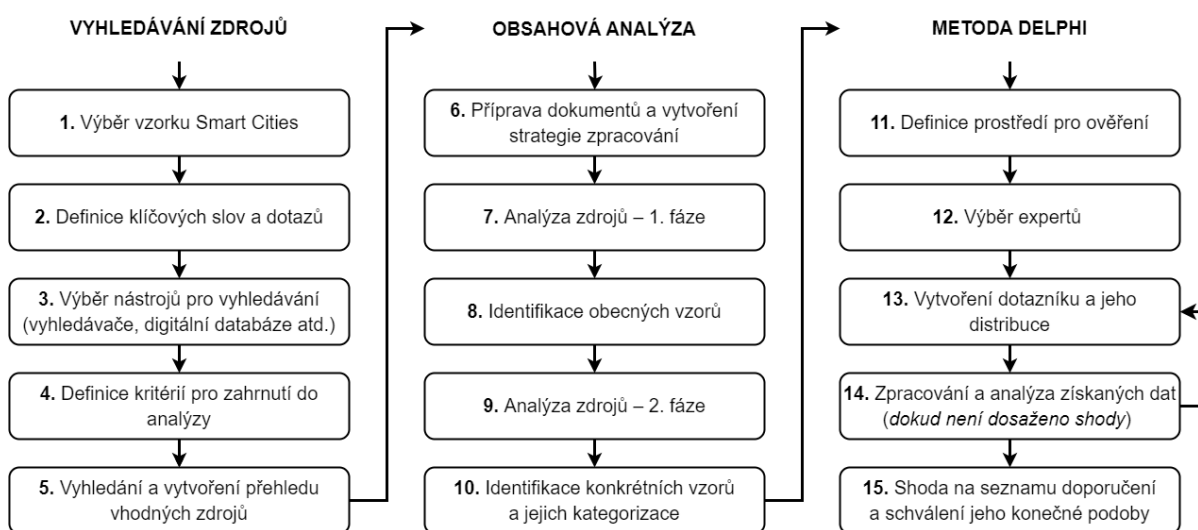
Smart samospráva se obecně věnuje podoblastem spojeným s řízením města. Pod tímto si lze představit škálu různých činností, jako je např. využívání technologií k pomoci s územním plánováním, nebo provázanost online prostředí k vyřizování občanských záležitostí, tzn. poskytování online veřejných služeb jako jsou petice, různá hlasování, případně oslovování konkrétních občanů na základě datové analytiky. Dále lze do této oblasti zařadit informování občanů o věcech spojených s chodem města.

Ve smyslu klimatického prostředí, nebo také životního prostředí, se zde objevuje oblast Smart prostředí. Ta se věnuje např. přírodním podmínkám (klimatu, zeleni ve městě atd.), různým druhům znečištění a jejich monitorování pomocí senzorů typu IoT, GPS, řízení přírodních zdrojů (senzory opět sledují, jestli je něčeho hodně nebo málo) a dalším podobným tématům.

Smart bezpečnost se chápe spíše ve smyslu kybernetické bezpečnosti a ochrany soukromí. Tak to chápe Lai a Cole (2023) ve studii *Measuring progress of Smart cities: Indexing the Smart city indices*. SCI se pak věnuje bezpečnosti jako celkové bezpečnosti ve městě, když jedním z ukazatelů je např. to, zda kamery pomáhají občanům cítit se bezpečně, resp. tento index zmiňuje veřejnou bezpečnost obecně. Hashem (2016) pak ve svém modelu uvádí oblast Smart policie, která znázorňuje využití kamer ve městě, což koresponduje s SCI. Otázkou však je, zda oblast policie nepatří spíše do Smart samosprávy.

3 METODIKA A POSTUP ŘEŠENÍ

Metodický postup směřující ke splnění cíle této diplomové práce, tzn. *analyzovat problematiku big dat, práci s těmito daty a jejich ochranu ve strategiích a projektech Smart Cities ve vybraných světových městech a na základě těchto zjištění navrhnout modely a doporučení pro tuto problematiku*, je zobrazen na Obrázek 7. Postup se dá rozdělit do 3 hlavních částí a celkem 15 dílčích kroků. První část se věnuje vyhledávání zdrojů pro analýzu, která je částí druhou. Poslední část, tedy ověření, je provedeno metodou Delphi. Těmto krokům se podrobněji věnují následující kapitoly.



Obrázek 7: Metodika a postup řešení cíle práce. Zdroj: *Vlastní zpracování, draw.io*.

3.1 Výběr vzorku Smart Cities

Vzorek reprezentativních měst pro analýzu byl vybrán na základě aktuální zprávy pro SCI 2023. Během zkoumání literatury a dalších zdrojů bylo v několika z nich odkazováno právě na tento index (odkazuje na něj například i ročenka *Smart Prague Index*). Tento index byl vybrán také z toho důvodu, že v roce 2023 se změnila metodika způsobu výpočtu indexu, která by měla vést ke zdokonalení žebříčku a lepšímu reflektování aktuálních požadavků na Smart Cities, např. co se týká udržitelnosti. Z tohoto důvodu je ale doporučeno neporovnávat momentální hodnocení s výsledky měst z minulých let.

Pro města byly v rámci této práce vytyčeny čtyři regiony, ze kterých by byla města vybírána. Ty jsou následující: Evropa, Severní Amerika + Kanada, Asie a zbytek světa. Aby byl výběr dostatečně reprezentativní, tak pro něj byla nastavena některá omezení. Prvním omezením je, že z každého regionu může být vybráno *maximálně deset měst*. V rámci Evropy, Asie a zbytku světa potom může být z jedné země vybráno *jenom jedno město*. Výběr vzorku měst by se tedy

měl skládat maximálně ze 40 měst z celkových 141 měst v SCI 2023. Výběr měst byl prováděn za rok 2023 a byl ukončen k 30.11.2023.

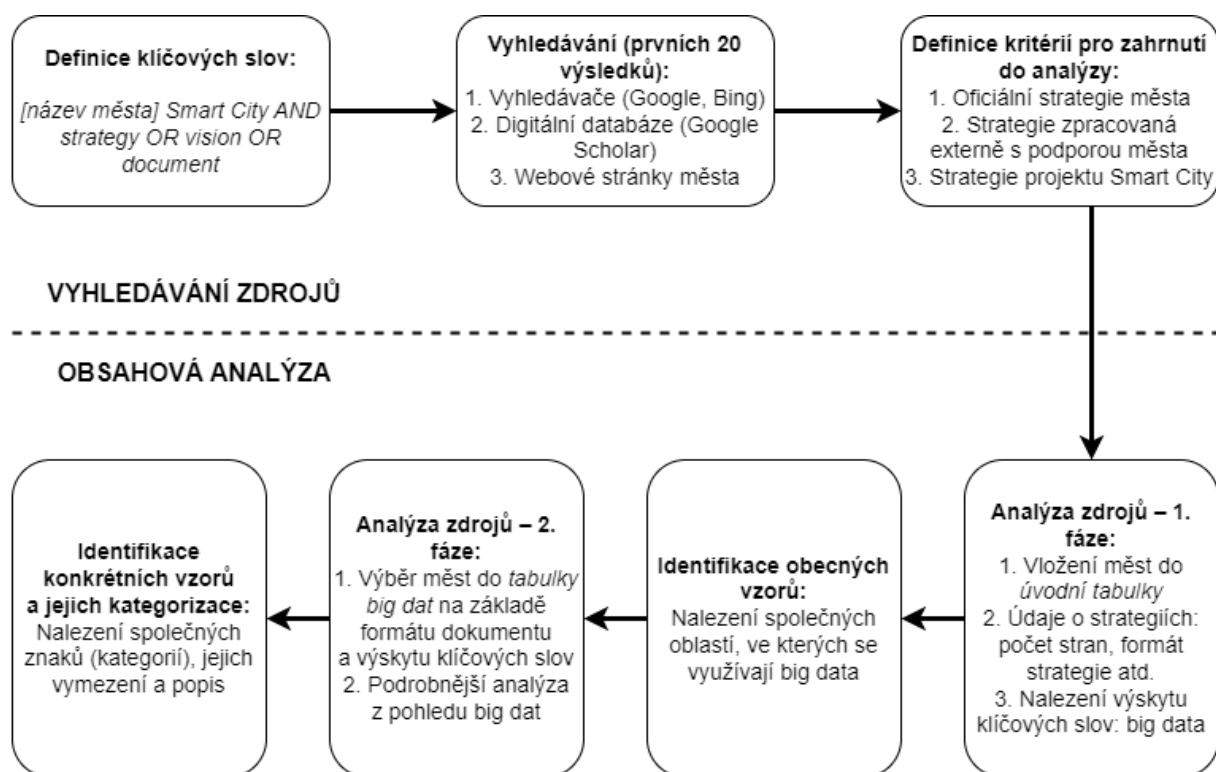
Způsob, jakým byla vybrána města pro úvodní tabulku, která obsahuje výběr všech 40 měst (viz kapitola 3.3), lze nalézt v Příloze I. Tento výběr probíhal podle pořadí měst SCI 2023, když se postupovalo město po městu od první pozice směrem k dalším, a to při použití omezení z předchozího odstavce. Pro druhou tabulku, tzn. podrobnější analýzu, byla vybrána pouze města, která měla vlastní Smart City strategii, která zároveň byla ve formátu psaného textu nebo prezentace a zmiňovala v nějakém kontextu pojem big data (viz kapitola 3.4). Podrobný postup je zobrazen v Příloze II.

3.2 Vyhledávání strategií a postup analýzy

Podrobný postup, jakým byly vyhledávány a analyzovány strategie pro všech 40 měst je uveden na následujícím obrázku, viz Obrázek 8. V prvním obdélníku jsou definována klíčová slova pro vyhledávače. Druhý obdélník řeší, kde jsou klíčová slova vyhledávána. Jako vyhledávače byly vybrány dva nejpoužívanější, a to Google a Bing. Dále byla využita Google Scholar databáze a vyhledávání v rámci oficiálních stránek vybraných měst.

Třetí obdélník řeší kritéria, při kterých se nalezený dokument použije v analýze. Jsou zde tři různé možnosti, buď když se jedná o oficiální strategii města, strategii zpracovanou externí firmou s podporou města anebo o strategii projektu Smart City, kde se pod projektem Smart City rozumí nějaký projekt se speciálním názvem, který se v dané lokalitě stará o vývoj Smart City (např. v Praze projekt Golemio).

Následující obdélníky ve druhé řadě se pak věnují analytické činnosti. V rámci obdélníku Analýza zdrojů – 1.část se řeší zápis údajů a poznatků zjištěných z dokumentů zahrnutých do analýzy do úvodní tabulky. Na základě první tabulky pak dojde k identifikaci oblastí, ve kterých strategie Smart City řeší big data a tyto oblasti budou vystupovat ve druhé tabulce. Postup končí identifikací a kategorizací konkrétních vzorů na základě druhé tabulky.



Obrázek 8: Postup vyhledávání strategií a jejich analýzy. Zdroj: *Vlastní zpracování, draw.io.*

Je ještě nutné dodat, že pro možnost porovnání byly hledány převážně strategie dostupné ve formátu PDF, kdy většina strategií je na internetu dostupná právě v této verzi. Některá města však mají i více či méně interaktivní webovou stránku či dokonce stránky, které bývají dosti specifické a je obtížné je mezi sebou porovnat. Zároveň tyto webové stránky bývají k dispozici současně s obsáhlejší strategií, např. ve formátu PDF. Z těchto důvodů s nimi diplomová práce nepracuje.

3.3 Zápis dat do úvodní tabulky

Prvním krokem obsahové analýzy dokumentů je vytvoření úvodní tabulky. Její obsah zachycuje Tabulka 4, kde je vidět přehled všech 40 měst definovaného vzorku, a to včetně sloupce ID (vytvořeného v rámci diplomové práce), pořadí v indexu (hodnota z SCI 2023), city HDI (městské HDI převzaté ze SCI 2023), názvu města, počtu obyvatel (hodnota převzatá z portálu World Population Review), země, ze které město pochází, region a country HDI (národní HDI převzaté ze SCI 2023).

Tabulka 4: Úvodní tabulka. Zdroj: *Vlastní zpracování, MS Excel.*

ID	Pořadí v indexu	City HDI	Český název	Město	Počet obyvatel (2023)		Region	Country HDI	Smart city dokument		Název dokumentu	Rok vydání strategického dokumentu	Platnost strategického dokumentu		Počet stran	Počet zmínek	
					Země				Jazyk				Má big data?	Oblast použití big data			
1	1	0,989	Zurich	Zurich	1 431 538	Švýcarsko	Evropa	0,962	Psaný text	Angličtina/němčina	Smart city strategie	2018	2035 (součást další strate	16	Ne		
2	2	0,98	Oslo	Oslo	1 085 992	Norsko	Evropa	0,961	Nemá								
3	3	0,98	Canberra	Canberra	472 304	Austrálie	Zbytek světa	0,951	Psaný text	Angličtina	Digitální strategie	2020	2022	31	Ne		
4	4	0,967	Kodaň	Copenhagen	1 381 005	Dánsko	Evropa	0,948	Prezentace	Angličtina a dánština	Smart city	2015	Neurčeno	50	Ano	10	Platforma, doprava, městské plánování, odpad
5	6	0,973	Londýn	London	9 648 110	Spojené království	Evropa	0,929	Psaný text	Angličtina	Smart city plán	Neurčeno	2020	54	Ano	5	Analýza dat, doprava
6	7	0,939	Singapur	Singapore	6 080 859	Singapur	Asie	0,939	Psaný text	Angličtina	Smart project	2018	Neurčeno	34	Ano	1	big data analytics
7	8	0,96	Helsinky	Helsinki	1 337 786	Finsko	Evropa	0,94	Nemá								
8	10	0,972	Stockholm	Stockholm	1 700 066	Švédsko	Evropa	0,947	Psaný text	Švédština	Smart city strategie	2021	2040	39	Ano	10	Doprava, platforma, analýza, AI, machine learning, real-time
9	11	0,972	Hamburg	Hamburg	1 787 520	Německo	Evropa	0,942	Psaný text	Angličtina	Digitální strategie	2020	Neurčeno	31	Ano	5	Big data analytics, IoT, městské cíle
10	12	0,907	Peking	Beijing	21 766 214	Čína	Asie	0,768	Prezentace	Čínština	Smart city	2019	Ne lze přeložit	41	Ne lze přeložit		
11	13	0,911	Abu Dhabi	Abu Dhabi	1 566 999	Spojené arabské emiráty	Zbytek světa	0,911	Prezentace	Angličtina a arabština	Digitální transformační plán	2019	2029+	79	Ano		Big data analytics, big data management, big data engines
12	14	0,96	Praha	Prague	1 323 339	Česká republika	Evropa	0,889	Prezentace	Čeština	Koncept smart city	2017	2030	88	Ano	20	Turismus, big data analytics, sociální média, firmy, doprava, energie,
13	15	0,962	Amsterdam	Amsterdam	1 174 025	Nizozemsko	Evropa	0,941	Psaný text	Holandština	Digitální město	2019	2025	24	Ne		
14	16	0,952	Soul	Seoul	9 988 049	Korea	Asie	0,925	Prezentace	Angličtina	Smart city status a strategie	Neurčeno	Neurčeno	21	Ano	1	E-vláda
15	19	0,946	Hongkong	Hong Kong	7 684 801	Honk Kong	Asie	0,949	Psaný text	Angličtina/čínština	Smart city blueprint	2020	Neurčeno	36	Ano	3	Big data analytics, real-time
16	21	0,938	New York	New York	7 888 121	USA	Severní Amerika	0,921	Psaný text	Angličtina	Budování smart city	2015	Neurčeno	24	Ne		
17	22	0,951	Auckland	Auckland	1 673 220	Nový Zéland	Zbytek světa	0,937	Nemá								
18	26	0,959	Reykjavik	Reykjavik	137 618	Island	Evropa	0,959	Prezentace	Islandština	Strategie zaměstnanosti a inovací	2022	2028	27	Ne		
19	29	0,916	Taipei	Taipei City	2 754 196	Tchaj-wan	Asie	0,916	Psaný text	Čínština a angličtina	Smart city	Neurčeno	Neurčeno	35	Ano	4	Big data analytics, AI, policie, zásoby vody
20	30	0,9	Rijád	Riyadh	7 682 430	Saúdská arábie	Zbytek světa	0,875	strategie	Pouze státní							
21	34	0,949	Boston	Boston	617 459	USA	Severní Amerika	0,921	Nemá								
22	39	0,94	Washington D.C.	Washington D.C.	631 693	USA	Severní Amerika	0,921	Nemá								
23	40	0,943	Ottawa	Ottawa	1 010 391	Kanada	Severní Amerika	0,936	Psaný text	Angličtina	Smart city	2017	Neurčeno	30	Ano	3	Analýza dat, big data analytics
24	42	0,944	Vancouver	Vancouver	672 857	Kanada	Severní Amerika	0,936	Prezentace	Angličtina	Digitální strategie	2013	2016+	36	Ne		
25	48	0,943	Toronto	Toronto	2 903 456	Kanada	Severní Amerika	0,936	Psaný text	Angličtina	Digitální infrastruktura - strategický rámec	2022	Neurčeno	108	Ne		
26	50	0,931	Los Angeles	Los Angeles	3 769 485	USA	Severní Amerika	0,921	Psaný text	Angličtina	Smart city	2020	2028	54	Ne		
27	53	0,942	Denver	Denver	699 288	USA	Severní Amerika	0,921	Psaný text	Angličtina	Smart city challenge	2016	Neurčeno	33	Ne		
28	55	0,94	Seattle	Seattle	725 487	USA	Severní Amerika	0,921	Psaný text	Angličtina	Smart city	2020	Neurčeno	32	Ano	3	Big data analytics, big data systems, doprava
29	59	0,855	Doha	Doha	658 344	Katar	Zbytek světa	0,855	Nemá								
30	61	0,929	Chicago	Chicago	2 608 425	USA	Severní Amerika	0,921	Prezentace	Angličtina	IT strategický plán	2021	Neurčeno	167	Ne		
31	72	0,951	Tokyo	Tokyo	37 194 105	Japonsko	Asie	0,925	Prezentace	Japonština	Implementační strategie smart city	2019	2040	69	Ano	5	Ekonomika, sociální problémy, rozvoj města, řešení lokálních problémů, doprava, výzkum
32	88	0,839	Bangkok	Bangkok	11 069 982	Thajsko	Asie	0,8	Psaný text	Thajština	Development s digitálními technologiemi	2018	2022	12	Ne		
33	89	0,858	Kuala Lumpur	Kuala Lumpur	8 621 724	Malajsie	Zbytek světa	0,803	Prezentace	Angličtina	Smart city plán	2021	2025	21	Ne		
34	91	0,919	Tel Aviv	Tel Aviv	4 420 855	izrael	Zbytek světa	0,919	nepřístupný								
35	96	0,816	Maskat	Muscat	1 650 319	Omán	Zbytek světa	0,816	strategie	Pouze státní							
36	100	0,744	Hanoj	Hanoi	5 253 385	Vietnam	Asie	0,703	Nemá								
37	102	0,759	Jakarta	Jakarta	11 248 839	Indonésie	Asie	0,705	Prezentace	Angličtina	Budování smart city	Neurčeno	Neurčeno	20	Ano	2	Digitální technologie a inovace, analýza
38	105	0,73	Dillí	Delhi	32 941 309	Indie	Asie	0,633	Nemá								
39	117	0,896	Nikósie	Nicosia	412 156	Kypř	Zbytek světa	0,896	Prezentace	Řečtina	Smart City strategie	2018	2028	19	Ne		
40	118	0,757	Medellín	Medellín	4 102 308	Kolumbie	Zbytek světa	0,752	Prezentace	Španělština	Smart city	2021	2030	31	Ne		

Další sloupce jsou již odvozeny z analýzy vybraných dokumentů. Sekce začíná sloupcem *Smart City dokument*. V tomto sloupci byly identifikovány dva hlavní typy, a to buď psaný text nebo prezentace. Další hodnotou v tomto sloupci je to, zda město má pouze státní (národní) Smart City strategii. Speciální hodnotou zde je, že město strategii sice má, ale je z ČR nepřístupná. Poslední možnou hodnotou sloupce je, že strategii nemá.

Dalším sloupcem je sloupec *Jazyk*, ve kterém je dokument napsaný. Zde je důležité dát pozor na znaménko / a písmeno *a*. Zatímco lomítko znamená, že je k dispozici verze ve více jazycích, písmeno *a* znamená, že dokument je z části napsaný v prvním a z části ve druhém jazyce. Následuje sloupec *Název dokumentu*, kde byly vytyčeny určité typy názvů, se kterými se dále v problematice pracuje. Některé typy se opakují, některé strategie mají více speciální názvy a celkově se zde pracuje se 17 různými druhy názvů v rámci 28 měst. Další dva sloupce souvisí s *Rokem vydání a Platností dokumentů*. U těchto dvou sloupců, a hlavně u druhého zmíněného, došlo k velkému výskytu chybějících hodnot. Následující sloupec pak zmiňuje *Počet stran dokumentu*. Poslední tři sloupce se věnují big data tématice, tzn. jestli se v dokumentu big data vůbec zmiňují, jestli ano, tak kolikrát, a v poslední řadě v jakých oblastech. Na závěr je důležité dodat, že u města Curych se pracuje s německou verzí dokumentu a u města Hongkong se pracuje s anglickou verzí.

V Tabulka 5 jsou k dispozici názvy vybraných strategických dokumentů pro každé město. Ze 40 vybraných měst jich pouze 30 mělo vlastní strategii. Tabulka je řazena podle pořadí měst v úvodní tabulce. V tabulce není přítomna strategie pro Peking s názvem *Smarter Beijing 2019* (v čínštině), protože se s ní nedalo k nemožnosti překladu, kopírování a vyhledávání v textu pracovat. Taktéž chybí strategie pro Tel Aviv, která je z lokace ČR nedostupná.

Tabulka 5: Přehled měst a jejich strategií. Zdroj: *Vlastní zpracování*.

Český název	Dokument
Curych	Strategie SMART CITY ZÜRICH
Canberra	ACT Digital Strategy
Kodaň	Copenhagen Smart City: The challenge
Londýn	SMART LONDON PLAN
Singapur	Smart Nation: The Way Forward
Stockholm	Strategi för Stockholm som smart och uppkopplad stad
Hamburg	DIGITAL STRATEGY FOR HAMBURG
Abu Dhabi	THE 3 DIMENSIONS Digital TRANSFORMATION PLAN
Praha	Koncepce Smart Prague do roku 2030

Český název	Dokument
Amsterdam	Een Digitale Stad voor én van iedereen
Soul	Smart Seoul Status & Strategies
Hongkong	Hong Kong Smart City Blueprint 2.0
New York	BUILDING A SMART + EQUITABLE CITY
Reykjavík	Nýsköpun alls staðar
Taipei	Smart Taipei, One City
Ottawa	Smart city 2.0
Vancouver	City of Vancouver: Digital Strategy
Toronto	DIGITAL INFRASTRUCTURE: STRATEGIC FRAMEWORK
Los Angeles	SmartLA 2028: Technology for a better Los Angeles
Denver	BEYOND TRAFFIC: DENVER THE SMART CITY CHALLENGE
Seattle	Smart Seattle: A Prototype for the New Century's Digital City
Chicago	2021 IT Strategic Plan
Tokyo	スマート ~ 東京版 東京実施戦略 Society 5.0 の実現に向けて ~
Bangkok	การพัฒนาพื้นที่กรุงเทพมหานคร ด้วย เทคโนโลยีดิจิทัล เพื่อบรรลุวิสัยทัศน์ การเป็นมหานครแห่งเอเชีย
Kuala Lumpur	KUALA LUMPUR: SMART CITY MASTER PLAN
Jakarta	Building a SMART CITY 4.0: A Collaboration for a Better Jakarta
Nikósie	Nicosia Municipality Presentation Smart City
Medelín	MEDELLÍN: CIUDAD INTELIGENTE

3.4 Zápis dat do tabulky big dat

Dalším krokem obsahové analýzy byl zápis dat do druhé tabulky, viz Tabulka 6. Sem vstoupila tedy jen města, která obsahují big data, resp. se nějak ve své strategii věnují této problematice. Druhá tabulka byla z většiny vytvořena na základě posledního sloupce z úvodní tabulky, tedy ve smyslu, že sloupce byly identifikovány na základě oblastí, ve kterých se zmiňovala big data.

První čtyři sloupce jsou všeobecné, kdy se zde vyskytly sloupce pořadí v indexu, český název, země, region a v jakém formátu je Smart City dokument. Další sloupce se pak věnují oblastem big dat. Oblasti jsou rozděleny do *Dopravy*, *Big data analytics* (včetně analýzy dat), *Věci spojené s řízením města* (obsahující městské plánování, projekty samosprávy, městské cíle, rozvoj města, odpadní hospodářství a policii), *Ekonomika a business* a *Technologické zmínky* (zahrnující IoT, real-time data, AI a machine learning). Tabulku pak uzavírá sloupec *Další specifické vlastnosti*, který obsahuje oblasti využití big dat zmíněné jen u jednoho konkrétního města.

Tabulka 6: Tabulka big dat. Zdroj: *Vlastní zpracování, MS Excel.*

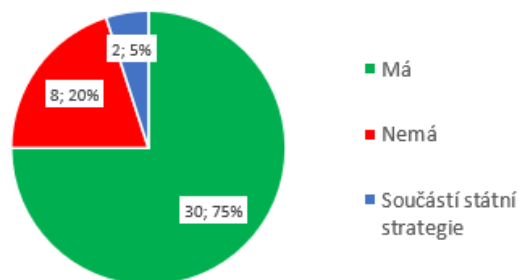
Pořadí v indexu	Český název	Země	Region	Smart city dokument	Doprava	Big data analytics	Věci spojené s řízením města	Ekonomika, business	Technologické zmínky	Další pojmy související s Big data	
										problematikou	Další specifické vlastnosti
4	Kodaň	Dánsko	Evropa	Prezentace	1	0	1	0	0		Řízení odpadu
6	Londýn	Spojené království	Evropa	Psaný text	1	1	0	0	0		x
7	Singapur	Singapur	Asie	Psaný text	0	1	0	0	0		x
10	Stockholm	Švédsko	Evropa	Psaný text	1	1	0	0	1		x
11	Hamburg	Německo	Evropa	Psaný text	0	1	1	0	1		x
13	Abu Dhabi	Spojené arabské	Zbytek světa	Prezentace	0	1	0	0	0		x
14	Praha	Česká republika	Evropa	Prezentace	1	1	0	1	0		Turismus, sociální média, energie
16	Soul	Korea	Asie	Prezentace	0	0	1	0	0		x
19	Hongkong	Honk Kong	Asie	Psaný text	0	1	0	0	1		x
29	Taipei	Tchaj-wan	Asie	Psaný text	0	1	1	0	1		Zásoby vody
40	Ottawa	Kanada	Severní Amerika	Psaný text	0	1	0	0	0		x
55	Seattle	USA	Severní Amerika	Psaný text	1	1	0	0	0		x
72	Tokyo	Japonsko	Asie	Prezentace	1	0	1	1	0		Výzkum
102	Jakarta	Indonésie	Asie	Prezentace	0	1	0	0	1		0

4 ANALÝZA A VYHODNOCENÍ ZÍSKANÝCH DAT

Tato kapitola se v jednotlivých dílčích kapitolách zabývá analýzou a vyhodnocením získaných dat, nejprve pomocí grafů v programu MS Excel a následně ve formě modelů vytvořených v programovacím jazyku Python.

4.1 Porovnání strategií

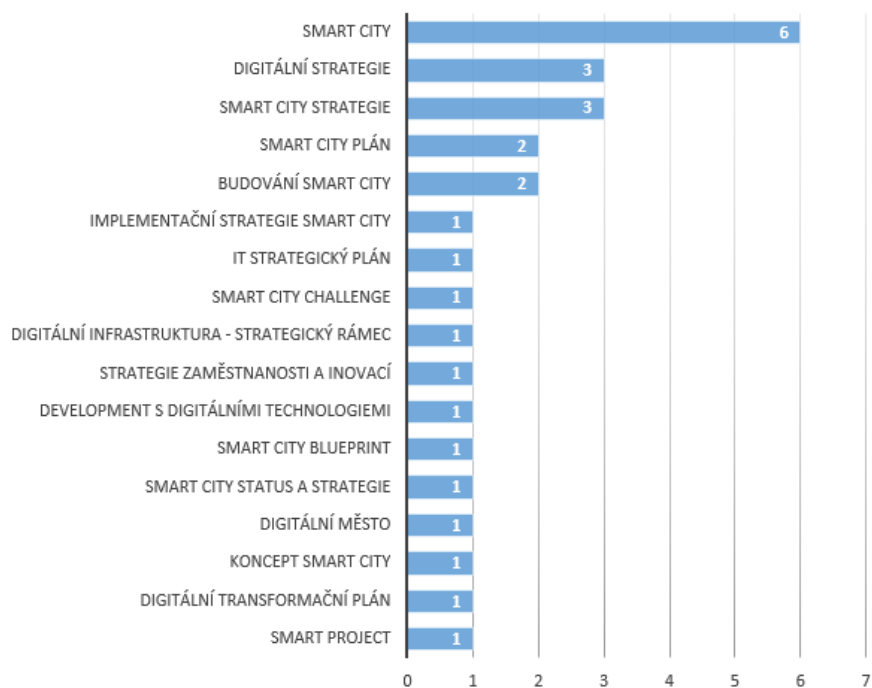
Jak lze vidět na Obrázek 9, ne všechna města mají vlastní Smart City strategii. Konkrétně jen 75 % z vybraných měst má svůj vlastní stahovatelný dokument pro strategii. Zde je třeba také zmínit dvě města a to Peking, který strategii sice má, ale dokument nelze v našich podmínkách přeložit, a Tel Aviv, který dokument pravděpodobně má, ale z ČR je nepřístupný. V dalších kapitolách se tedy pracuje pouze s výběrem 28 měst, které strategii mají, tedy výběr ze 30 měst bez zmíněných dvou výjimek.



Obrázek 9: Porovnání strategií měst. Zdroj: *Vlastní zpracování, MS Excel.*

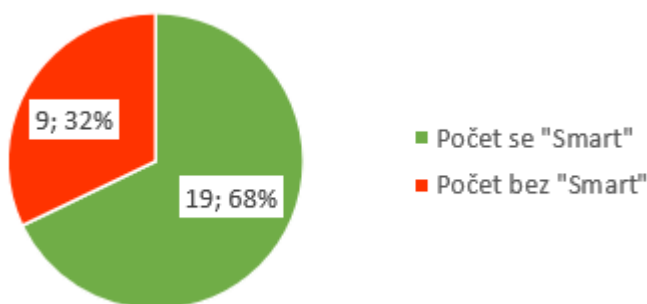
4.2 Název dokumentu

Jedním ze zkoumaných atributů je název dokumentu, kdy bylo analyzováno, zda se mezi všemi názvy vyskytuje nějaký vzor, resp. určující klíčová slova. Jejich přehled je k dispozici viz Obrázek 10. Nejčastější název je ve smyslu Smart City (zde se bere v úvahu jak například typ názvu Smart Seattle, tak Medelín: Smart City) kdy se tento název objevil 6×. Další podobnost, a to vždy po třech výskytech, bylo identifikováno u názvů Smart City strategie a Digitální strategie. Poslední zmiňované v sobě neobsahuje názvosloví Smart City, nicméně dané strategie nesly znaky typické pro Smart City strategie. Po dvou se vyskytly názvy Smart City plán a Budování Smart City. Další názvy už byly unikátní a vyskytly se pouze jednou.



Obrázek 10: Názvy dokumentů. Zdroj: *Vlastní zpracování, MS Excel.*

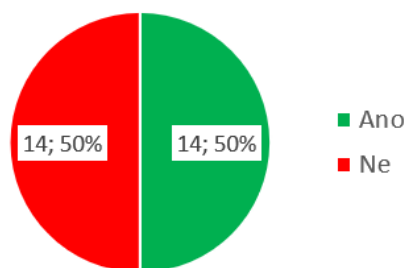
Pohledem na Obrázek 11 lze zjistit, že z celkového počtu pouze 19 strategií obsahuje ve svém názvu slovo „Smart“. Ostatní dokumenty mají různé názvy, kdy se převážně zmiňuje pojem digitalizace, respektive ve tvaru *digitální*. I přes jiné názvy však dané dokumenty nesly, alespoň částečně, znaky Smart City strategií.



Obrázek 11: Počet výskytů s pojmem „Smart“ v názvu. Zdroj: *Vlastní zpracování, MS Excel.*

4.3 Výskyt big dat

Obrázek 12 zobrazuje výskyty klíčového slova big data v daném dokumentu. Z grafu lze vyčíst, že přesně 50 %, měst zmiňuje ve svém strategickém dokumentu big data.

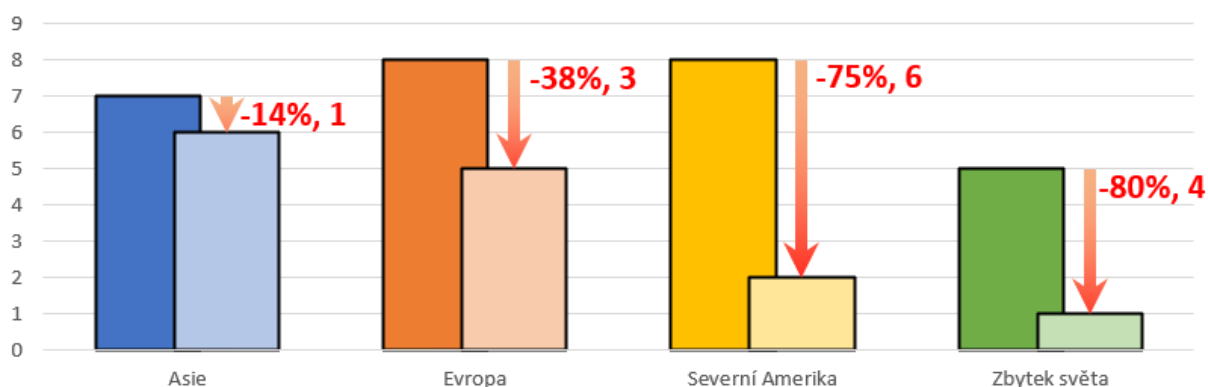


Obrázek 12: Výskyt big dat ve strategiích. Zdroj: *Vlastní zpracování, MS Excel.*

4.4 Počty z regionů

Ačkoliv se práce věnuje 40 městům, kdy každý region byl omezen na 10 měst, výběr pro analýzu byl nejprve zúžen z důvodu neexistence vlastní strategie města a poté pak na základě výskytu klíčového slova big data. Tato kapitola se tedy věnuje znázornění či změně proporcí po těchto úpravách. Na Obrázek 13 je v levém sloupci k vidění výběr z měst majících strategii. Evropa a Severní Amerika mají v zúženém výběru shodně po osmi městech, Asie má o jedno méně a zbytek světa má jenom pět čili polovinu z původního počtu. Dříve bylo zmíněno, že dvě města řeší svojí strategii pouze jako součást státní strategii, což se vyskytlo u dvou zemí z Blízkého východu patřících právě do zbytku světa. Také je nutné zde zmínit, že dvě výjimky (Peking, Tel Aviv) nejsou zahrnuty v tomto ani dalších obrázcích.

V pravém sloupci u každého regionu lze vidět pouze města zmiňující big data. Z obrázku je dále patrné, že pojem big data nejčastěji řeší města Asie a poté Evropy. Šipky v grafu zobrazují, jak se změnili počty měst po zaměření pouze na strategie obsahující big data. Dále lze vyčíst, že nejméně se big data řešila v rámci zbytku světa a pak Americe, nejvíce naopak v rámci Asie.



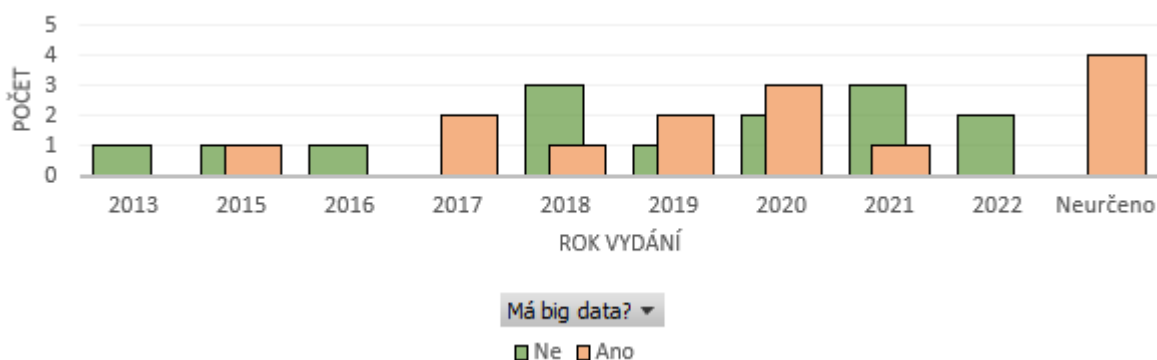
Obrázek 13: Porovnání počtů měst z regionů (s big daty proti bez big dat). Zdroj: *Vlastní zpracování, MS Excel.*

4.5 Rok vydání

Na Obrázek 14 jsou ukázány roky vydání u strategií. Na obrázku lze vidět, že nejvíce strategií bylo vydáno v roce 2020 (5 strategií), dále pak v roce 2018 (4 strategie), 2021 (4 strategie) a v roce 2019 byly vydány 3 strategie. V dalších rocích to je po dvou a méně. U čtyř měst nebylo jasné, ze kterého roku strategie pochází.

Pravý sloupec pak řeší rok vydání u měst obsahující big data. Nejstarší rok vydání je 2015 a to u Kodaně. Nejvíce dokumentů bylo vydáno v roce 2020 (3 strategie) a v počtu o jednu více pak nešlo určit rok vydání. Nejnovější dokument patří Stockholmu a pochází z roku 2021.

Pro porovnání je v levém sloupci znázorněná situace pro města a jejich strategie neobsahující big data. Nejvíce strategií bylo vydáno v letech 2018 a 2021. Nejnovější dokumenty pochází z roku 2022 a patří Torontu a Reykjavíku. Nejstarší dokument neobsahující big data z roku 2013 patří Vancouveru. Při porovnání těchto dvou sloupců lze vyvodit závěr, že big data nejsou trendem novějších strategií. U všech dokumentů neobsahujících data šlo určit, v jakém roce byly vydány.

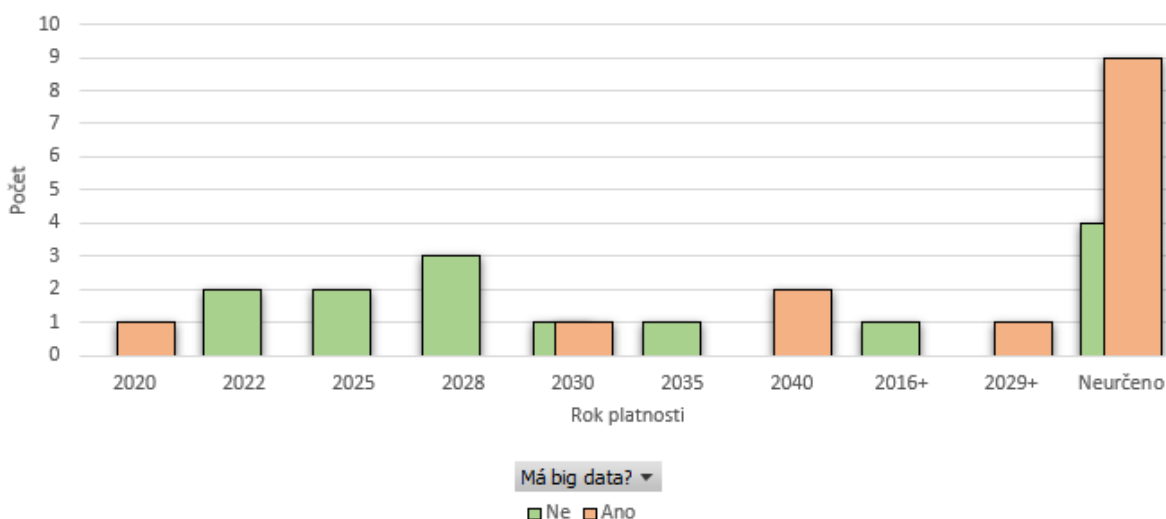


Obrázek 14: Roky vydání. Zdroj: *Vlastní zpracování, MS Excel*

4.6 Rok platnosti

Předchozí kapitola porovnávala rok vydání, tato kapitola naopak porovnávala roky platnosti (rozmezí let), pro které byly strategické dokumenty vytvořeny. To lze vidět viz Obrázek 15, kde 13 dokumentů nemělo žádnou omezenou dobu platnosti. Tři města mají podle odvozených údajů již po platnosti dokumentů, jeden dokument patřící Londýnu měl platnost do roku 2020. Dva další dokumenty, patřící městům Canberra a Bangkok, pak mají platnost určenou do roku 2022. Dva další dokumenty pak mají stanovenou platnost na konkrétní rok a dále, na obrázku znázorněné jako rok se znaménkem +. Jedná se tedy o rok 2016 a více u Vancouveru a o rok 2029 a více u Abu Dhabi. Nejdělsí rok platnosti je 2040, a to u Tokya.

Pravý sloupec se odkazuje opět jen na města obsahující pojem big data. Zde jde opět vidět, že nejvíc měst nemá určený rok platnosti u svého dokumentu. Jeden dokument je už po své platnosti, a to Londýn. Pro porovnání jsou v levém sloupci znázorněna města neobsahující big data. Horní hranice platnosti je rok zde 2035, kdy Smart City strategie je součástí dalšího strategického dokumentu, který má rok platnosti do této doby. Jedná se o švýcarské město Curych.



Obrázek 15: Roky platnosti strategií. Zdroj: *Vlastní zpracování, MS Excel.*

4.7 Forma dokumentu

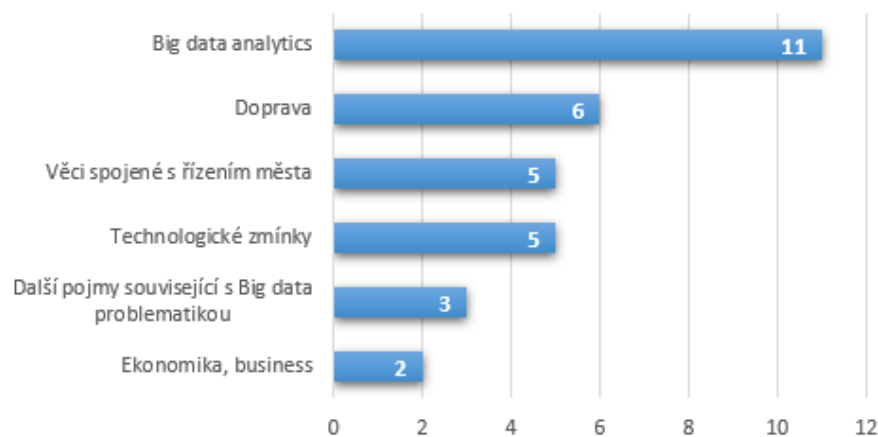
Tato kapitola je věnována tomu, jakým způsobem je napsána strategie města. Zde se vyskytují dva způsoby, a to buď strategie zaznamenané ve stylu prezentace anebo ve stylu psaného textu. Rozdělení je záleženo na tom, jak byla stránka vedena a kolik plochy zabíral text v porovnání s obrázky či ukázkami. Jak ukazuje Tabulka 7, více strategií je zaznamenáno jako psaný text, konkrétně 57 % (16 v absolutním čísle), oproti 12 strategiím vedeným spíše ve stylu prezentace. Lze vidět, že i v případě srovnání měst obsahujících big data s městy neobsahující big data došlo ke stejnému poměru.

Tabulka 7: Forma dokumentu. Zdroj: *Vlastní zpracování.*

Forma strategie	Města mající strategie	Města obsahující big data	Města neobsahující big data
Prezentace	12; 43 %	6; 43 %	6; 43 %
Psaný text	16; 57 %	8; 57 %	8; 57 %

4.8 Použití v oblastech

Následující Obrázek 16 ukazuje, jaké oblasti byly nejvíce používanými v souvislosti s big daty. Jedenáct měst zmínilo big data v souvislosti s Big data analytikou. Druhou nejpoužívanější oblastí je doprava, která byla zmíněna 6×. S pěti zmínkami následují věci spojené s řízením města (zahrnuje městské plánování, projekty samosprávy, městské cíle, rozvoj města a policii) a Technologické zmínky (včetně pojmů jako IoT, real-time, AI a machine learning). Následují Další pojmy související s big data problematikou (zahrnuje např. big data management, big data engines, big data platformy nebo big data systémy) a přehled zakončuje oblast Ekonomika a business.



Obrázek 16: Součty použití big dat v oblastech. Zdroj: *Vlastní zpracování, MS Excel.*

4.9 Analýza v Pythonu

Pro analýzu měst v Pythonu bylo vybráno zkoumání pomocí výpočtu korelace atributů a použití shlukové a textové analýzy. Součástí kapitoly je i popisná statistika vytvořená v Pythonu. Kódy k jednotlivým částem a modelům jsou k nalezení v přílohách diplomové práce, tzn. Přílohy III, IV, V a VI.

4.9.1 Příprava dat

Analýza v Pythonu vychází ze stejných souborů, jako je tomu v případě porovnání v Excelu. Nicméně vzhledem k nutnosti různých výpočtů je třeba dané zdrojové tabulky upravit. Jednak je třeba tabulky obecně upravit a poté je nutné připravit data i v rámci jednotlivých modelů.

V přípravě dat u úvodní tabulky se odstranil sloupec „ID“ (identifikátor je použit jen v rámci diplomové práce), „Město“ (tabulka obsahuje český název, který se používá místo tohoto sloupce), „Země“ a „Oblast použití big dat“ z důvodu nepotřebných údajů pro analýzu pomocí Pythonu. Dále byl odstraněn sloupec „Platnost strategického dokumentu“ z důvodu, že mnoho

hodnot bylo nastaveno na Neurčeno. Dále byla u hodnot obsahujících desetinná místa zaměněna čárka za tečku. U sloupce „Rok vydání strategického dokumentu“ byly naopak hodnoty Neurčeno (objevy 4×) nastaveny nejprve na prázdné, poté byly pomocí Pythonu dopočítány chybějící hodnoty průměrnou hodnotou, která v tomto případě činí 2018. U sloupce počet zmínek big dat byla nulová hodnota (objevy 14×) nastavena na hodnotu 0. Soubor byl také převeden z XLSX do formátu CSV.

Také pro druhou tabulku big dat byl formát změněn na CSV. V rámci přípravy tabulky big dat došlo k odstranění sloupce „Další specifické vlastnosti“, který není vhodný pro další analýzu. Poté byl přidán sloupec z první tabulky „Počet zmínek big data“. Další změny v této tabulce v této fázi neproběhly, konkrétní úpravy a výběru dat z tabulky se pak věnují další kapitoly.

4.9.2 Popisná statistika

Jako první jsou v Tabulka 8 zapsané popisné statistiky pro úvodní tabulku. Je třeba zmínit, že atribut „Počet zmínek big data“ je nereprezentativní kvůli záznamům s nulovou hodnotou. Nicméně se tento atribut vyskytuje i v popisné statistice pro tabulku big dat, která je dále. Nejzajímavější výsledky se asi týkají atributu roku vydání, či počtu stran, který je ale ovlivněn tím, že strategie jsou rozdělené na prezentaci nebo psaný text. Po rozdělení lze zjistit, že typ prezentace má v průměru o 15 stránek více. Konkrétně je u prezentace v průměru 52 stránek, u psaného textu je to stránek 37.

Tabulka 8: Výstupy popisných statistik z 28 měst úvodní tabulky. Zdroj: *Vlastní zpracování.*

Stat.	Pořadí v indexu	City HDI	Počet obyvatel (2023)	Country HDI	Rok vydání strategického dokumentu	Počet stran	Počet zmínek big data
COUNT	28	28	28	28	28	28	28
MEAN	40,36	0,93	5002037,61	0,91	2018,64	43,61	2,75
STD	35,74	0,06	7276888,12	0,06	2,18	33,17	4,46
MIN	1	0,76	137618	0,705	2013	12	0
25%	12,50	0,93	1133116,5	0,91475	2018	24	0
50%	27,50	0,94	2197972,5	0,925	2018,50	32,50	0,50
75%	56,50	0,96	7735631	0,94125	2020	51	4,25
MAX	118	0,99	37194105	0,962	2022	167	20

Tabulka 9 poté zobrazuje situaci pro tabulku big dat. Co se týká atributu „Počet zmínek big data“, tak v průměru je ve strategii zmíněn 5,5×, medián pak má hodnotu 4,5. Největší hodnota výskytu je 20, kde se jedná o Prahu. Porovnáním pořadí v indexu s předchozí tabulkou jsou zde

nižší hodnoty. Z toho by se dalo usoudit, že města řešící ve Smart City strategii problematiku big dat se nacházejí v indexu výše.

Tabulka 9: Výstupy popisných statistik ze 14 měst tabulky big dat. Zdroj: *Vlastní zpracování.*

Statistika	Pořadí v indexu	Počet zmínek big data	Doprava	Big data analytics	Věci spojené s řízením města	Ekonomika, business	Technologické zmínky	Další pojmy související s Big data problematikou
COUNT	14	14	14	14	14	14	14	14
MEAN	28,43	5,50	0,43	0,79	0,36	0,14	0,36	0,21
STD	29,15	5,00	0,51	0,43	0,50	0,36	0,50	0,43
MIN	4	1	0	0	0	0	0	0
25%	10,25	3	0	1	0	0	0	0
50%	15	4,50	0	1	0	0	0	0
75%	37,25	5	1	1	1	0	1	0
MAX	102	20	1	1	1	1	1	1

4.9.3 Korelace

Kód v Pythonu pro korelaci je obsažen v Příloze V. U hodnot došlo ke konverzi textu na čísla, konkrétně tomu bylo u atributů „*Smart City dokument*“, „*Má big data?*“ a „*Region*“. U prvního zmiňovaného byla hodnota *Psaného textu* nahrazena hodnotou nula, zatímco *Prezentace* byla nahrazena hodnotou jedna. U „*Má big data?*“ byla nahrazena odpověď „*Ne*“ nulou a odpověď „*Ano*“ jedničkou. U posledního sloupce byla Evropa označena jako 0, Asie jako 1, Zbytek světa jako 2 a Severní Amerika číslem 3. Dále byla na datech provedena standardizace a normalizace.

Tabulka 10 zobrazuje situaci pro úvodní tabulku. V ní je vidět zvýšená korelace mezi pořadím v indexu a City či Country HDI, což je očekávaná korelace. Další případ je mezi City HDI a Country HDI, což je také očekávaný závěr. Další hodnoty jsou už spíše zanedbatelné, nicméně jde vidět slabší vliv mezi regionem a pořadím v indexu a regionem a vlastností „*Má big data?*“. Zajímavé jsou však záporné korelační koeficienty mezi některými atributy, obzvláště pořadím v indexu a hodnotami HDI, ze kterých lze usuzovat, že čím vyšší je HDI, tím nižší je pořadí Smart City v indexu. Tento vztah by mohl být vstupem pro další podrobnější výzkum. Nicméně

lze předpokládat, že hodnotící rámce Smart City indexů jsou nastaveny komplexně, a když se město zaměří jen na jednu oblast, resp. atribut, tak to nepovede k vyšší pozici v pořadí indexu.

Tabulka 10: Výstup korelace 28 měst úvodní tabulky. Zdroj: *Vlastní zpracování.*

	Pořadí v indexu	Region	City HDI	Počet obyvatel (2023)	Country HDI	Smart City dokument	Rok vydání str. dokumentu	Počet stran	Má big data?
Pořadí v indexu	1	0,486	-0,754	0,141	-0,714	0,285	0,079	-0,024	-0,410
Region		1	-0,279	-0,128	-0,144	-0,060	-0,110	0,162	-0,329
City HDI			1	-0,211	0,905	-0,311	-0,029	0,034	0,140
Počet obyvatel (2023)				1	-0,278	0,150	-0,015	-0,014	0,152
Country HDI					1	-0,377	0,003	0,015	0,174
Smart City dokument						1	-0,090	0,267	-0,055
Rok vydání str. dokumentu							1	0,186	-0,058
Počet stran								1	0,048
Má big data?									1

V následující Tabulka 11 je korelace pro tabulku big dat. Zde jde vidět lehce zvýšená korelace mezi atributem „Počet zmínek big data“ a „Big data analytics“ a taktéž mezi prvním zmíněným a dopravou. Hodnota skoro 0,6 je i mezi atributy „Big data analytics“ a „Věci spojené s řízením města“. I v této tabulce je několik záporných korelací mezi atributy, když v této tabulce, kde je pouze 14 měst, lze některé vztahy považovat za diskutabilní s ohledem na velikost vzorku, resp. by tato zjištění měla být ověřena v dalším výzkumu.

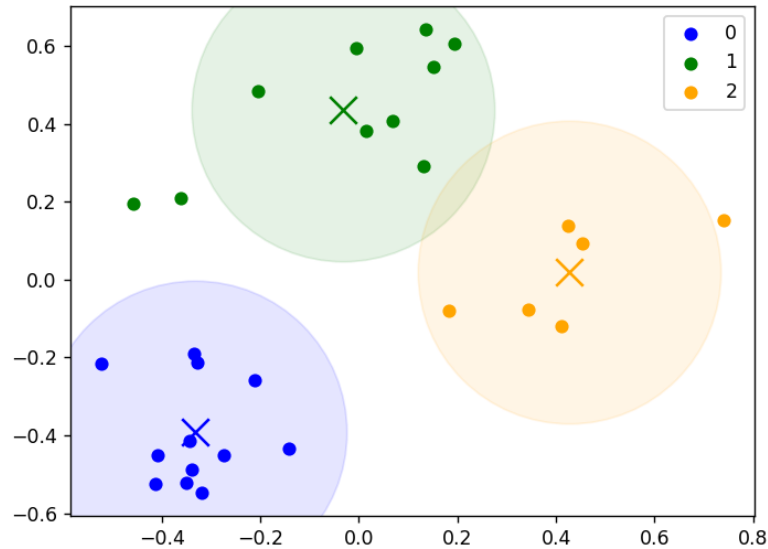
Tabulka 11: Výstup korelace 14 měst tabulky big dat. Zdroj: *Vlastní zpracování.*

	Pořadí v indexu	Region	Smart City dokument	Počet zmínek big data	Doprava	Big data analytics	Věci spojené s řízením města	Ekonomika, business	Technologické zmínky	Další pojmy big data problematiky
Pořadí v indexu	1	0,427	0,301	-0,167	-0,028	-0,053	-0,014	0,273	0,150	0,022
Region		1	-0,017	-0,410	-0,318	0,145	-0,294	-0,117	-0,290	0,310
Smart City dokument			1	0,311	0,137	-0,658	0,244	0,444	-0,323	0,134
Počet zmínek big data				1	0,619	-0,098	0,055	0,552	-0,018	0,212
Doprava					1	-0,252	-0,083	0,406	-0,336	0,282
Big data analytics						1	-0,592	-0,321	0,347	0,166
Věci spojené s řízením města							1	0,129	0,211	-0,309
Ekonomika, business								1	-0,225	-0,085
Technologické zmínky									1	-0,085
Další pojmy big data problematiky										1

4.9.4 Shluková analýza

Kód pro shlukovou analýzu lze nalézt v Příloze VI. U hodnot došlo ke stejné konverzi textových hodnot na čísla, jako tomu bylo u korelační analýzy. Na datech taktéž byla udělána normalizace a standardizace.

Obrázek 17 ukazuje situaci pro úvodní tabulku. Pro ni se jeví jako optimální rozdělení do tří shluků. Při čtyřech shlucích se shluky začali překrývat přes sebe. Na obrázku lze vidět, že optimální je pouze modrý shluk 0. Ostatní shluky mají v sobě zahrnuta města, které jsou umístěna daleko od středu.



Obrázek 17: Shlukování 28 měst úvodní tabulky. Zdroj: *Vlastní zpracování, PyCharm.*

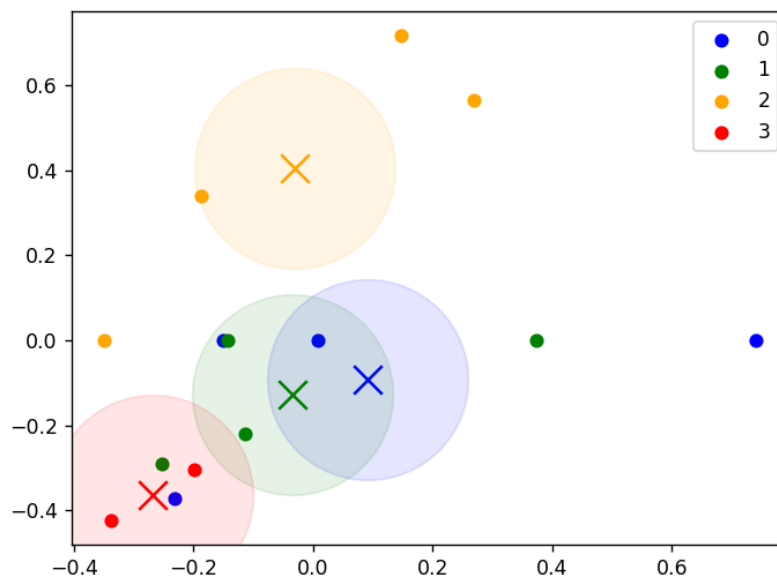
V následující Tabulka 12 lze vidět rozdělení měst do jednotlivých shluků. Nejpočetnější je Shluk 0, nejméně početný naopak Shluk 2. Závěry jsou takové, že sloupce „*Má big data?*“ či „*Počet zmínek big data*“ neměly na rozdělení do shluků takový vliv. Shluk 0 obsahuje všechna evropská města a s výjimkou tří i asijská města. Shluk 1 poté obsahuje všechna města Severní Ameriky, k nimž se přidávají dvě města ze zbytku světa, a to Canberra a Abu Dhabi. Ačkoliv se s atributem jazyka ve shlukování nepracovalo, je zajímavé, že pro druhý shluk je typické, že všechna města mají jako jazyk angličtinu, čímž se odlišují právě dvě zmíněná města ze zbytku světa od zbylých tří, které jsou ve Shluku 2. Roli ve shlukování ovšem hrálo také pořadí v indexu a City HDI s pár výjimkami. Shluk 0 měl průměr pořadí v indexu 13,17 (medián 12,5), Shluk 1 pak měl průměr pořadí v indexu 38,6 (medián 45) a poslední shluk měl průměrnou hodnotu 97,67 (medián 95,5). Co se týče City HDI, Shluk 0 měl průměrnou hodnotu 0,959 (medián 0,961), Shluk 1 měl průměr 0,940 (medián 0,941) a Shluk 2 nabýval průměru 0,843 (medián 0,849). Shluk 1 a Shluk 2 však měly svoje výjimky, které se značně lišily od uvedených hodnot. V případě prvního zmíněného to byla Canberra a v případě Shluku 2 pak Tokyo.

Tabulka 12: Rozřazení do shluků (úvodní tabulka). Zdroj: *Vlastní zpracování*.

Shluk – 0	Shluk – 1	Shluk – 2
Curych	Canberra	Tokyo
Kodaň	Abu Dhabi	Bangkok
Londýn	New York	Kuala Lumpur
Singapur	Ottawa	Jakarta
Stockholm	Vancouver	Nikósie
Hamburg	Toronto	Medelín
Praha	Los Angeles	
Amsterdam	Denver	
Soul	Seattle	
Hongkong	Chicago	
Reykjavík		
Taipei		

Na Obrázek 18 jsou výsledky shlukování pro tabulku big dat. Výsledky shlukování této tabulky nejsou dobré, shluky se při různých pokusech počtu shluků překrývaly přes sebe a některé případy byly výrazně odlehle od sebe. Jako optimální se jevílo shlukování o čtyřech shlucích, jelikož při pěti a více shlucích už došlo k situaci, kdy se v jednom shluku nacházelo pouze jedno město. Při pohledu na podrobnější vlastnosti shluků se však dají v oblastech použití big dat najít jisté vzory. Je možné vyvodit závěr, že do přiřazení do shluků vstupují vlastnosti nespojené s použitím big dat v určených oblastech, anebo některé specifické oblasti ovlivňují výsledné shlukování.

Na obrázku lze vidět, že nejbliže si jsou k sobě města červeného shluku (shluk 3). Nějaký vzor se dá pozorovat i ve shluku 2, kdy se v podstatě jedná o města ležící v levém horním rohu. Shluky 0 a 1 se překrývají, jedno město od každého z nich pak leží i v oblasti červeného shluku 3.



Obrázek 18: Shlukování 14 měst tabulky big dat. Zdroj: *Vlastní zpracování, PyCharm.*

V

Tabulka 13 je zaznamenáno rozřazení do shluků. Na obrázku výše bylo zjištěno, že města ve shluku 3 jsou si nejbližší, jedná se o Londýn a Stockholm, která byla ve stejném shluku také v případě shlukování úvodní tabulky. Dalším zmíněným shlukem byl shluk 2, kdy se jedná o města zhruba ve stejné oblasti, do kterých nezasahují jiné shluky. Zde se nachází Singapur, Abu Dhabi, Ottawa a Seattle, které se s výjimkou prvního zmíněného nacházely spolu rovněž i v prvním shlukování. Při pohledu na pořadí v indexu tento atribut nehrál takovou roli jako tomu bylo v případě prvního shlukování.

Tabulka 13: Rozřazení do shluků (tabulka big dat). Zdroj: *Vlastní zpracování.*

Shluk – 0	Shluk – 1	Shluk – 2	Shluk – 3
Hamburg	Kodaň	Singapur	Londýn
Hongkong	Praha	Abu Dhabi	Stockholm
Taipei	Soul	Ottawa	
Jakarta	Tokyo	Seattle	

Podrobnější popis obsahuje Tabulka 14, která zobrazuje vlastnosti Shluku 0. Čím se odlišuje od ostatních shluků je převážně atribut Technologických zmínek, kde všude má hodnotu 0, a použitím Big data analytics. Naopak, města ve své strategii nezahrnují použití v dopravě a Ekonomice a businessu.

Tabulka 14: Popis Shluku 0. Zdroj: *Vlastní zpracování.*

Shluk 0									
Pořadí v indexu	Český název	Region	Smart City dokument	Doprava	Big data analytics	Věci spojené s řízením města	Ekonomika, business	Technologické zmínky	Další pojmy související s Big data problematikou
11	Hamburg	Evropa	Psaný text	0	1	1	0	1	0
19	Hongkong	Asie	Psaný text	0	1	0	0	1	0
29	Taipei	Asie	Psaný text	0	1	1	0	1	0
102	Jakarta	Asie	Prezentace	0	1	0	0	1	0

Další shluk, zobrazený v Tabulka 15, se vyznačuje použitím big dat v oblasti dopravy, naopak tato města nezmiňují až na jedno z nich Big data analytics. Dalším znakem je, že všechna města v tomto shluku mají dokument ve formě prezentace.

Tabulka 15: Popis Shluku 1. Zdroj: *Vlastní zpracování.*

Shluk 1									
Pořadí v indexu	Český název	Region	Smart City dokument	Doprava	Big data analytics	Věci spojené s řízením města	Ekonomika, business	Technologické zmínky	Další pojmy související s Big data problematikou
4	Kodaň	Evropa	Prezentace	1	0	1	0	0	0
14	Praha	Evropa	Prezentace	1	1	0	1	0	0
16	Soul	Asie	Prezentace	0	0	1	0	0	0
72	Tokyo	Asie	Prezentace	1	0	1	1	0	0

Tabulka 16 popisuje Shluk 2. Tento shluk se vyznačuje především používáním big data analytics, přičemž další oblasti až na výjimky nevyužívá.

Tabulka 16: Popis Shluku 2. Zdroj: *Vlastní zpracování.*

Shluk 2									
Pořadí v indexu	Český název	Region	Smart City dokument	Doprava	Big data analytics	Věci spojené s řízením města	Ekonomika, business	Technologické zmínky	Další pojmy související s Big data problematikou
7	Singapur	Asie	Psaný text	0	1	0	0	0	0
13	Abu Dhabi	Zbytek světa	Prezentace	0	1	0	0	0	1
40	Ottawa	Severní Amerika	Psaný text	0	1	0	0	0	0
55	Seattle	Severní Amerika	Psaný text	1	1	0	0	0	1

Tabulka 17 zobrazuje situaci pro poslední shluk. Tento shluk obsahuje dvě evropská města, která mají strategický dokument ve formě psaného textu, s podobným pořadím v indexu. Obě města taktéž využívají big data v dopravě a v návaznosti na pojem big data analytics, naopak obě města nepoužívají big data ve spojení s řízením města či Ekonomikou a businessem.

Tabulka 17: Popis Shluku 3. Zdroj: *Vlastní zpracování.*

Shluk 3									
Pořadí v indexu	Český název	Region	Smart City dokument	Doprava	Big data analytics	Věci spojené s řízením města	Ekonomika, business	Technologické zmínky	Další pojmy související s Big data problematikou
6	Londýn	Evropa	Psaný text	1	1	0	0	0	0
10	Stockholm	Evropa	Psaný text	1	1	0	0	1	1

4.10 Projekty Smart City

Co se týká Smart City projektů, které jsou zpravidla součástí Smart City strategií, ale z důvodu omezení rozsahu diplomové práce nebyly podrobněji analyzovány, jsou stručně představeny v této kapitole. Při vyhledávání nějakých konkrétních projektů často dochází k problému toho,

že je složité dostat se k detailnějším dokumentům souvisejícím s touto problematikou. Dalším problémem je, že některé zdroje sice zmiňují big data v rámci nějakého projektu, nicméně po vyhledání projektu se v dalších zdrojích big data už nezmiňují. U některých projektů lze zase nalézt zprávy o plánování či začátku daného projektu, jenže pak už o nich nelze nalézt další informace. Zpráva k indexu SCG mapuje několik měst a obsahuje i několik krátkých popisů řešení v rámci různých oblastí, když došlo i na zmíněná témata big dat.

Prvním městem zmiňujícím big data je v této zprávě Soul, kde byla v roce 2022 nainstalována řada sensorů. Tyto sensory sbírají informace v různé škále městských elementů, jako například hluk, prach, intenzita nočního světla a dopravní provoz. Tato data jsou následně předána do integrovaného veřejného big data úložiště, které využívá služeb AI. Taktéž se zde zmiňuje veřejně-privátní big data platforma, přes kterou komunikují obyvatelé i privátní sektor, jako jsou např. telekomunikace, finance a logistické firmy. Dalším výskytem využití big dat je Šen-čen v Číně. Zde Huawei vymyslelo řešení využívající rozpoznávání obličeje a big data analytics k nahrazení manuální identifikace pasažérů na letišti. Posledním projektem je zde Khon Kaen v Thajsku. Cílem v Khon Kaenu je vývoj platformy, která by využívala blockchainu a big dat analytics k centralizaci klíčových zdravotních dat a zefektivnění svých procesů pro veřejné i soukromé poskytovatele zdravotních služeb (Eden Strategy Institute, 2021).

Co se týče ochrany dat, tak ta se zde zmiňuje pouze v souvislosti s Berlínem, ačkoliv je v úvodu zprávy ochrana dat zmiňována jako jedna výzva v souvislosti se Smart Cities. Berlínský senát považuje ochranu údajů za základní právo a základ svobodné demokratické společnosti. Klíčová zásada nového technologického nasazení v Berlíně spočívá v tom, že jeho použití je v souladu se zákony na ochranu údajů, aby si občané zachovali kontrolu nad svými osobními údaji (Eden Strategy Institute, 2021).

4.11 Ochrana dat ve strategiích

Ochraně dat není ve strategiích věnováno moc prostoru, tudíž se práce pokusí zmapovat tuto problematiku v rámci této podkapitoly. Pro zkoumání, resp. nalezení zmínek o ochraně dat ve Smart City strategiích byla brána v potaz města s big daty. Strategie Kodaně se nezmiňuje o ochraně dat. V Londýně se zmiňuje v souvislosti s „Oyster card“ (kartou se dají platit různé dopravní prostředky v Londýně), nicméně věta pouze zmiňuje, že jsou dodrženy standardy ochrany dat. Jaké standardy to jsou už zde není uvedeno. Singapur se naopak této problematice věnuje mnohem více. Stockholm taktéž problematice věnuje větší část své strategie. Hamburg se problematice věnuje ve dvou odstavcích. Nejprve je pojem použit ve významu obecné věty

potřeby vysoké úrovně ochrany dat, nicméně druhý odstavec se věnuje využití cloudových technologií a strategie apeluje na podniky a organizace, aby se aktivně podílely na ochraně dat na jejich platformách. Zbylá města s big daty, tedy Abu Dhabi, Praha, Seoul, Hong Kong, Taipei, Ottawa, Seattle, Tokyo a Jakarta se ve svých strategiích této problematice nevěnují vůbec.

4.11.1 Singapur

Ochraně dat jsou ve strategii věnovány dvě kapitoly, s přesahem i do třetí. První s názvem „*Ensuring reliable and secure cyber foundations*“ a následující nazvaná „*Maximising the value of data in trusted environment*“. První z nich se odkazuje na další dokumenty a strategie. Konkrétně zmiňovaná zde je *Singapore Cybersecurity Strategy* z roku 2016, v níž jsou vytyčeny priority pro budování odolné infrastruktury a vytvoření bezpečnějšího kybernetického prostoru. Mimo to byly vypracovány kyberbezpečnostní plány pro 11 sektorů kritických informačních infrastruktur (Critical Information Infrastructure, CII). V dalším odstavci jsou zmiňovány *Cybersecurity Act* a *CII Protection Programme* a další iniciativy v oblasti kyberbezpečnosti. V dokumentu je taktéž k dispozici krátký popis, co obnáší a čemu se věnuje jak National Cybersecurity Strategy 2016, tak Cybersecurity Act 2018. Dále je zde popsáno, že Smart Nation projekty spolupracují se Smart Nation Cybersecurity Committee k zajištění bezpečnosti již od návrhu, takže kybernetická bezpečnost je začleněna od samého počátku včetně poskytování poradenství ohledně zásad zabezpečení, návrhu a implementace. Na mezinárodní úrovni pak v rámci IoT spolupracuje s International IoT Security Initiative (Smart Nation and Digital Government Office, 2018).

Další kapitola se pak mimo kyberbezpečnosti věnuje už samotné ochraně dat, resp. se věnuje důležitosti dat obecně. Například v souvislosti s integrovaným data management frameworkem řeší, že nebude docházet k fúzi datasetů specifickým centrům pro ochranu dat. Infocomm Media Development Authority také vyvinula Data Protection Trustmark standards, který pomáhá podnikům ověřit, že dodržují standardy a postupy v oblasti ochrany osobních údajů. Certifikát potvrzuje platnost ochrany údajů v organizaci, čímž posiluje její pověst a konkurenceschopnost. Také zde existuje Personal Data Protection Act z roku 2012, který stanovuje základní standard ochrany osobních údajů v celé ekonomice. Poslední zmínka o ochraně dat je v souvislosti se spojením se světem, kdy Singapur je zapojen do Cross-Border Privacy Rules (CBPR) a Privacy Recognition for Processors (PRP) Systems. Společnosti, které získají certifikát CBPR nebo PRP, se pak budou odlišovat jako důvěryhodní zpracovatelé údajů, kteří dodržují regionálně schválené a uznávané standardy (Smart Nation and Digital Government Office, 2018).

4.11.2 Stockholm

V rámci strategického dokumentu Stockholmu je tomuto tématu věnována kapitola *Bezpečnost informací a ochrana soukromí*. Kapitola se v prvním odstavci věnuje tomu, že v rámci Smart City dochází ke vzniku velkého množství informací. Pokračuje popisem, že pro práci se správou a bezpečností informací již existují různé systémy a odborné znalosti, je však potřeba zajistit, aby všechna nová řešení byla v souladu s platnými směnicemi o bezpečnosti informací, kdy dále zmiňuje současné zákony, a hlavně nařízení EU o ochraně údajů. Další odstavec se věnuje přezkoumání pokynů pro bezpečnost a klasifikaci informací, zvláštní pozornost by se pak měla podle strategie věnovat nakládání s daty, u nichž se propojují informace z více zdrojů dat. Taktéž zmiňuje potřebu vyvinutí společných pracovních metod a postupů a společnou strukturu informací, aby byla zajištěna konzistentnost shromažďovaných informací. Třetí odstavec se věnuje rozdělení odpovědnosti za vlastnictví informací, které se shromažďují a využívají. Rovněž řeší otázky bezpečnosti s ohledem na komunikaci, která bude probíhat, a to především bezdrátově a na různých místech města. Poslední odstavec se věnuje otázce nového vývoje, kdy je při zpracování informací potřeba zvážit přínosy pro uživatele a osobní integritu. Otázky ochrany soukromí by měly být zohledněny ve všech nových řešeních vyvíjených v rámci Smart City. Pro řešení této problematiky by pak město Stockholm chtělo zřídit radu pro otázky ochrany soukromí a etické otázky související s novými digitálními řešeními, která by poskytovala pokyny v této oblasti. Podoba této rady, včetně členů a mandátu, by měla být definována jako součást prioritních projektů (Stockholms stad, 2021).

5 NÁVRH DOPORUČENÍ A OPATŘENÍ PRO DANOU PROBLEMATIKU

Tato kapitola obsahuje úvod pro návrh doporučení a opatření pro danou problematiku. Tvoří ji shrnutí zjištění identifikovaných v předchozí kapitole.

5.1 Problematika big dat v rámci Smart Cities

Při analýze strategií bylo v práci vytyčeno několik kategorií, ve kterých big data v rámci Smart Cities zaznívají. Typické oblasti, resp. komponenty, pro Smart City jsou: Smart lidé, Smart život, Smart průmysl/ekonomie, Smart síť (elektřina, teplo atd.), Smart zdravotnictví, Smart doprava, Smart samospráva, Smart prostředí a Smart bezpečnost. V souvislosti s použitím big dat byly identifikovány tyto oblasti: Big data analytics, Doprava, Věci spojené s řízením města (zahrnující městské plánování, projekty samosprávy, městské cíle, rozvoj města a policie) Technologické zmínky, Další pojmy související s Big data problematikou (IoT, Real-time, AI a machine learning) a Ekonomika, business. Zde jde vidět převážně přesah do Smart dopravy a Smart samosprávy. Co se týče Big data analytics a návaznosti na konkrétní oblasti, většinou se zmiňuje v souvislosti s dopravou či využitím v soukromém sektoru. Hamburg pak zmiňuje tuto oblast v souvislosti s monitorováním půdy k identifikaci změn v městském prostoru. Praha zmiňuje analýzu big dat v souvislosti se sociálními médii a digitálními platformami, za účelem zlepšení kvality života a usnadnění každodenních činností. Hong Kong má Big data analytics provázaný na zdravotnictví a oblast infrastruktury, respektive samosprávy. Některé zmínky o big data analytics však nejsou navázané na žádné konkrétní využití, ale slouží jen jako „buzzword“, který města využívají.

Trendem pak pro Smart City je tedy pojem Big data analytics a použití big dat v dopravě, dále pak i věci co se týče řízení města – městské plánování, různé cíle a rozvoj města, použití v rámci policie a například řízení odpadu. Speciální oblastí využití big dat je např. turismus a využití sociálních médií (Praha), řízení zásob vody (Taipei) a výzkum (Tokyo). Zatímco první tři zmíněné (Big data analytics, doprava, řízení města) jsou to hlavní, čemu by se měla města věnovat v souvislosti s big daty, inspirací mohou být i zmíněná speciální využití v rámci Prahy, Taipei a Tokya. Praha se ve své strategii také odkazuje na big data v turismu a řízení toku turistů a návštěvníků v Amsterdamu (Beautiful Noise project) a projekt La Sagrada Familia z Barcelony. Inspirací může být i využití různých sensorů, jak zmiňuje zpráva Top 50 Smart City Governments, v případě Soulu. Ty měří a sbírají data jako je např. hluk, prach, intenzita nočního světla a dopravní provoz, kdy jsou tato data následně analyzována pomocí AI.

AI je také oblast, která se poslední dobou dostává hodně do podvědomí a svoje využití ve Smart Cities v návaznosti na big data určitě má. Stockholm ve své strategii zmiňuje big data a AI v souvislosti s řízením dopravy. Taipei naopak zmiňuje AI a big data analytics v rámci spotřeby elektřiny, samotná big data a AI v rámci „AI smart patrol system“ a identifikace poznávacích značek v rámci policejního oddělení. Dalším použitím big dat a AI v Taipei je v rámci zásoby vody.

Některá doporučení jsou sice ve strategiích vypsány, ale nejsou úplně rozvinuty do detailu, jako tomu je např. u strategie Abu Dhabi. Nepříliš přesná a vesměs obecná byla i zmínka v rámci Ottawy, kde se sice zmiňuje, jak je tato oblast užitečná (převážně v rámci predikce, určení trendů a řešení problému) a je nutné do ní investovat a být v rámci analýzy big dat leaderem, ale není to však provázáno na konkrétní oblast. Jakarta je na tom podobně, kdy ve své strategii zmiňují, že big data a jejich analýza experty pomáhá v procesu rozhodování.

5.2 Problematika ochrany dat v rámci Smart Cities

Ochraně dat nebylo ve strategiích věnováno tolik pozornosti. Určitým vzorem by tak mohla být strategie Singapuru a Stockholmu. Singapurská strategie se věnuje převážně odkazování na další dokumenty, související s bezpečností. Věnuje se tomu, jaké standardy dodržovat, jaké instituce zde působí a jak mohou tyto zmínky pomoci soukromému sektoru. Na druhé straně Stockholm v této oblasti podrobný nebyl, strategie zmínila nařízení EU a potřebu prozkoumání současných zákonů. Strategie zmiňuje i konzistentnost dat, což je v zásadě důležitý požadavek, co se kvality dat týká, kdy v souvislosti s big daty a jejich povahou však tady může docházet k částečnému rozporu, především s ohledem na vlastnosti big dat, které vyžadují jiný přístup k jejich ochraně. Dobrý a doporučený nápad je však zřízení rady pro otázky soukromí a etiky, která by se podílela a byla součástí prioritních projektů. Podobná zmínka je i v Singapurské strategii, kde se zmiňuje, že projekty od začátku spolupracují se Smart Nation Cybersecurity Committee a tento orgán poskytuje v této oblasti i poradenství. Toto souvisí i se závěrem Berlínského senátu, pro který je klíčový, aby nasazení nových technologií bylo v souladu se zákony na ochranu údajů. Důležitým zdrojem by pro města taktéž mohl být dokument *Cyber-en IT-crisisplan: Voorbereid op digitale noodsituaties* (Krizový plán v oblasti kybernetiky a IT) Haagu.

6 OVĚŘENÍ A DISKUSE K DANÉ PROBLEMATICE

Tato kapitola se věnuje ověření doporučení, která byla identifikována na základě analýzy dat získaných a popsanych v předchozích kapitolách, tzn. především kapitolách 4 a 5. Součástí této kapitoly je tedy nejprve popis zvolené metody, její praktická realizace, vyhodnocení zjištění a následná diskuse k nim.

6.1 Ověření pomocí metody Delphi

Ověření identifikovaných doporučení bylo realizováno s využitím metody Delphi. Tato metoda je založena na anonymním více kolovém expertním odhadu odborníků. Při realizaci dílčích kroků je využíváno statistických metod. Obdobně jako u brainstormingu či u expertního panelu se Delphi účastní určitý počet nezávislých expertů. Největší odlišností od obou výše zmíněných metod je anonymní spolupráce expertů (je tak odstraněna psychologická bariéra plynoucí z bezprostřední reakce účastníků v přímém kontaktu). Postoje expertů o možných vývojích jsou upřesňovány v několika kolech pomocí zpětné vazby prostřednictvím poskytnuté informace o ostatních získaných poznacích. Experti jsou seznámeni s postoji ostatních expertů z panelu a musí jim buď přizpůsobit své postoje (s novými daty a argumenty je změnit), nebo vyvrátit pomocí vlastní argumentace expertní názory kolegů (stále je zachována anonymita). Výsledky jsou následně statisticky zpracovány (Štědroň et al., 2012, str. 44-45). V rámci této metody se doporučuje provedení dvou až tří iterací, při dalším nárůstu vzrůstá statistická chyba metody (Smejkal a Rais, 2009, str. 110).

Prvním krokem metody Delphi byla definice prostředí pro ověření. Jelikož byla zvolena online forma této metody s dotazníkem v elektronické podobě, tak museli mít experti k dispozici popis toho, v jaké modelové situaci, včetně možných omezení, pro Smart City mají odhodnocovat předložená doporučení. Jelikož vzorek měst byl vybrán na základě pořadí SCI 2023, tzn. hlavně úspěšná města, která mají dostatek zdrojů, tak i prostředí bylo vymezeno v tomto kontextu. Konkrétně se tedy jedná o modelové Smart City, které má zájem pracovat s big daty a zajistit jejich bezpečnost, tzn., že lze předpokládat, že má k dispozici dostatečné technologické, lidské i finanční zdroje, stejně jako infrastrukturu, kterou pro Smart City projekty může využít.

Dalším krokem bylo nalezení vhodných expertů, když hlavním kritériem pro jejich výběr byla jejich odbornost pro ověřovanou oblast, tzn. Smart Cities, informační systémy a procesy, datové požadavky, a bezpečnost a ochrana dat. Nebylo nutné, aby daný expert měl znalosti ze všech oblastí, ale aby byla vytvořena taková skupina expertů, jejichž znalosti a zkušenosti se budou

doplňovat, což podporuje i zvolená metoda Delphi. Celkem bylo vytipováno a osloveno sedm expertů a všichni s účastí souhlasili. V Tabulka 18 je zaznamenán jejich stručný popis.

Tabulka 18: Seznam expertů a jejich odbornost. Zdroj: *Vlastní zpracování*.

ID	Země	Pracovní pozice	Roky zkušeností	Profesní zkušenosti
1	ČR	Odborný asistent, datový analytik, e-government konzultant	15	Expert má dlouholeté zkušenosti v oblastech e-governmentu, big a open linked dat, bezpečnosti dat, datové analytiky a analýzy dopadů moderních technologií ve veřejném sektoru.
2	Německo	Státní zaměstnanec	12	PhD v informatice, 4 roky zaměstnaný na univerzitě jako výzkumní pracovník – výzkum informačních systémů, 8 let praxe jako systémový analytik a projektant ve veřejném sektoru
3	Rakousko	Profesor, projektový manažer	25	Informatika, informační systémy, data modelling a projektový management zaměřený na modelování a správu dat. Odborník má rozsáhlé zkušenosti s projektovým řízením ve veřejném sektoru.
4	Polsko	Výzkumní pracovník	5	Výzkumní pracovník v oblasti informatiky, informačních systémů a e-governmentu se zaměřením na Smart Cities/Villages/Regions.
5	Slovensko	IT konzultant a datový specialista	20+	Odborník má více než 20 let zkušeností s informačními systémy, webovými aplikacemi a dalšími HW a SW řešeními. Zaměřený na požadavky na data, výměnu dat a posuzování bezpečnosti. Podílel se na projektech soukromého i veřejného sektoru.
6	Švýcarsko	PhD student	3	Smart Cities a e-government
7	ČR	Odborný asistent	25	ICT systémy, počítačové sítě

Dalším krokem bylo vytvoření dotazníku pro první kolo, a to včetně vymezení cílů ověřování a pokynů pro experty. Experti mohli doporučení hodnotit od 0-3 (0 – nehodnoceno, 1 – málo, 2 – mírně, 3 – velmi), tedy jak moc souhlasí s daným doporučením. Dotazník byl sestrojen na základě poznatků zjištěných jak ze studia literatury, tak zejména s pomocí poznatků zjištěných v rámci analýzy strategických dokumentů měst. Doporučení jsou rozdělena do dvou hlavních částí, kdy první část se věnuje problematice big dat a druhá část problematice ochrany dat. Co

se týká části big dat, dotazník zjišťuje, ve kterých oblastech Smart Cities by měla být big data využita, jaké jsou požadavky na big data systémy a jaké ostatní technologické aspekty související s big daty by měly být využity. Ochrana dat se pak zabývá tím, jak by měla města přistupovat k ochraně dat v rámci strategie a jaké aspekty ochrany dat by mohla Smart Cities využít. Detailní přehled doporučení je v následující Tabulka 19. Tento seznam byl použit pro první kolo hodnocení pomocí metody Delphi. Experti měli 14 dní na ohodnocení doporučení a vložení komentářů a návrhů na nová doporučení, případně odstranění existujících doporučení.

Tabulka 19: Úvodní seznam doporučení pro řešenou problematiku. Zdroj: *Vlastní zpracování.*

BIG DATA
Big data by měla být použita v těchto oblastech Smart City...
Smart lidé – oblasti jako sociální interakce, vzdělání, veřejný život atd.
Smart život – oblasti kvality života, jako je kultura, bydlení, turismus atd.
Smart průmysl/ekonomika
Smart síť (elektrina, teplo atd.)
Smart zdravotnictví
Smart doprava
Smart samospráva
Smart prostředí
Smart bezpečnost
Hlavní požadavky pro big data systémy by měla obsahovat...
Škálovatelná a odolná infrastruktura – dokáže pojmout rostoucí objemy dat a zároveň vydrží různé výkyvy v objemu toku dat
Správa dat – získávání, zpracování a spojení z různých zdrojů
Ukládání dat
Výpočetní infrastruktura – jedním z nejpoužívanějších použití oblastí big dat je big data analytics, v některých případech i v reálném čase. Toto musí být bráno na vědomí při tvorbě výpočetní infrastruktury
Ochrana dat
Další technologické aspekty spojené s big daty ve Smart Cities by měla obsahovat...
IoT
AI
Machine Learning
Real-time (analýza dat atd.)
Cloud
Big data analytics

OCHRANA DAT
Smart City strategie by měly...
Vyhnout se zmínce o ochraně dat
Zmínit ochranu dat, ale obecně – bez detailů ohledně implementace
Být transparentní ohledně ochrany dat – ukázat detailní řešení, jak jsou data chráněna
Odkazovat se na ostatní standardy a zákony spojené s ochranou dat – například GDPR
Odkazovat se na další dokumenty, které se věnují ochraně dat v daném městě/zemi
Mít separátní dokument/strategii pro ochranu dat
Ochrana dat ve Smart City by měla obsahovat...
Experta nebo tým expertů pro ochranu dat, která by dohlížela na Smart City projekty
Spolupracovat s komisí zaměřenou na ochranu dat/kyberbezpečnost
Vytvářet či podporovat ochranné známky, které by dělaly firmy ve městě důvěryhodné

Navracené dotazníky z prvního kola byly zpracovány a analyzovány, kdy na základě výpočtů průměrů a směrodatné odchylky byla pro druhé kolo odstraněna tato doporučení „*Vyhnout se zmínce o ochraně dat*“, „*Zmínit ochranu dat, ale obecně – bez detailů ohledně implementace*“ a „*Vytvářet či podporovat ochranné známky, které by dělaly firmy ve městě důvěryhodné*“. Tato doporučení měla mezi experty nízké průměrné skóre na úrovni 1 – málo. Na základě zpětné vazby expertů taktéž byla přidána nová doporučení. AI se tak rozdělila do dvou, a to na *tradiční* a *generativní AI*. Cloud se pak rozdělil na *Cloud computing* a *Cloudové úložiště*. Ve stejné kategorii byly přidány pojmy *Datová jezera* a *NoSQL databáze*. Stejně tak škálovatelná a odolná infrastruktura pro big data systémy byla rozdělena do dvou položek. K rozdělení došlo také u *Úložiště dat*. Dále byla přidána celá jedna kategorie týkající se požadavků na *Big data systémy* od organizace National Institute of Standards and Technology (NIST). To však ale znamenalo i odstranění *Ochrany dat* z hlavních požadavků, jelikož pro toto byla vytvořena nová kategorie. V poslední části se pak přidalo pět nových položek, a to využití nějakého standardu, školení (zaměstnanců, firem a občanů) a vydávání pravidelných zpráv v rámci ochrany dat. Poté byl do úvodu přidán odstavec více osvětlující problematiku a téma. Nakonec byla některá doporučení více rozepsána a upřesněna.

S těmito úpravami bylo realizováno druhé kolo metody Delphi, kdy na vyplnění dotazníků měli experti jeden týden. Navracené dotazníky byly opět zpracovány a řídicí osoba (autor této práce) připravil upravený dotazník pro další kolo, jelikož ani po druhém kole nedošlo mezi experty ke shodě. Pro poslední třetí kolo dotazníku již nebyla přidána žádná nová kategorie. Došlo však k odstranění těch méně významných doporučení, konkrétně těch, které měly průměr pod 2. V případě doporučení od NISTu v kategorii „*Požadavky na poskytovatele služeb*“ došlo

k odstranění doporučení „*Nutnost podporovat starší a pokročilé softwarové balíčky*“ a v sekci „*Požadavky na řízení životního cyklu dat*“ došlo k odstranění „*Nutnost podporovat archivy z více webových stránek*“. V případě ochrany big dat pak došlo k odstranění doporučení „*Vzdělávání (např. školení nebo příručky) firem ve Smart City v otázkách ochrany dat*“.

6.2 Konečný seznam doporučení

V Tabulka 20 je k dispozici závěrečný, potvrzený seznam doporučení, na kterém se experti shodli ve třetím kole metody Delphi. Co se týká souhlasu expertů se seznamem doporučení, v šesti bodové stupnici byl ohodnocen čtyřmi experty jako Agree/Souhlasím (5. stupeň) a třemi experty Agree Strongly/Velmi souhlasím (6. stupeň). Celkem je to tak ohodnocení s průměrem 5,429, přičemž hodnota 6 je nejlepší pozitivní varianta.

V prvním sloupci tabulky je přidáno označení pro přehlednost. Druhý sloupec pak obsahuje doporučení, třetí sloupec průměr, kde je zeleně zvýrazněno 10 nejvyšších hodnot. V posledním sloupci je pak směrodatná odchylka. V oblastech, ve kterých by měla být využívána big data, byla nejlépe hodnocena *Smart doprava* (D 1.1.7), kde všemi experty bylo toto doporučení hodnocené nejvyšší možnou hodnotou. Poté následuje *Smart průmysl/ekonomika* (D 1.1.3) a jako třetí se stejným průměrem *Smart lidé* a *Smart život* (D 1.1.1 a D 1.1.2). V hlavních požadavcích pro big data je pak nejvýše hodnoceno *Úložiště dat pro citlivá data* (D 1.2.4), poté *Škálovatelná infrastruktura* (D 1.2.1) a *Výpočetní infrastruktura* (D 1.2.6).

Co se týče ochrany dat, tak v sekci „*Smart City strategie by měly...*“ je nejlépe hodnoceným doporučením *Být transparentní ohledně ochrany dat* (D 2.1.1) a *Mít samostatný dokument (strategii) pro ochranu dat* (D 2.1.4). Za nimi pak je ještě doporučení D 2.1.2, tedy *Odkazovat se na ostatní standardy a zákony spojené s ochranou dat*. V poslední kategorii, „*Ochrana dat ve Smart City by měla obsahovat...*“, je pak nejlépe hodnocené doporučení D 2.2.1 – *Experta nebo tým expertů pro ochranu dat, kteří by dohlíželi na Smart City projekty* s průměrem 3 a *Vzdělávání (např. školení nebo příručky) zaměstnanců Smart City v otázkách ochrany dat* (D 2.2.4).

Co se týká jednotlivých kategorií, experti hodnotili oblast ochrany dat jako důležitější než oblast big dat. V rámci dílčích kategorií byla v oblasti big dat nejlépe hodnocena kategorie D 1.2, tedy *Hlavní požadavky pro big data systémy by měly obsahovat*. V rámci ochrany dat se pak s průměrem 3, tzn., že se všichni experti shodli na nejvyšším možném hodnocení, jako klíčová zdá kategorie D 2.1 – *Smart City strategie by měly*, která se zabývá otázkou, jak by měla Smart City strategie přistupovat k problematice ochrany dat.

Tabulka 20: Konečný seznam doporučení pro řešenou problematiku. Zdroj: *Vlastní zpracování.*

Číslování	Doporučení	Průměr	Směrodatná odchylka
D 1	BIG DATA	2,286	0,452
D 1.1	Big data by měla být využita v těchto oblastech Smart City...	2,286	0,452
D 1.1.1	Smart lidé – oblasti jako sociální interakce, vzdělání, veřejný život atd.	2,714	0,452
D 1.1.2	Smart život – oblasti kvality života, jako je kultura, bydlení, turismus atd.	2,714	0,452
D 1.1.3	Smart průmysl/ekonomika – podpora businessu a příprava podmínek pro růst (např. tvorba platformy/sdílení dat, tvorba/podpora infrastruktury, optické kabely, 5G atd.)	2,857	0,350
D 1.1.4	Smart síť – oblasti jako je přívod elektřiny, tepla, vody atd.	2,429	0,495
D 1.1.5	Smart zdravotnictví – podpora funkčnosti jak nemocnic, tak propojení/podpora zdravotních služeb soukromého sektoru	2,429	0,495
D 1.1.6	Smart mobilita – zahrnuje veřejnou dopravu a sdílení dopravních služeb (auta, kola atd.)	2,429	0,495
D 1.1.7	Smart doprava – řízení dopravy (semaforey, přechody pro chodce atd.), smart parkování a sdílení informací o dopravě	3,000	0,000
D 1.1.8	Smart samospráva – oblasti spojené s řízením města, jako je např. územní plánování, vyřizování občanských záležitostí, volby, sdílení informací atd.	2,143	0,350
D 1.1.9	Smart prostředí – přírodní podmínky (klíma, zeleň), znečištění, řízení přírodních zdrojů	2,286	0,452
D 1.1.10	Smart bezpečnost – oblasti, které dělají město více bezpečné, jako je například podpora kamer, chytrá policie, mapy kriminality atd.	2,571	0,495
D 1.2	Hlavní požadavky pro big data systémy by měly obsahovat...	2,714	0,452
D 1.2.1	Škálovatelná infrastruktura – dokáže pojmout rostoucí objemy dat	2,714	0,452
D 1.2.2	Odolná infrastruktura – vydrží různé výkyvy v objemu toku dat	2,571	0,495
D 1.2.3	Správa dat – podporuje získávání, zpracování a propojování dat z různých zdrojů	2,571	0,495
D 1.2.4	Úložiště dat pro citlivé data – speciální zabezpečené místo pro citlivá data	2,857	0,350
D 1.2.5	Úložiště dat pro ostatní big data – platforma pro ostatní typy big dat	2,143	0,350
D 1.2.6	Výpočetní infrastruktura – jedním z nejčastějších využití oblastí big dat je big data analytics, v některých případech i v reálném čase, což musí být bráno na vědomí při budování výpočetní infrastruktury	2,714	0,452

Číslování	Doporučení	Průměr	Směrodatná odchylka
D 1.3	Smart Cities by měla zahrnout tyto NIST Big data obecné požadavky...	2,143	0,350
D 1.3.1	Požadavky na poskytovatele převodů a zpracování	2,143	0,350
D 1.3.1.1	Musí podporovat diverzifikované výpočetně náročné, statistické a analytické zpracování a techniky machine learning	2,143	0,350
D 1.3.1.2	Musí podporovat dávkové analytické zpracování a analytické zpracování v reálném čase	2,286	0,452
D 1.3.1.3	Musí podporovat zpracování rozsáhlého a různorodého obsahu dat a modelování	2,143	0,350
D 1.3.1.4	Musí podporovat zpracování dat v pohybu (streamování, načítání nového obsahu, sledování atd.)	2,143	0,350
D 1.3.2	Požadavky na poskytovatele služeb	2,143	0,350
D 1.3.2.1	Musí podporovat starší a pokročilé výpočetní platformy	2,000	0,000
D 1.3.2.2	Musí podporovat starší a pokročilé distribuované výpočetní clustery, koprocory, vstupně-výstupní zpracování (infrastruktura)	2,000	0,535
D 1.3.2.3	Musí podporovat pružný přenos dat (sítě)	2,143	0,350
D 1.3.2.4	Musí podporovat starší, rozsáhlé a pokročilé distribuované ukládání dat (úložiště)	2,286	0,452
D 1.3.2.5	Musí podporovat starší a pokročilé spustitelné programování: aplikace, nástroje, utility a knihovny (SW)	2,000	0,535
D 1.3.3	Požadavky na zabezpečení a ochranu osobních údajů	2,429	0,495
D 1.3.3.1	Musí poskytovat ochranu a zachovávat bezpečnost a soukromí citlivých údajů	2,571	0,495
D 1.3.3.2	Musí podporovat sandbox, řízení přístupu a víceúrovňovou autentizaci řízenou zásadami u chráněných dat	2,286	0,700
D 1.3.4	Požadavky na řízení životního cyklu dat	2,286	0,452
D 1.3.4.1	Musí podporovat správu kvality dat včetně předzpracování, shlukování dat, klasifikace a transformace formátů	2,143	0,350
D 1.3.4.2	Musí podporovat dynamické aktualizace dat, uživatelských profilů a odkazů	2,143	0,639
D 1.3.4.3	Musí podporovat životní cyklus dat a politiku dlouhodobého uchovávání, včetně původu dat	2,286	0,452
D 1.3.4.4	Musí podporovat ověřování dat	2,286	0,452
D 1.3.4.5	Musí podporovat lidský faktor pro ověření dat	2,000	0,000
D 1.3.4.6	Musí podporovat prevenci ztráty nebo poškození dat	2,429	0,495
D 1.3.4.7	Musí podporovat trvalý identifikátor a sledovatelnost dat	2,286	0,452
D 1.3.4.8	Musí podporovat standardizaci, agregaci a normalizaci dat z různých zdrojů	2,143	0,350

Číslování	Doporučení	Průměr	Směrodatná odchylka
D 1.4	Další technologické aspekty spojené s big daty ve Smart Cities by měla obsahovat...	2,286	0,452
D 1.4.1	IoT	2,571	0,495
D 1.4.2	Tradiční AI	2,429	0,495
D 1.4.3	Generativní AI	2,286	0,452
D 1.4.4	Machine Learning	2,286	0,452
D 1.4.5	Real-time (analysis etc.)	2,143	0,350
D 1.4.6	Cloudová úložiště – pouze ty formy nasazení cloudu, které splňují bezpečnostní požadavky veřejného sektoru	2,000	0,000
D 1.4.7	Cloud computing – pouze ty formy nasazení cloudu, které splňují bezpečnostní požadavky veřejného sektoru	2,143	0,639
D 1.4.8	Big data analytics	2,286	0,452
D 1.4.9	Datová jezera	2,000	0,535
D 1.4.10	NoSQL databáze	2,143	0,639
D 2	OCHRANA DAT (všechna data ve Smart Cities)	2,857	0,350
D 2.1	Smart City strategie by měly...	3,000	0,000
D 2.1.1	Být transparentní ohledně ochrany dat – ukázat detailní řešení, jak jsou data chráněna	2,857	0,350
D 2.1.2	Odkazovat se na ostatní standardy a zákony spojené s ochranou dat – například GDPR	2,714	0,452
D 2.1.3	Odkazovat se na další dokumenty, které se věnují ochraně dat v daném městě/zemi	2,286	0,452
D 2.1.4	Mít samostatný dokument (strategii) pro ochranu dat	2,857	0,350
D 2.2	Ochrana dat ve Smart City by měla obsahovat...	2,571	0,495
D 2.2.1	Experta nebo tým expertů pro ochranu dat, kteří by dohlíželi na Smart City projekty	3,000	0,000
D 2.2.2	Spolupracovat s komisí zaměřenou na ochranu dat/kyberbezpečnost	2,286	0,452
D 2.2.3	Využití nějakého existujícího standardu nebo frameworku (např. NIST).	2,429	0,495
D 2.2.4	Vzdělávání (např. školení nebo příručky) zaměstnanců Smart City v otázkách ochrany dat	2,857	0,350
D 2.2.5	Vzdělávání (např. školení nebo příručky) občanů Smart City v otázkách ochrany dat	2,286	0,452
D 2.2.6	Vydávání pravidelných kontrolních zpráv (výročních atd.) v oblasti ochrany dat	2,000	0,535

6.3 Diskuse

Na základě výsledků ověření seznamu doporučení lze tvrdit, že vybraní experti se zaměřovali spíše na problematiku bezpečnosti big dat než samotná big data, když lze předpokládat, že této problematice zřejmě rozumí více a je pro ně bližší. U oblasti big dat se zaměřovali na konkrétní komponenty Smart City, tzn. oblasti aplikace chytrých služeb jako Smart doprava, kde si lze využívání big dat v praxi zřejmě lépe představit než u spíše technologicky pojatých doporučení. Nicméně i tato zjištění mají svůj přínos, protože se lze v dalším výzkumu zaměřit na konkrétní služby, která z big dat mohou čerpat, a až v dalším kroku řešit různé technologické aspekty. Při konkrétním využitím ve formě služeb lze vyjít např. z prací autorů Al Nuaimi et al. (2015), Hashem (2016), Lim et al. (2018) nebo Wu et al. (2018).

Dále lze diskutovat to, že by bylo vhodné získat více vstupních dat, což ale zabere i více času, který však nebyl v rámci této diplomové práce k dispozici. Města byla vybrána na základě SCI z roku 2023, kdy bylo vybráno 40 měst z celkového počtu 141. K analýze by se tedy mohla vzít všechna města z tohoto indexu a dohledat zbylých 101 Smart City strategií, pokud ty existují, analyzovat tyto dokumenty a provést analýzu na tomto zvětšeném datovém souboru. Větší vzorek lze uvažovat i u počtu expertů, kteří doporučení hodnotili, ale i zde by byl problémem čas a případné další náklady spojené se zapojením více expertů.

Dále zde lze uvést kroky zpracování a analýzy dat v této práci, když se např. mohly vyskytnout chyby v průběhu zápisu údajů do tabulek. V některých dostupných dokumentech totiž někdy bylo složité odvodit daný údaj (především co se týká roku vydání a roku platnosti strategie). U některých měst bylo také na dané téma k dispozici několik dokumentů, kdy je složité určit který z nich by se dal považovat za oficiální strategii, na kterou by se mělo odkazovat. Během zpracování dat v Pythonu pomocí shlukové analýzy byl proveden pokus o identifikaci vzorů, kdy se daly v jednotlivých shlucích najít určité vzory, převážně co se využívání big dat týká. Z důvodu rozsahu diplomové práce ale nebyla tato část více rozebrána, když se tomuto tématu mohou věnovat některé budoucí práce.

Posledním bodem diskuse je přístup k úpravám a změnám v dotaznících pro jednotlivá kola na základě zpětné vazby od expertů. Zde je nutné zmínit, že v dotaznících byly některými experty zmíněny i další poznatky, které byly příliš specifické anebo se nehodily do tématu či struktury dotazníku. Jeden z expertů zmiňuje např. potřeby týkající se organizačních a finančních aspektů a také provedení analýzy současného stavu a připravenosti Smart City, což se úplně netýká problematiky big dat a ochrany dat jako takové. Nicméně je to jeden z kroků, který by měl být

v rámci Smart Cities realizován. Jako příklad expert uvádí studii, zda město má dostatečné finanční a lidské zdroje na takovéto projekty, zda projekty v rámci big dat jsou podporované zastupiteli a vedením města atd. Stejný expert dále zmiňoval také doporučení týkající se financování a doporučení zabývající se zúčastněnými stranami (stakeholdery), jejich potřebami, dovednostmi a kompetencemi, což je opět jeden z kroků, kterým by města měla věnovat pozornost. Další expert zmiňoval potřebu provedení počáteční analýzy před tvorbou strategie, ve které by města v rámci projektů hodnotila: 1) jde-li pouze o návrh ke zvážení, 2) jde-li o konkrétní návrh, který obsahuje podrobné objasnění požadavků, 3) zda je daný projekt již implementován, ale měl by být vylepšen, aby mohl čelit novým výzvám vyplývajícím z big dat a ochrany dat a 4) zda projekt využívá již existující infrastrukturu, HW a SW prostředky. S tím souvisí i zhodnocení implementovaných řešení pomocí různých metod, což je téma, které bylo ve strategiích často opomíjeno.

Pomocí metody Delphi byl tedy vytvořen a experty ověřen seznam doporučení. Nicméně je nutné zmínit, že tato doporučení slouží převážně větším městům, jelikož analýza získaných dat vycházela hlavně z poznatků větších měst umístěných v různých žebříčcích, ačkoliv by se některé z nich daly použít i na města menší. Toto vše je tedy nutné vzít v úvahu při praktické aplikovatelnosti doporučení ve Smart City projektech, případně jejich použití ve strategiích.

ZÁVĚR

Cílem této diplomové práce byla analýza problematiky big dat, práce s těmito daty a jejich ochrana ve strategiích a projektech Smart Cities ve vybraných světových městech, která pomocí zvolených metodických kroků vedla k vytvoření modelů a doporučení pro tuto problematiku.

Cíl práce byl splněn, když v přechozích kapitolách jsou prezentovány jak modely vytvořené v Pythonu, tak ověřený seznam doporučení. Pro práci bylo vybráno 40 měst na základě zprávy SCI 2023, z nichž pouze 30 mělo svoji vlastní strategii. Dvě z nich však nebyly z ČR dostupné, proto se v analýze pracovalo pouze s 28 strategiemi nesoucí. U nich se zjistilo, že 68 % z nich, tedy 19 strategií, mělo ve své strategii pro tuto problematiku název „Smart“. Mimoto strategie pracují převážně se slovem „Digitální“. Z vybraných strategií se pak polovina měst zmiňovala v nějakém smyslu o big datech, která byla nejčastěji zmíněna ve spojitosti s big data analytics a dopravou. Při tvorbě modelů v Pythonu pak byla klíčovou částí shluková analýza, která se na základě použití big dat snažila města shlukovat. Po detailním prozkoumání měst v jednotlivých shlucích se našly určité společné znaky, výsledky však nebyly na tomto vzorku měst průkazné.

Dále byl na základě poznatků zjištěných jak v rámci analýzy literatury, tak na základě analýzy strategií, vytvořen seznam doporučení. Ten byl pomocí tří kol metody Delphi ověřen a schválen sedmi experty. Tato doporučení však slouží převážně větším městům, jelikož doporučení byla vytvořena na základě analýzy měst umístěných na předních příčkách pořadí věnujících se problematice Smart Cities. Některá doporučení však mohou najít využití i v městech menších.

V závěru práce pak byla k dané problematice provedena diskuse. Zde se nachází jak vymezení možných nedostatků v rámci práce, převážně co se kvality dat týká, tak vytyčení bodů, kde by se dala práce rozvinout, ale z důvodu rozsahu diplomové práce tomu už nemohl být věnován prostor. Také se zde nachází i další poznatky zjištěné od expertů v rámci metody Delphi, které nenašly uplatnění v navrženém seznamu doporučení, avšak je důležité tyto informace zmínit.

Závěrem lze říci, že problematika big dat byla sice řešena v polovině vybraných strategií, ale není v nich řešena v takovém rozsahu či detailu, v jakém by to bylo možné. Co se týká problematiky ochrany dat, výsledky byly ještě horší. Ochrana dat se neřešila ve strategiích skoro vůbec a když už, tak to byly jenom krátké zmínky. Zde je proto třeba vyzdvihnout strategie Singapuru a Stockholmu, které se problematice ochrany dat věnovaly nejvíce.

Tato diplomová práce má jak významné praktické přínosy ve formě ověřených doporučení pro strategie a projekty Smart Cities, tak přínosy spočívající v propojení pojmů big data, ochrana dat, strategie, projekty a Smart City a společném pohledu na ně.

POUŽITÁ LITERATURA

- AL NUAIMI, Eiman, AL NEYADI, Hind, MOHAMED, Nader a AL-JAROODI, Jameela. Applications of big data to Smart Cities. *Journal of Internet Services and Applications*, 2015, 6.1: 1-15.
- ANGELIDOU, Margarita. Smart City policies: A spatial approach. *Cities*, 2014, 41: S3-S11.
- AROCKIA, Panimalar, VARNEKHA Shree. S a VENESHIA, Kathrine. A. The 17 V's of Big Data. *International Research Journal of Engineering and Technology*, 2017, 4.9: 329-333.
- AZURE. Databáze NoSQL – co je NoSQL? MICROSOFT. *Microsoft Azure* [online]. c2024. Dostupné z: <https://azure.microsoft.com/cs-cz/resources/cloud-computing-dictionary/what-is-nosql-database>
- BARLOW, Mike a LEVY-BENCHETON, Cornelia. *Smart Cities, Smart Future: Showcasing Tomorrow*. Wiley, 2018. ISBN 9781119516217.
- BAWANY, Narmeen Zakaria a SHAMSI, Jawwad A. Smart City architecture: Vision and challenges. *International Journal of Advanced Computer Science and Applications*, 2015, 6.11: 1-11.
- BERISHA, Blend, MĚZIUI, Endrit a SHABANI, Isak. Big data analytics in Cloud computing: an overview. *Journal of Cloud Computing*, 2022, 11.1: 24.
- BERRONE, Pascual a RICART, Joan Enric. IESE Cities in Motion Index. *Business School - University of Navarra* [online]. 2022. Dostupné z: <https://www.iese.edu/media/research/pdfs/ST-0633-E.pdf>
- BOTELHO, Bridget a BIGELOW, Stephen J. TECHTARGET. *What is big data?* [online]. 2022. Dostupné z: <https://www.techtargget.com/searchdatamanagement/definition/big-data>
- COLLIS, David J. a RUKSTAD, Michael G. Can you say what your strategy is?. *Harvard Business Review*, 2008, 86.4: 82-90.
- ČESKO. Zákon č. 110/2019 Sb., o zpracování osobních údajů. In: *Sbírka zákonů*. 2019, částka 47, číslo 110, s. 890-963.
- ČESKO. Zákon č. 181/2014 Sb., o kybernetické bezpečnosti a o změně souvisejících zákonů (zákon o kybernetické bezpečnosti). In: *Sbírka zákonů*. 2014, částka 181, s. 1-29.
- DODGE, Martin a Rob KITCHIN. *Mapping Cyberspace*. Taylor & Francis, 2003. ISBN 9781134639007.

- DOLEŽAL, Jan a KRÁTKÝ, Jiří. *Projektový management v praxi: Naučte se řídit projekty!*. Grada, 2016. ISBN 9788024756936.
- CUZZOCREA, Alfredo. Big data lakes: models, frameworks, and techniques. In: *2021 IEEE International Conference on Big Data and Smart Computing (BigComp)*. IEEE, 2021, p. 1-4.
- EDEN STRATEGY INSTITUTE. *Top 50 Smart City Governments* [online]. 2021. Dostupné z:
https://static1.squarespace.com/static/5b3c517fec4eb767a04e73ff/t/6063814af4d39b693379597d/1617133979623/Eden_Top+50+Smart+City+Governments+2020-21_DIGITAL.pdf
- FAVARETTO, Maddalena, DE CLERCQ, Eva, SCHNEBLE, Christophe Olivier a ELGER, Bernice Simone. What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PloS one*, 2020, 15.2: e0228987.
- FISCHER, Eric A. *Cybersecurity issues and challenges: In brief* [online]. Congressional Research Service, 2014. Dostupné z: <https://a51.nl/sites/default/files/pdf/R43831.pdf>
- FREEDMAN, Lawrence. *Strategy: A history*. Oxford University Press, 2015.
- GIFFINGER, Rudolf, FERTNER, Christian, KRAMAR, Hans, KALASEK, Robert, PICHLER-MILANOVIĆ, Nataša a MEIJERS, Evert. *Smart Cities. Ranking of European medium-sized cities. Final Report*. Centre of Regional Science (SRF), Vienna University of Technology, 2007.
- IKEGWU, Anayo Chukwu, NWEKE, Henry Friday, ANIKWE Chioma Virginia, ALO, Uzoma Rita a OKONKWO, Obikwelu Raphael. Big data analytics for data-driven industry: a review of data sources, tools, challenges, solutions, and research directions. *Cluster Computing*, 2022, 25.5: 3343-3387.
- HALEGOUA, Germaine. *Smart Cities*. MIT Press, 2020.
- HARRISON, Colin a DONNELLY, Ian Abbott. A theory of Smart Cities. In: *Proceedings of the 55th Annual Meeting of the ISSS-2011*. 2011, 55.1, p. 1-15.
- HASHEM, Ibrahim Abaker Targio, et al. The role of big data in Smart City. *International Journal of Information Management*, 2016, 36.5: 748-758.
- HENDL, Jan. *Big data: věda o datech – základy a aplikace*. Praha: Grada Publishing, 2021. Průvodce. ISBN 978-80-271-3031-3.

- HOLUBOVÁ, Irena, KOSEK, Jiří, MINAŘÍK, Karel a NOVÁK, David. *Big Data a NoSQL databáze*. Profesional. Praha: Grada, 2015. ISBN 978-80-247-5466-6.
- HUSSIEN, Abou_el_ela Abdou. Fifty-Six Big Data V's Characteristics and Proposed Strategies to Overcome Security and Privacy Challenges (BD2). *Journal of Information Security*, 2020, 11: 304-328.
- HOWARTH, Josh. EXPLODING TOPICS. *5 Key Big Data Trends (2024 & 2025)* [online]. 2023. Dostupné z: <https://explodingtopics.com/blog/big-data-trends>
- IMD BUSINESS SCHOOL. *IMD Smart City Index Report 2023* [online]. 2023. Dostupné z: <https://www.imd.org/wp-content/uploads/2023/04/smartcityindex-2023-v7.pdf>
- INNOVATURE BPO. *Top 10 Essential Big Data Trends for 2023* [online]. 2023. Dostupné z: <https://innovatureinc.com/top-10-essential-big-data-trends/>
- ISARSOFT GMBH. *Smart City Strategy: The Foundation of Every Smart City* [online]. 2023. Dostupné z: <https://www.isarsoft.com/article/smart-city-strategy-the-foundation-of-every-smart-city>
- KEARNEY. *The distributed geography of opportunity - 2023 Global Cities Report* [online]. 2023. Dostupné z: <https://www.kearney.com/documents/291362523/299003325/The+distributed+geography+of+opportunity-the+2023+Global+Cities+Report.pdf>
- KHAN, Zaheer, ANJUM, Ashiq a KIANI, Saad Liaquat. Cloud based big data analytics for smart future cities. In: *2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing*. IEEE, 2013. p. 381-386.
- KITCHIN, Rob. *Getting smarter about Smart Cities: Improving data privacy and data security*. Data Protection Unit, Department of the Taoiseach, Dublin, Ireland, 2016.
- KITCHIN, Rob. The real-time city? Big data and smart urbanism. *GeoJournal*, 2014, 79: 1-14.
- LAI, Calvin Ming Tsun a COLE, Alistair. Measuring progress of Smart Cities: Indexing the Smart City indices. *Urban Governance*, 2023, 3.1: 45-57.
- LIM, Chiehyeon, KIM, Kwang-Jae a MAGLIO, Paul P. Smart Cities with big data: Reference models, challenges, and considerations. *Cities*, 2018, 82: 86-99.
- LISDORF, Anders. *Demystifying Smart Cities: Practical Perspectives on How Cities Can Leverage the Potential of New Technologies*. Apress, 2019. ISBN 9781484253779.

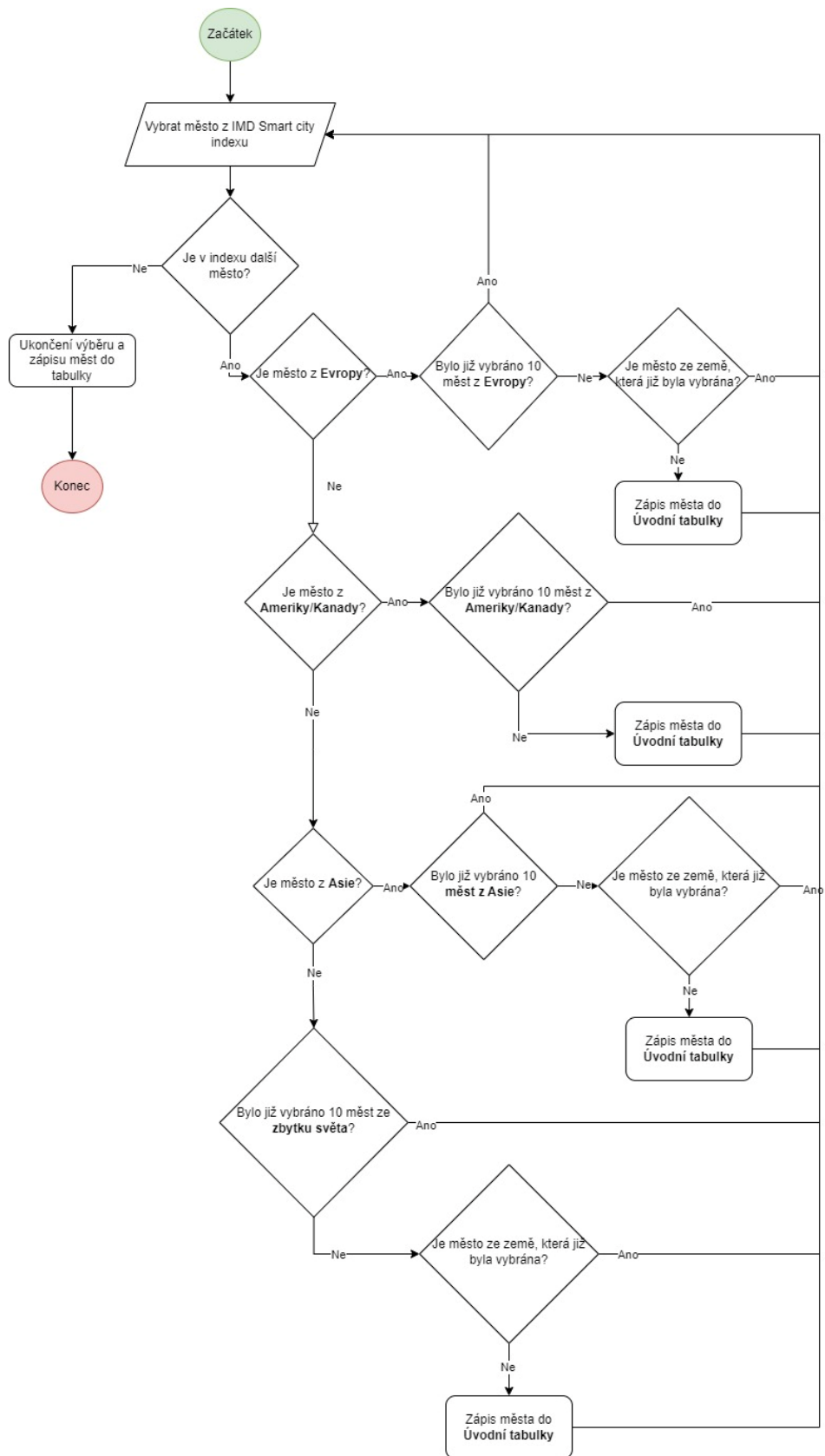
- LI, DeRen, CAO, JianJun a YAO, Yuan. Big data in Smart Cities. *Science China. Information Sciences*, 2015, 58.10: 1-12.
- MCGHEE, Pamela a MCALINEY, Peter. *Painless Project Management: A Step-by-Step Guide for Planning, Executing, and Managing Projects*. Wiley, 2007. ISBN 9780470175187.
- MICROSOFT LEARN. Relační data vs. data NoSQL. MICROSOFT. *Microsoft Learn* [online]. 2023. Dostupné z: <https://learn.microsoft.com/cs-cz/dotnet/architecture/cloud-native/relational-vs-nosql-data>
- MILOSLAVSKAYA, Natalia a TOLSTOY, Alexander. Big data, fast data and data lake concepts. *Procedia Computer Science*, 2016, 88: 300-305.
- MINISTERSTVO PRO MÍSTNÍ ROZVOJ ČR. Metodika Smart Cities. *Ministerstvo pro místní rozvoj ČR* [online]. 2018. Dostupné z: https://mmr.gov.cz/getmedia/f76636e0-88ad-40f9-8e27-cbb774ea7caf/metodika_smart_cities.pdf.aspx?ext=.pdf
- MOORE, Matt. The 7 V's of Big Data. IMPACT TECH, INC. *Impact* [online]. c2023. Dostupné z: <https://impact.com/marketing-intelligence/7-vs-big-data/>
- NEVIMA, Jan, MAJEROVÁ, Ingrid, LEBIEDZIK, Marian, TUREČKOVÁ, Kamila, VAŇOVÁ, Anna et al. *SMART technologie pro zvyšování kvality života v městech a regionech*. [Průhonice]: Professional Publishing, 2023. ISBN 978-80-88260-62-2.
- OZKAYA, Erdal. *Cybersecurity: The Beginner's Guide: A Comprehensive Guide to Getting Started in Cybersecurity*. Erdal Ozkaya, 2019. ISBN 9781789806939.
- PELTON, Joseph N. a SINGH, Indu B. *Smart Cities of today and tomorrow: Better technology, infrastructure and security*. Cham: Springer International Publishing, 2018. ISBN 978-3-319-95821-7.
- POTANSKI, Tom. DEVS DATA. *Big Data Terms Every Manager Should Know* [online]. 2023. Dostupné z: <https://devsdata.com/big-data-terms-every-manager-should-know/>
- PRAKASH, Kolla Bhanu, NAYAK, Janmenjoy, MADHAV, B. T. P., PADMANABAN, Sanjeevikumar a BALAS, Valentina Emilia (Eds.). *Big data analytics and intelligent techniques for Smart Cities*. Boca Raton: CRC Press, 2021. ISBN 978-0367753559.
- ROBINSON, Scott. TECHTARGET. *What is big data analytics?* [online]. 2023. Dostupné z: <https://www.techtargget.com/searchbusinessanalytics/definition/big-data-analytics>

- ROSATI, Umberto a CONTI, Sergio. What is a smart city project? An urban model or a corporate business plan?. *Procedia-Social and Behavioral Sciences*, 2016, 223: 968-973.
- ROSS, Ron, Victoria PILLITTERI, Richard GRAUBART, Deborah BODEAU a MCQUAID, Rosalie. Developing Cyber-Resilient Systems: A Systems Security Engineering Approach. *NIST Special Publication 800-160* [online]. NIST - National Institute of Standards and Technology, 2021, 310. Dostupné z: <https://csrc.nist.gov/pubs/sp/800/160/v2/r1/final>
- SHARIFI, Ayyoob. A critical review of selected Smart City assessment tools and indicator sets. *Journal of Cleaner Production*, 2019, 233: 1269-1283.
- SLAVÍK, Jakub. *Smart City v praxi: jak pomocí moderních technologií vytvářet město příjemné k životu a přátelské k podnikání*. Praha: Profi Press, 2017. ISBN 978-80-86726-80-9.
- SMART NATION AND DIGITAL GOVERNMENT OFFICE. *Smart Nation: The Way Forward – Executive Summary*. Smart Nation and Digital Government Office, 2018.
- SMEJKAL, Vladimír a Karel RAIS. *Řízení rizik ve firmách a jiných organizacích*. 3., rozšířené a aktualizované vydání. Grada Publishing, 2009. ISBN 9788024770055.
- STOCKHOLMS STAD. *Strategi för Stockholm som smart och uppkopplad stad: Bilaga 1. Strategi*. Stockholms stad, 2021.
- SUNYAEV, Ali. Cloud computing. In: *Internet Computing*. Springer, Cham, 2020, p. 195-236.
- ŠTĚDRONĚ, Bohumír, Martin POTŮČEK, Jaroslav KNÁPEK a Petr MAZOUCH. *Prognostické metody a jejich aplikace*. C.H. Beck, 2012. ISBN 9788071791744.
- TORRIE, Ralph a MORSON, Nadia. *SUSTAINABLE CITIES INDEX REPORT*. CORPORATE KNIGHTS, 2023. Dostupné z: <https://www.corporateknights.com/wp-content/uploads/2023/04/2023-Sustainable-Cities-Index-Report.pdf>
- WU, Shiann Ming, CHEN, Tsung-chun, WU, Yenchun Jim a LYTRAS, Miltiadis. Smart Cities in Taiwan: A perspective on big data applications. *Sustainability*, 2018, 10.1: 106.
- WU, Yenchun Jim a CHEN, Jeng-Chung. A structured method for smart city project selection. *International Journal of Information Management*, 2021, 56: 101981.
- YANG, Chaowei, HUANG, Qunying, LI, Zhenlong, LIU, Kai a HU, Fei. Big Data and cloud computing: innovation opportunities and challenges. *International Journal of Digital Earth*, 2017, 10.1: 13-53.

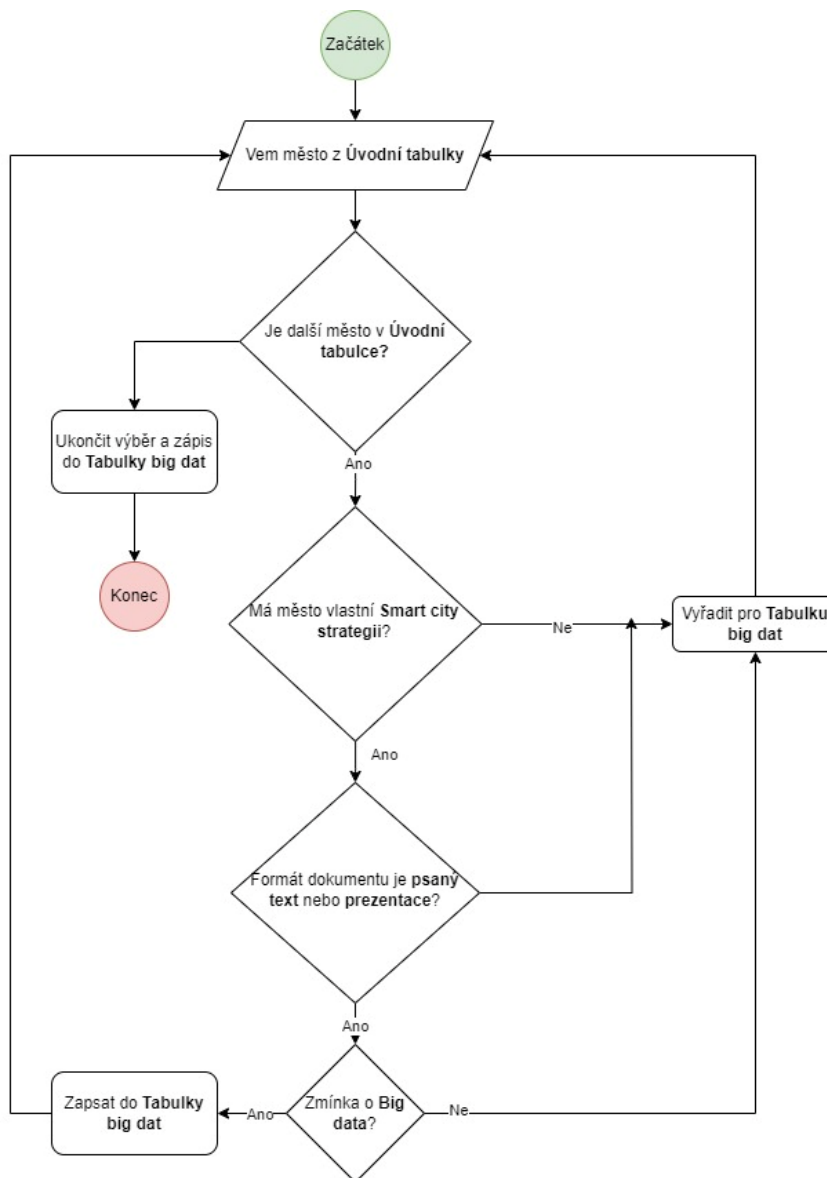
SEZNAM PŘÍLOH

Příloha I – Způsob výběru do Úvodní tabulky.....	I
Příloha II – Způsob výběru měst do tabulky Big dat	II
Příloha III – Python (chybějící hodnoty).....	II
Příloha IV – Python (popisná statistika).....	III
Příloha V – Python (korelace).....	III
Příloha VI – Python (shluková analýza).....	IV

PŘÍLOHA I – ZPŮSOB VÝBĚRU DO ÚVODNÍ TABULKY



PŘÍLOHA II – ZPŮSOB VÝBĚRU MĚST DO TABULKY BIG DAT



PŘÍLOHA III – PYTHON (CHYBĚJÍCÍ HODNOTY)

```
1 import pandas as pd
2
3 data = pd.read_csv("PythonTabulka1C.csv", sep=';')
4 chybejiciHodnoty = data.isnull().sum()
5 print(chybejiciHodnoty)
6
7 data['Rok vydání strategického dokumentu'].fillna(int(data['Rok vydání strategického dokumentu'].mean()), inplace=True)
8 data['Počet zmínek big data'].fillna(int(0), inplace=True)
9
10 data.to_csv("PythonTabulka1CC.csv", sep=';', index=True, encoding="utf-8-sig")
```

PŘÍLOHA IV – PYTHON (POPISNÁ STATISTIKA)

```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 #Načtení souborů
5 data1 = pd.read_csv("PythonTabulka100.csv", sep=';')
6 data2 = pd.read_csv("PythonTabulka20.csv", sep=';')
7
8 #Popisná statistika
9 ds1 = data1.describe()
10 print(ds1)
11 df = pd.DataFrame(ds1)
12 df.to_csv("PopisnaStatistika.csv", sep=';', index=True)
13
14 ds2 = data2.describe()
15 print(ds2)
16 df = pd.DataFrame(ds2)
17 df.to_csv("PopisnaStatistika2.csv", sep=';', index=True)
18
19 data1.info()
20 data2.info()
21
22 ds2 = data2["Počet zmínek big data"].describe()
23 print(ds2)
24
25 #Grafy
26 plt.boxplot(data1["Počet zmínek big data"])
27 plt.title("Rozložení počtu zmínek big dat")
28 plt.show()
```

PŘÍLOHA V – PYTHON (KORELACE)

```
1 import pandas as pd
2 from sklearn import preprocessing
3
4 #Načtení souborů
5 data1 = pd.read_csv("PythonTabulka100.csv", sep=';')
6 data2 = pd.read_csv("PythonTabulka20.csv", sep=';')
7
8 #Transformace hodnot
9 data1["Smart city dokument"] = data1["Smart city dokument"].replace("Psaný text", 0).replace("Prezentace", 1)
10 data2["Smart city dokument"] = data2["Smart city dokument"].replace("Psaný text", 0).replace("Prezentace", 1)
11 data1["Má big data?"] = data1["Má big data?"].replace("Ne", 0).replace("Ano", 1)
12 data1["Region"] = data1["Region"].replace("Evropa", 0).replace("Asie", 1).replace("Zbytek světa", 2).replace("Severní Amerika", 3)
13 data2["Region"] = data2["Region"].replace("Evropa", 0).replace("Asie", 1).replace("Zbytek světa", 2).replace("Severní Amerika", 3)
14
15 #Výběr sloupců
16 vyber1 = data1[["Pořadí v indexu", "Region", "Číslo HDI", "Počet obyvatel (2023)", "Country HDI",
17 "Smart city dokument", "Rok vydání strategického dokumentu", "Počet stran", "Má big data?"]]
18 vyber2 = data2[["Pořadí v indexu", "Region", "Smart city dokument", "Počet zmínek big data", "Doprava",
19 "Big data analytics", "Věci spojené s řízením města", "Ekonomika, business", "Technologické zmínky", "Další pojmy související s Big data problematikou"]]
20
21 #Standardizace a normalizace
22 standard1 = preprocessing.scale(vyber1)
23 standard2 = preprocessing.scale(vyber2)
24 normal1 = preprocessing.normalize(standard1)
25 normal2 = preprocessing.normalize(standard2)
26
27 #Korelace
28 df1 = pd.DataFrame(normal1).corr()
29 df1.to_csv("Korelace1.csv", sep=';', index=True)
30 df2 = pd.DataFrame(normal2).corr()
31 df2.to_csv("Korelace2.csv", sep=';', index=True)
```

PŘÍLOHA VI – PYTHON (SHLUKOVÁ ANALÝZA)

```
1 import numpy as np
2 import pandas as pd
3 from matplotlib import pyplot as plt
4 from sklearn.cluster import KMeans
5 from sklearn import preprocessing
6
7 #Načtení souborů
8 data1 = pd.read_csv("PythonTabulka1CC.csv", sep=';')
9 data2 = pd.read_csv("PythonTabulka2C.csv", sep=';')
10
11 #Transformace hodnot
12 data1["Smart city dokument"] = data1["Smart city dokument"].replace("Psaný text", 0).replace("Prezentace", 1)
13 data2["Smart city dokument"] = data2["Smart city dokument"].replace("Psaný text", 0).replace("Prezentace", 1)
14 data1["Region"] = data1["Region"].replace("Evropa", 0).replace("Asie", 1).replace("Zbytek světa", 2).replace("Severní Amerika", 3)
15 data1["Má big data?"] = data1["Má big data?"].replace("Ne", 0).replace("Ano", 1)
16 data2["Region"] = data2["Region"].replace("Evropa", 0).replace("Asie", 1).replace("Zbytek světa", 2).replace("Severní Amerika", 3)
17
18 #Výběr sloupců
19 vyber1 = data1[["Pořadí v indexu", "Region", "City HDI", "Počet obyvatel (2023)", "Country HDI", "Smart city dokument",
20 "Rok vydání strategického dokumentu", "Počet stran", "Má big data?"]]
21 vyber2 = data2[["Pořadí v indexu", "Region", "Smart city dokument", "Počet zmínek big data", "Doprava", "Big data analytics", "Věci spojené s řízením města",
22 "Ekonomika, business", "Technologické změny", "Další pojmy související s Big data problematikou"]]
23
24
25 #Standardizace a normalizace
26 standard1 = preprocessing.scale(vyber1)
27 standard2 = preprocessing.scale(vyber2)
28
29 normal1 = preprocessing.normalize(standard1)
30 normal2 = preprocessing.normalize(standard2)
```

```
31
32 #Shluková analýza
33 #Úvodní tabulka
34 colors = ['blue', 'green', 'orange', 'red', 'brown', 'purple', 'pink', 'black', 'yellow']
35 kmeans1 = KMeans(init="random", n_clusters=3, n_init=10, max_iter=300, random_state=42)
36 label1 = kmeans1.fit_predict(normal1)
37 clusters1 = kmeans1.labels_
38 centers1 = kmeans1.cluster_centers_
39
40 print("První tabulka")
41 j = 0
42 while(j < len(clusters1)):
43     print(data1["Český název"].values[j], clusters1[j])
44     j+=1
45
46 u_labels1 = np.unique(label1)
47
48 for i in u_labels1:
49     plt.scatter(normal1[label1 == i, 0], normal1[label1 == i, 1], label=i, c=colors[i])
50     plt.scatter(centers1[i, 0], centers1[i, 1], c=colors[i], marker='x', edgecolor='black', s=200, alpha=1)
51     plt.scatter(centers1[i, 0], centers1[i, 1], c=colors[i], s=25000, alpha=0.10)
52 plt.legend()
53 plt.show()
```

```
54
55 #Tabulka big dat
56 kmeans2 = KMeans(init="random", n_clusters=4, n_init=10, max_iter=300, random_state=42)
57 label2 = kmeans2.fit_predict(normal2)
58 clusters2 = kmeans2.labels_
59 centers2 = kmeans2.cluster_centers_
60
61 print("Druhá tabulka")
62 j = 0
63 while(j < len(clusters2)):
64     print(data2["Český název"].values[j], clusters2[j])
65     j+=1
66
67 u_labels2 = np.unique(label2)
68
69 for i in u_labels2:
70     plt.scatter(normal2[label2 == i, 0], normal2[label2 == i, 1], label=i, c=colors[i])
71     plt.scatter(centers2[i, 0], centers2[i, 1], c=colors[i], marker='x', edgecolor='black', s=200, alpha=1)
72     plt.scatter(centers2[i, 0], centers2[i, 1], c=colors[i], s=10000, alpha=0.10)
73 plt.legend()
74 plt.show()
```