

Univerzita Pardubice

Fakulta ekonomicko-správní

Otevřená data a datové portály na úrovni regionů a měst

Bakalářská práce

2022

Veronika Greyová

Univerzita Pardubice
Fakulta ekonomicko-správní
Akademický rok: 2021/2022

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Veronika Greyová**
Osobní číslo: **E19845**
Studijní program: **B0688A140004 Informatika a systémové inženýrství**
Specializace: **Informační a bezpečnostní systémy**
Téma práce: **Otevřená data a datové portály na úrovni regionů a měst**
Zadávací katedra: **Ústav systémového inženýrství a informatiky**

Zásady pro vypracování

Cílem práce je popsat charakteristiky zdrojů otevřených dat na úrovni regionů a měst a analyzovat vybrané přístupy řešení.

Osnova:

- Základní pojmy řešené problematiky.
- Charakteristiky zdrojů otevřených dat.
- Analýza existujících přístupů řešení.
- Formulace závěrů a doporučení.

Rozsah pracovní zprávy: **cca 35 stran**
Rozsah grafických prací:
Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam doporučené literatury:

- BOČEK, Jan, Jakub MRÁČEK a Jindřich MYNARZ. (2012). *Otevřená data: příležitost pro Českou republiku*. Praha: Fond Otakara Motejla. ISBN 978-80-87725-03-0.
- MRÁČEK, Jakub, Jan BOČEK a Jáchym ČEPICKÝ. (2014). *Jak otevírat data?*. Praha: Fond Otakara Motejla. ISBN 978-80-87725-15-3.
- DAVIES, Tim, Stephen B. WALKER, Mor RUBINSTEIN a Fernando PERINI (Eds.). (2019). *The State of Open Data: Histories and Horizons*. Cape Town and Ottawa: African Minds and International Development Research Centre. ISBN 978-1-55250-612-7.
- HAWKEN Scott, Hoon HAN a Chris PETTIT (Eds.). (2020). *Open Cities | Open Data: Collaborative Cities in the Information Era*. Singapore: Palgrave Macmillan. ISBN 978-981-13-6605-5.
- CHARALABIDIS, Yannis, Charalampos ALEXOPOULOS, Enrico FERRO, Marijn JANSSEN, Thomas LAMPOLTSHAMMER a Anneke ZUIDERWIJK. (2018). *The World of Open Data: Concepts, Methods, Tools and Experiences*. Cham: Springer. ISBN 978-3-319-90850-2.

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

Datum zadání bakalářské práce: **1. září 2021**
Termín odevzdání bakalářské práce: **30. dubna 2022**

L.S.

prof. Ing. Jan Stejskal, Ph.D.
děkan

RNDr. Ing. Oldřich Horák, Ph.D.
vedoucí ústavu

Prohlášení:

Prohlašuji:

Práci s názvem *Otevřená data a datové portály na úrovni regionů a měst* jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena 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. 4. 2022

Veronika Greyová

Poděkování

Tímto bych ráda poděkovala rodině, díky jejichž podpoře mi bylo umožněno studovat a svému vedoucímu práce, panu Ing. et Ing. Martinu Lněničkovi, Ph.D. za trpělivost a ochotu.

ANOTACE

Práce se zabývá charakteristikami otevřených dat a analýzou datových portálů na úrovni měst. Obsahuje vymezení základní pojmů, funkce a služby datových portálů včetně ukázek z praxe. Vybrané datové portály hlavních měst v Evropě jsou ohodnoceny na základě metodiky indexu GODI. V závěru jsou diskutovány výsledky a představena doporučení s ohledem na zjištění i aktuální trendy.

KLÍČOVÁ SLOVA

otevřená data, datové portály, charakteristiky, analýza, GODI metodika

TITLE

Open data and data portals at the regional and city level

ANNOTATION

The thesis deals with the characteristics of open data and analysis of data portals at the city level. It contains the definition of basic concepts, features and services of data portals, including examples from practice. Selected data portals of capital cities in Europe are evaluated on the basis of the GODI index methodology. Finally, the results are discussed and recommendations are presented with regard to the findings and current trends..

KEYWORDS

open data, data portals, characteristics, analysis, GODI methodology

Obsah

Úvod	10
1 Základní pojmy	12
1.1 Historie otevřených dat	12
1.2 Éra otevřeného vládnutí	12
1.3 Koncepce otevřených dat	13
1.4 Potenciál otevřených dat	15
1.5 Metodika publikace otevřených dat ve veřejné správě ČR	16
1.5.1 Datový katalog	16
1.5.2 Role při publikaci a katalogizaci otevřených dat.....	16
1.6 Právní legislativa otevřených dat	17
1.7 Datové portály	18
2 Charakteristiky zdrojů otevřených dat	20
2.1 Ukládání a správa otevřených dat	20
2.1.1 Webová úložiště.....	20
2.1.2 Placený cloud.....	21
2.1.3 Vlastní server	21
2.1.4 Systémy pro správu otevřených dat	21
2.2 Pětihvězdičkové schéma	22
2.3 Formáty otevřených dat	23
2.4 Funkce a služby datových portálů	23
2.5 Datové portály a aplikace nad otevřenými daty	24
2.5.1 Datové portály v ČR	25
2.5.2 Otevřená data a datové portály na úrovni EU.....	26
3 Analýza vybraných portálů	27
3.1 Global Open data Index	27

3.2	Postup analýzy	30
3.3	Výsledky a zjištění.....	30
3.3.1	Obecný popis portálů	30
3.3.2	Vyhledání evropských portálů	31
3.3.3	Hodnocení vybraných portálů.....	33
3.3.4	Z pohledu GODI kritérií	34
4	Doporučení a aktuální trendy	36
4.1	Best practices EU.....	37
4.2	Best practices USA	38
	Závěr.....	40
	Použité zdroje	42

Seznam tabulek

Tabulka 1: Přehled evropských měst a jejich datových portálů. Zdroj: autor.31

Tabulka 2: Výsledné ohodnocení vybraných datových portálů na úrovni měst. Zdroj: autor..33

Seznam zkratk

API	Application Programming Interface
ČR	Česká republika
CSV	Comma Separated Values
EU	Evropská unie
GODI	Global Open Data Index
HDP	Hrubý domácí produkt
ICT	Informační a komunikační technologie
ISVS	Informační Systém Veřejné Správy
JSON	JavaScript Object Notation
OGD	Open Government Data
PDF	Portable Document Format
PSI	Public Sector Information
RDF	Resource Description Framework
SPARQL	Simple Protocol and RDF Query Language
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
USA	Spojené státy americké
XML	Extensible Markup Language

ÚVOD

Informační technologie a jejich využívání je dnes vnímáno jako norma a něco, co je každodenní a běžné. Je normální mít doma počítač, notebook, tablet nebo mobilní telefon. Je normální využívat sociální sítě a další komunikační prostředky. A je normální využívat internet a pomocí něj přistupovat k různým webovým portálům a zdrojům dat a informací.

Mezi klíčové trendy moderní veřejné správy patří otevřenost, transparentnost a poskytování služeb, které vycházejí vstříc potřebám občanů. Zpravidla se jedná o online služby dostupné přes webová rozhraní, kde je často vyžadována autorizace a autentizace uživatele. V těchto případech se jedná o služby, kde je nutné ověřit identitu uživatele, který má pak online přístup k odpovídajícím datům v informačních systémech a registrech veřejné správy. Toto je stěžejní oblast e-governmentu, neboli elektronické veřejné správy, kdy je občan nebo podnik v nějaké oficiální formě interakce s institucí veřejné správy.

V posledních letech však nabývá na významu nový pohled na e-government, kdy služby nejsou cíleny na konkrétní subjekty, resp. nevychází z povinností daných zákony a dalšími právními předpisy, ale instituce veřejné správy sami nabízejí data a informace, které může kdokoliv využít. Tento přístup je součástí otevřeného vládnutí, kdy se jednotlivé instituce, obce, města, kraje i stát snaží být proaktivní a transparentní vůči svým občanům. Zpravidla se jedná o zveřejňování dat a informací, na které mají občané právo podle nějaké právní normy. V České republice (ČR) je to Zákon č. 106/1999 Sb. o svobodném přístupu k informacím. Dnes se tato data shromažďují na tzv. datových portálech (open data portály), které může provozovat stát, kraje, města, obce, jednotlivé instituce nebo neziskové organizace.

Klíčové pro tyto portály je, že by měly dodržovat pravidla pro zveřejňování otevřených dat, a zároveň nabízet takové funkce, které uživatelům umožní s těmito daty snadno a rychle pracovat a využívat je pro svoje potřeby. Nicméně ne všechny datové portály toto splňují.

Cílem této práce je popsat charakteristiky zdrojů otevřených dat na úrovni regionů a měst a analyzovat vybrané přístupy řešení. Datové portály na úrovni regionů a měst byly zvoleny proto, že města a obce jsou svým občanům blíže a lze předpokládat, že zveřejňují taková data, která mají pro jejich občany význam. Z tohoto důvodu je nutné analyzovat to, jak tyto portály splňují pravidla pro otevřená data a formulovat takové závěry, které umožní jejich další rozvoj a zlepšování nabídky služeb pro občany.

V první části jsou rozepsány základní pojmy. Poté následují charakteristiky zdrojů otevřených dat. Hlavní část práce je věnována analýze existujících řešení, tzn. výběru, rozdělení, popisu, porovnání a vyhodnocení výsledků datových portálů na úrovni regionů a měst podle vybraného postupu. V poslední kapitole jsou pak zformulovány závěry z provedené analýzy a představena doporučení směřující ke zlepšení datových portálů s ohledem na dodržování pravidel pro otevřená data, včetně zasazení do kontextu aktuálních trendů.

1 ZÁKLADNÍ POJMY

1.1 HISTORIE OTEVŘENÝCH DAT

Otevřená data se dostala do popředí zájmu vlád po roce 2009, kdy tehdejší prezident Spojených států amerických Barack Obama představil svou koncepci otevřeného vládnutí (Davies et al., 2019). V evropském kontextu lze zmínit např. spor několika britských autorit s místním ministerstvem o jízdní řády, které bylo možno do té doby využívat pouze v rámci komerčního využití s velmi drahou licenci. K byrokratickému odboji se přidal i tehdejší premiér Gordon Brown. Tlak byl nakonec úspěšný. Ministerstvo data zveřejnilo a to zdarma. Bylo tedy možné je využít kýmkoliv na cokoliv. Tento krok měl velkou odezvu. Začaly být vytvářeny úspěšné aplikace a zvedl se zájem o další data z různých oblastí (Boček et al., 2012).

1.2 ÉRA OTEVŘENÉHO VLÁDNUTÍ

I přes přetrvávající předsudky a strach z otevírání dat všem zdarma, stále více států, měst i obcí dnes zpřístupňuje svá data online. Nicméně je potřeba si ale uvědomit, že otevřená éra vládnutí je sama o sobě nová myšlenka, prakticky nový princip. Nejenom v minulosti lze nalézt mnoho příkladů kdy vlády nebo jiné orgány nestály o to, aby někde mohl nahlédnout do struktury jejich fungování nebo nakládání s finančními prostředky. Myšlenka toho, že lidé tu nejsou pro stát, ale naopak stát je tu pro lidi, k nám přichází uceleně až v režimu demokracie, která je v novodobé historii také poměrně nová. Co má tato historie společného s otevřenými daty? Prostý fakt, že otevřená data vyžadují změnu smýšlení. Jak obyčejných lidí, tak lidí, kteří mají v rukách odpovědnost za svůj stát. Pokud má stát sloužit lidem, být jejich oporou, pak je naopak potřeba, aby lidé měli přístup k informacím. Svobodně, bez překážek a zdarma (Davies et al., 2019; Harari, 2012).

Současná éra otevřeného vládnutí se vyznačuje tím, že občané již nemusí žádat různé instituce o zveřejnění dat v otevřených formátech, ale tyto instituce je zveřejňují automaticky a rovněž je udržují aktuální. Státy, města, ale i obce zároveň vytvářejí různé metodiky a doporučení, jak postupovat při zveřejňování otevřených dat. Zkušenosti z různých oblastí dokazují, že podpora otevřeného vládnutí je cesta ke zvyšování transparentnosti a důvěryhodnosti veřejného sektoru v očích široké veřejnosti (Charalabidis et al., 2018).

1.3 KONCEPCE OTEVŘENÝCH DAT

Jak již bylo zmíněno, otevřená data jsou poměrně nová, a proto lze nalézt více definic. Podle Open Knowledge¹, která ve variantě 2.1 říká, že „*Data jsou otevřená, pokud k nim může každý volně přistupovat, používat je, upravovat a sdílet za podmínek, které nejvíce zabezpečí zachování původu, tedy vlastnictví a otevřenost.*“ Což zjednodušeně znamená, že data jsou k dispozici online a kdokoliv je může využívat, měnit a sdílet bez jakýchkoliv omezení. Otevřenost je tedy klíčová vlastnost, když při zákazu jejich zpracování, či komerčního užití se jedná o veřejná data, ale nikoli o otevřená data (Myška et al., 2014). Je potřeba, aby data byla otevřená technicky i právně. Což znamená z technického požadavku (Open Definition, 2014):

- otevřenou licenci nebo stav,
- být v maximálním možném rozsahu původního zdroje,
- volný a snadný přístup k bezplatnému stažení přes internet s minimálním úsilím a náklady,
- strojovou čitelnost,
- a otevřený formát.

Z těchto právních musí umožňovat (Open Definition, 2014):

- volné užití ve smyslu licencí,
- redistribuci,
- modifikaci,
- oddělení práv od jiných děl, aniž by to znamenalo omezení.

Data nesmí nikoho diskriminovat ani být zpoplatněna (Open Definition, 2014).

Definice od Sunlight Foundation² formuluje deset principů, ve kterých jsou podrobně vymezeny podmínky, které musí data veřejné správy splňovat, aby je bylo možné považovat za otevřená data. Těmito podmínkami jsou (Ten principles for opening up government information, 2017):

¹ Open Knowledge – mezinárodní nezisková organizace orientovaná na otevřená data.

² Sunlight Foundation – nezisková organizace věnující se otevřeným datům na vládní úrovni.

1. Úplnost – data musí být v maximálním možném rozsahu, který je definován právním předpisem, vládním usnesením či poskytovatelem dat. Například seznam všech památkově chráněných objektů v dané obci.
2. Snadná dostupnost – data musí být snadno dostupná internetu a dohledatelná běžně dostupnými ICT nástroji a prostředky. Neměla by existovat omezení pro různá zařízení, např. osobní počítač, tablet, mobilní telefon atd.
3. Strojová čitelnost – data a jejich struktura by měly být ve formátu vhodném pro strojové zpracování.
4. Používání standardů s volně dostupnou specifikací (otevřené standardy) – data musí být ve formátu, který je volně (zdarma) dostupný pro libovolné použití nebo do takového formátu převoditelný volně (bezplatně) dostupnou aplikací.
5. Přístupnost za jasně definovaných podmínek užití dat (licence) s minimem omezení – je potřeba, aby podmínky využívání dat byly jasně, zřetelně a veřejně definovány.
6. Dostupnost dat při vynaložení minima nákladů na jejich získání – poskytovatelé jsou v souvislosti s poskytováním dat oprávněni žádat úhradu maximálně ve výši, která nesmí přesáhnout náklady spojené s jejich zpřístupněním uživateli. Poskytovatel dat může i jednorázově vyžádat úhradu za mimořádně náročné pořízení dat, pokud si uživatel zpřístupnění těchto dat vyžádá.

Dále je vhodné, aby otevřená data byla:

7. Primární (původní) – data, která jsou zveřejněna původcem dat v podobě, v jaké byla původcem jako primární (původní) vytvořena.
8. Zveřejněná bez zbytečného odkladu – zveřejnění dat není zdržováno jinými než nezbytně nutnými činnostmi. Nezbytně nutné činnosti jsou provedeny v čase, který umožní jejich zveřejnění bez nepřiměřeně dlouhé prodlevy.
9. Neomezující přístup – data musí být dostupná způsobem, který nediskriminuje jednotlivce nebo skupinu osob.
10. Stále dostupná – data jsou dostupná online po dobu uvedenou jejich poskytovatelem.

Podle této definice se za otevřená data veřejné správy považují taková data, která vyhovují alespoň bodům 1-6 výše.

1.4 POTENCIÁL OTEVŘENÝCH DAT

Otevřená data mají zejména potenciál v podobě zpětné vazby, resp. kontroly vůči poskytovateli dat a přispívají ke zvyšování transparentnosti a otevřenosti celého veřejného sektoru vůči svým občanům. Znovu využívání otevřených datových sad je realizováno především přes webové či mobilní aplikace, které tato data využívají. Mezi hlavní důvody pro publikování dat náleží (Mráček et al., 2014):

- Lepší služby – nad otevřenými daty se dají vytvářet modely či využívat nejrůznější analytické nástroje, především vizualizace dat, díky které je možné odhalit nové souvislosti a vztahy mezi daty.
- Ekonomický potenciál – aplikace mohou generovat významné zisky a vracet tak státu náklady na otevřená data v daních hned několikanásobně. Jejich ekonomický potenciál může spočívat i ve schopnosti peníze ušetřit.
- Transparentnost – mohou zprůhlednit fungování státu a zefektivnit jeho fungování.
- Efektivnější správa – procesy mohou fungovat rychleji a efektivněji.

U nás jsou data stále na půli cesty. Máme v tomto ohledu mnoho nedostatků, jako je omezený přístup k nim, či zveřejnění ve strojově nečitelném formátu, jako je například PDF. Přitom otevřená data mohou pomoci v boji s korupcí. Stačí si jen představit, jak těžké by muselo být jen nějakou korupci provést, pokud by se zveřejňovala výběrová řízení, kompletně. Včetně nabídek od firem. A nad těmito daty by fungoval inteligentní srovnávač či přímo aplikace zaměřená na pomoc při rozhodování. Bylo by možné nastavit a posoudit priority. Někdy ve velmi vzdálené budoucnosti by lidé mohli mít i možnost hlasovat a vyjádřit svůj názor. Stát chce koupit to nejlevnější? Jistě bude spousta lidí, co s touto prioritou na ceně bude souhlasit, ale co když se některá z nabídek výrazně více vyplatí v budoucnosti? Vláda by mohla získat hodnotnou zpětnou vazbu a údaje o lidech ve svém státě, ke kterým se jinak nedostane. Navíc by takové možnosti mohli vést i k diskusi nad těmito tématy ve společnosti, které jsou obecně dobré pro usměrnění názorů (Hawken et al., 2020).

1.5 METODIKA PUBLIKACE OTEVŘENÝCH DAT VE VEŘEJNÉ SPRÁVĚ ČR

1.5.1 Datový katalog

Ve veřejném sektoru jsou data ukládána v různých Informačních Systémech Veřejné Správy (ISVS), ale i v listinné podobě (Myška et al., 2014). Katalog otevřených dat veřejné správy ČR, dále jen Datový katalog, je jednotné místo, kde je možné vyhledat otevřená data veřejné správy a získat údaje o nich (metadata). Datový katalog primárně k ukládání otevřených dat neslouží, ukládá pouze záznamy o nich, na základě kterých jsou jeho uživatelé schopni vyhledat data na jednom místě, bez potřeby procházet jednotlivé izolované webové portály jednotlivých orgánů veřejné správy. Zároveň záznamy v Datovém katalogu obsahují i popis dat, díky kterým je usnadněna práce s těmito daty (Klímeck, 2020; Klímeck, 2021).

1.5.2 Role při publikaci a katalogizaci otevřených dat

Při publikaci otevřených dat spolupracuje několik rolí, které jsou uvedeny níže. Pro úplnost je uvedena i role koncového uživatele, který data bude využívat a k vyhledávání využije Datový katalog. Vymezení rolí je následující (Ministerstvo vnitra ČR, 2015):

Poskytovatel dat

Poskytovatelem dat je pro oblast otevřených dat ve veřejné správě jakýkoliv orgán veřejné správy, který může a nemusí mít povinnost zveřejňovat data, ale v rámci svých kompetencí se rozhodne otevřená data zveřejnit. Pokud data pochází z ISVS, pak je poskytovatelem těchto dat správce ISVS.

Kurátor dat

Kurátor je osoba, starající se o zveřejnění otevřených dat a údržbu záznamů o datech v Datovém katalogu. Zpravidla je u záznamu o otevřených datech uveden kontakt právě na tuto osobu, na kterou by mělo být možno se v případě problémů obrátit.

Správce datového katalogu

Správce Datového katalogu je osoba či subjekt odpovědný za Datový katalog jako takový. Správce Datového katalogu může být zároveň i jeho provozovatel. Obecně se předpokládá, že roli správce Datového katalogu bude vykonávat Ministerstvo vnitra ČR.

Provozovatel Datového katalogu

Zpravidla provádí činnosti související se zajištěním běžného chodu a údržbou Datového katalogu a potřebného softwarového a hardwarového vybavení. Mimo jiné zodpovídá i za zajištění bezpečnosti Datového katalogu.

Redaktor

Redaktorem je zpravidla osoba, která se stará o obsahovou kontrolu vložených záznamů o otevřených datech veřejné správy ČR. Předpokládá se, že obsazování role redaktora spadá do působnosti Ministerstva vnitra ČR.

Koncový uživatel

Koncový uživatel je osoba, která 1) vyhledává záznamy o otevřených datech veřejné správy ČR v Datovém katalogu, 2) posílá správci Datového katalogu podněty na vytvoření nových záznamů, respektive úpravu stávajících záznamů v Datovém katalogu, a 3) využívá data v souladu s podmínkami jejich užití. Jako koncový uživatel může vystupovat kdokoliv.

1.6 PRÁVNÍ LEGISLATIVA OTEVŘENÝCH DAT

Předpokládá se, že informace jsou nedílnou součástí svobodného projevu. V rámci veřejného sektoru je problematika informací přímo zakotvená v zákoně č. 106/1999 Sb., o svobodném přístupu k informacím. Ten vymezuje i otevřená data. Právní úprava informací veřejného sektoru (Public Sector Information, PSI) je rozvětvená do dvou částí. První část je míněna jako přístup k informacím jako projev základního politického práva na informace. Druhá část má účel ekonomický, naplňovaný možností opětovného užití PSI pro další předem nedefinované účely, včetně užití komerčního (Zákonná definice otevřených dat, 2020).

Možnost opětovného užití PSI je v Česku založena zásadou legální licence, a je tedy limitována jen výjimkami a omezeními vyplývajícími z právních předpisů, jako je například ochrana osobních údajů, obchodní tajemství, nebo duševní vlastnictví (Myška et al., 2014). Obvyklými předpoklady otevřených dat jsou zveřejnění PSI online v otevřeném a strojově čitelném formátu, který má minimální právní překážky a takové podmínky užití, které nijak neomezují způsob a účel jejich dalšího využití. Tomu rovněž odpovídá zákonná definice v podobě § 3 odst. 11 zákona č. 106/1999 Sb., která navíc přináší podmínku registrace zveřejňovaných otevřených dat v Národním katalogu otevřených dat³, jehož účelem je umožnit efektivní vyhledávání mezi

³ Národní katalog otevřených dat je ISVS ve správě Ministerstva vnitra, fungující jako seznam otevřených dat.

publikovanými otevřenými daty. Vedle zavedení definice a zakotvení Národního katalogu přinesla tzv. Open Data novela provedená zákonem č. 298/2016 Sb. třetí zásadní novinku v podobě povinných otevřených dat, které byly určeny nařízením vlády č. 425/2016 Sb. V roce 2019 byla přijata směrnice EU č. 2019/1024, o otevřených datech a opakovaném použití informací veřejného sektoru, která s účinností od července 2021 nahradila směrnici z roku 2003 a její novelu.

První z nich spočívá v rozšíření působnosti směrnice na veřejné podniky ve smyslu definice vycházející ze směrnice 2014/25/EU, tedy na obchodní společnosti podnikající v klíčových oblastech fakticky ovládané státem. Druhá novinka spočívá v zavedení tzv. datových sad s vysokou socioekonomickou hodnotou, které budou stanoveny prováděcím předpisem k Open Data Směrnici, a které budou povinně zveřejňovány napříč EU. V současné době nejsou známy konkrétní informace, které budou takto zveřejňované. Open Data Směrnice nicméně stanoví alespoň obecné kategorie, z nichž mají být vybrány. Jedná se o informace o pozorování Země a životní prostředí, meteorologii, statistické údaje, nebo údaje o vlastnictví společností. Konečně třetí zásadní novinka spočívá ve výslovném zahrnutí údajů z vědeckého výzkumu prováděného z veřejné podpory do aplikačního rozsahu směrnice. Současná úprava česká otevřených dat stanoví jasné povinnosti určující, které informace mají být jako otevřená data publikovány. Zároveň však nechává otevřené dveře možnosti poskytovat otevřená data dobrovolně nad rámec těchto povinností (Komárková, 2017).

Další právní variantou pro otevřená data je licence Creative Commons⁴. Klade si za cíl rozšířit množství autorských děl dostupných veřejnosti k legálnímu využívání a sdílení. Konkrétně Creative Commons 4.0 je veřejná licence, nyní kompletně přeložena do českého jazyka. Díky této veřejné licenci je možné dát zájemci svolení k užití obecně, užití ke komerčním účelům nebo editaci (Myška et al., 2014).

1.7 DATOVÉ PORTÁLY

Úrovně otevřených dat mohou jít od těch nejmenších na úrovni obcí, přes úroveň měst, států či společenství. Existují i data na úrovni EU. Specifická jsou tím, že EU je společenství několika zcela odlišných států, které ale dohromady mají významnou obchodní sílu. Portál pro otevřená

⁴ Creative Commons je americká nezisková organizace pracující s autorskými právy.

data na úrovni EU nabízí jak data jednotlivých států, tak i agregovaná data, která se týkají celé EU (Oficiální portál evropských dat, 2022).

Datový portál (datový katalog), je zpravidla definován jako webová stránka, na které lze nalézt datové sady. Ty jsou nějak uspořádané, např. dle kategorií, a lze mezi nimi vyhledávat, filtrovat podle kritérií, hodnotit a komentovat, žádat o novou sadu, případně provádět různé analytické operace s daty, např. souhrny. Uživatelé někdy mohou na portál nahrávat i upravené datové sady nebo aplikace vytvořené nad zde zveřejněnými daty. Předpokládá se, že spolu s datovými sadami budou nalezeny na těchto portálech i metadata, která budou podávat informace o datové sadě. Zároveň portály nabízí i API rozhraní pro vývojáře. Kromě těchto funkcí, které lze na portálu využívat, je důležitá také použitelnost celého portálu, protože pokud uživatelů nemohou najít požadovaná data, tak je nemohou ani dále využívat (Máchová a Lněnička, 2017; Thorsby et al., 2017).

Nejrozšířenějším typem datového portálu je takový, kde datové sady splňují pravidla platná pro otevřená data, viz výše. Zdrojem těchto dat může být buď pouze jedna instituce (poskytovatel dat) nebo se na portálu může zaregistrovat více poskytovatelů a uživatelé tak mají na výběr mezi více datovými sadami na jednom místě. Toto jsou zpravidla portály na národní úrovni, kde většina datových sad platí pro celý stát. Postupně však portály s otevřenými daty začaly spouštět i regiony a města, čímž se dostávají blíže ke svým občanům, kterým jsou zpravidla bližší konkrétní datové sady z jejich města a okolí než z celého státu (Máchová a Lněnička, 2017; Thorsby et al., 2017).

2 CHARAKTERISTIKY ZDROJŮ OTEVŘENÝCH DAT

Tato kapitola se věnuje charakteristikám zdrojů otevřených dat ve formě datových portálů. Jak bylo zmíněno již výše, tak datové portály jsou webové stránky a platí pro ně pravidla pro tvorbu webových stránek, požadavky na přístupnost, použitelnost atd. Tyto charakteristiky nicméně nebudou dále rozebírány a pozornost bude věnována funkcím portálů a vlastnostem datových sad.

2.1 UKLÁDÁNÍ A SPRÁVA OTEVŘENÝCH DAT

Zde je nutné zmínit, že otevřená data nejsou jenom věc veřejného sektoru, ale data mohou volně zveřejňovat i neziskové organizace, podniky, občané atd. Otevřená data ve veřejném sektoru se nazývají Open Government Data (OGD) a jsou na ně při zveřejňování a správě kladeny zejména požadavky související s ochranou osobních údajů a bezpečností státu (např. vojenská zařízení nebo různé smlouvy).

Na základě toho je pak nutné přistupovat k volbě konkrétního řešení. Náklady na uložení a správu dat také nemusí být zrovna malé a je třeba porovnat i přínosy a náklady. Mráček et al. (2014) se doporučují v prvním kroku rozhodnout, zda se bude jednat pouze o manuální nebo automatizované zveřejňování otevřených dat. Při manuálním lze zveřejnit několik datových sad na zkoušku zájmu ze strany občanů a dalších uživatelů, např. na běžných webových stránkách města či obce. V případě zájmu lze pak přistoupit k automatizovanému řešení, kdy je zvoleno řešení, které celou správu otevřených dat usnadňuje a zároveň nabízí různé funkce, které mohou uživatelé používat.

2.1.1 Webová úložiště

Principem je, že provozovatel úložiště nabídne prostor za cenu zobrazování reklam. Velkou nevýhodou, obzvláště pro veřejný sektor je, že nepůsobí úplně důvěryhodně. Samo o sobě je to netypické řešení, ale splňuje požadavky a k tomu je zdarma. Pro nízkonákladové řešení stojí za úvahu. Patří sem např. nástroje Google. Je to kancelářský balík zdarma v rámci Google disku. K použití je třeba jen registrace Google účtu, který je také zdarma. Nejužívanějším pro otevřená data z Google balíku je část zvaná Tabulky Google. Tento nástroj má bohatou dokumentaci a práce s daty i metadaty je intuitivní. Navíc Google, společnost založená na vyhledávání umí tuto činnost i nad otevřenými daty. Nástroj je i takový katalog dat, tedy pravděpodobnost nálezu a využití dat je vysoká. Uživatelé potěší fakt, že jim jsou k dispozici analytické nástroje. Správci

otevřených dat zase odpadá starost o server, protože práce s daty probíhá na Google serveru, o který se stará jeho vlastník, včetně vytížení.

Mezi další vhodné aplikace typu webdisk patří také OneDrive nebo Dropbox. Mezi nevýhody všech těchto aplikací náleží především otázka bezpečnosti v souvislosti s OGD, malý prostor, pokud se používá pouze verze zdarma, a omezený počet funkcí, které by umožnily datové sady blíže prozkoumat. Za více prostoru a lepší funkce se musí platit a v tomto případě je výhodnější zvolit placený cloud.

2.1.2 Placený cloud

Velmi podobné řešení jako otevřená uložení, když za poplatek nabízí lepší služby. Nemusí tam být reklama, zaručí vyšší výkon, nonstop podporu či záruku proti výpadku služeb. Většinou je vše i smluvně ošetřeno a hlavní poskytovatelé cloudových služeb se zaměřují i na veřejný sektor a jeho požadavky. Lze takto získat lepší zabezpečení, propojení na další služby nebo možnost integrace s dalšími systémy.

Využít lze Google Cloud Platform, Amazon Web Services a jejich službu přímo zaměřenou na otevřená data, Microsoft a jejich službu Azure Open Datasets atd.

2.1.3 Vlastní server

Možnost, která se zpravidla vyplatí jen velkým institucím. Je potřeba mít někoho, kdo bude server spravovat a zodpovídat za něj. Mezi výhody lze zařadit kontrolu nad daty a jejich možnou správu. Nevýhod je ale hned několik. Mít vlastní server znamená nezanedbatelný nepřetržitý odběr energie, který je potřeba pro zajištění dostupnosti dat. Dále nutné aktualizace a napojení na další systémy (Mráček et al., 2014).

2.1.4 Systémy pro správu otevřených dat

Tyto systémy patří do kategorie online informačních systému a jsou přímo vytvořeny tak, aby odpovídaly požadavkům na zveřejňování otevřených dat. Jde zejména o strukturu katalogu, kdy je nutná podpora vyhledávání v katalogu datových sad podle různých kritérií, rozsah možného popisu a kategorizace každé datové sady, a nástroje pro aktivní práci s nimi.

Nejpoužívanějším systémem je CKAN (Máchová a Lněnička, 2017), který je open-source. Je oblíben pro svou flexibilitu v oblasti využívání vybraných modelů a funkcí, jež lze přizpůsobit podle požadavků provozovatele portálu. Dalším systémem je DKAN, který je opět open-source a zaměřuje se více na problematiku propojování otevřených dat (linked data). Existují rovněž komerční systémy jako Junar, Socrata nebo OpenDataSoft.

2.2 PĚTIHVĚZDIČKOVÉ SCHÉMA

Existuje několik metrik pro hodnocení kvality otevřených dat. Internetový vizionář a vynálezce webu Tim Berners-Lee sestavil hodnocení otevřenost dat na základě hvězdiček. Minimum je 1 hvězdička, nejlepší skóre je 5 hvězdiček. Je považován za standard, ze kterého vychází většina definic otevřenosti datových sad. Každé hvězdičkové hodnocení lze rozvést a hodnotit několik dalších kritérií, které se považuje za podmínku splnění daného počtu hvězdiček. V pětihvězdičkovém schéma vypadá hodnocení následovně (Berners-Lee, 2006):

* Publikace dat a jejich opětovné použití.

Minimum je zveřejnit alespoň nějaká data v jakémkoliv formátu. I kdyby mělo jít jen o naskenované dokumenty, i tak je to první krůček a v některých případech i dobré řešení. Například pro archivovaná data, která jsou zapsána ve strojově nerozlušitelném formátu, a nebylo by efektivní je převádět. I tak ale digitalizovaná data, online zveřejněná, šetří práci úředníkům i lidem, kteří nemusí na úřad.

** Strojově čitelná data.

Toto nejsou jen online data, jsou i strojově čitelná. Tedy srozumitelná nejen pro člověka, ale i pro počítač. Úplně jednoduše data nejsou v podobě obrázků, ale znaků, které zvládne načíst i stroj a dále s nimi pracovat.

*** Otevřené formáty.

Tři hvězdičky považujeme za standard pro veřejný sektor. Je to požadavek, aby datové sady mohly být použity běžnými programy. Data by tedy neměla být ve formátu, který dokáže přečíst jen jeden program, a ještě k tomu placený.

**** Univerzální identifikátory.

V českém prostředí už jsou 4 hvězdičky nadstandard. Univerzální identifikátor zdroje (Uniform Resource Identifier, URI) znamená přiřazení vlastnosti každému sloupci tabulky. Technologické řešení takovéto strojové struktury vyžaduje vysokou míru standardizace anebo kvalifikovanou ruční práci.

***** Propojená data.

Posledním krokem k otevřeným datům je určit, které informace v datech jsou společné, ideálně pomocí kódu. Následně je možné je pomocí dotazovacího jazyka spojit a začít provádět analýzy. A to navzdory tomu, že každý z údajů se nachází v jiné tabulce.

2.3 FORMÁTY OTEVŘENÝCH DAT

Formát otevřených dat je důležitý pro jejich využití. Udává totiž, jak budou data vypadat a jakým způsobem se s nimi bude pracovat. Formáty se liší i ve využitelnosti, když klíčové formáty jsou strojově zpracovatelné a jejich struktura umožňuje např. snadný přenos přes API. Mezi tyto formáty patří (Boček et al., 2012):

- XML – široce používaný formát pro výměnu dat, je možné ho libovolně editovat.
- JSON – snadno čitelný formát pro programovací jazyk, oproti XML nemá zaručenou formální strukturu, ale i tak je snadněji zpracovatelný pro počítač.
- RDF – formát, který umožňuje zobrazení dat ve formě, která je vhodná pro kombinování informací z více zdrojů. RDF data mohou být uchována jako JSON nebo XML. Tim Bernes-Lee považuje RDF za jeden z cílů projektů otevřených dat (Berners-Lee, 2006).
- CSV – jednoduchý formát pro data se stejnou strukturou, zpravidla bývá první řádek hlavička a pod ní všechny její hodnoty.
- Čistý text – formáty TXT jsou velmi dobře strojově čitelné, nicméně obvykle neobsahují informace o struktuře přímo v dokumentu.
- PDF – zpravidla text uložený v univerzální formě pro tisk. K jeho čtení je třeba mít správnou aplikaci. Není vhodný pro strojové čtení, ale spíše pro zpracování.
- Tabulkový procesor – výhodou je možnost použití informace za použití popisku buněk, zároveň je to často používaný formát u institucí veřejného sektoru. Jeho příkladem je například MS Excel. Nevýhoda spočívá v tom, že některá makra nebo formuláře mohou být v jiných programech nefunkční.
- Dokument – asi nejužívanější program pro dokument je MS Word. Tento formát je dobrý pro zobrazení relativně stálých informací. Zpravidla tyto formáty neprochází konverzí, jak byly vytvořeny, tak jsou zveřejněny. Z tohoto důvodu jsou tyto formáty sice levné, ale nemívají pevnou strukturu, takže je těžké je strojově číst.

2.4 FUNKCE A SLUŽBY DATOVÝCH PORTÁLŮ

Pro využívání otevřených dat je nutné, aby datový portál nabízel takové funkce a služby, které uživatelům rychle a efektivně umožní pracovat s datovými sadami. Tyto funkce a služby se však liší podle toho, jaký je cíl a potenciální uživatelé portálu. Některé portály jsou zaměřené pouze na zveřejňování dat, jejich vyhledání a stažení. Jiné portály nabízejí více funkcí také pro

vizualizace a analýzy datových sad. Další datové portály pak preferují podporu interakce v rámci komunity otevřených dat, tzn., že tyto portály nabízí více funkcí pro zpětnou vazbu, hodnocení datových sad, žádosti o nové datové sady, diskusní fóra, hackathony atd. Dále je nutné vzít v úvahu i cílovou skupinu uživatelů, především jejich znalosti, tzn., že provozovatel portálu může nabízet jednoduché funkce a služby pro začátečníky a středně pokročilé uživatele, nebo se může zaměřit na pokročilé uživatele a experty a tomu přizpůsobit datový portál (Hawken et al., 2020; Lněnička et al., 2021).

Lněnička et al. (2021) definovali tyto základní kategorie funkcí a služeb:

1. Učení se a prozkoumávání – portál poskytuje dokumentaci, návody, pokyny, sekci pro odborníky atd.,
2. Hledání a filtrování – portál poskytuje vyhledávací pole, procházení témat a kategorií, různé možnosti filtrování atd.,
3. Výběr a stažení – portál poskytuje metadata, statistiky, různé formáty, možnosti stahování atd.,
4. Analýza a propojení – portál poskytuje API a SPARQL, náhled datových sad, nástroje pro propojování datových sad atd.,
5. Vizualizace a interakce – portál poskytuje vizualizace dat v grafech a mapách včetně interaktivních prvků atd.,
6. Hodnocení a diskuse – portál poskytuje funkce pro komentování a hodnocení datových sad, veřejné diskusní fórum atd.,
7. Zpětná vazba a hlášení – portál poskytuje funkce pro zpětnou vazbu, hlášení problémů, kladení otázek atd.,
8. Žádost a návrh – portál poskytuje formulář žádosti o datovou sadu, seznam existujících žádostí a jejich stav, návržení nového zdroje (např. aplikace) atd.,
9. Sdílení a publikování – portál poskytuje seznam otevřených aplikací, možnost nahrávání zdrojů, kanál ke sdílení datových sad nebo aplikací atd.

2.5 DATOVÉ PORTÁLY A APLIKACE NAD OTEVŘENÝMI DATY

Kromě datových sad jsou na datových portálech klíčové aplikace vytvořené s využitím daných datových sad. Tato charakteristika portálů je zásadní pro usnadnění představy uživatelů o tom, co lze z otevřených dat vytvořit a jak je lze využívat v praxi. Vytvořené aplikace pak občanům

a dalším uživatelům mohou pomáhat v běžném životě, kdy jim mohou ušetřit čas nebo náklady na cestování. Snahou je i zvyšování transparentnosti veřejného sektoru. Aplikace se zpravidla dělí na ty webové, tzn. přístupné přes webový prohlížeč, a mobilní. Na portálech jsou rozčleněny do kategorií podle oblastí, např. doprava, zdraví atd., a lze je různě filtrovat. Většina systémů pro správu otevřených dat podporuje i práci s aplikacemi a např. i jejich propojení na konkrétní datové sady, ze kterých byly vytvořeny.

2.5.1 Datové portály v ČR

Jak bylo zmíněno již výše, tak datové portály pro OGD se zpravidla dělí podle územní oblasti, ke které se datové sady vztahují. Existuje několik globálních datových portálů, které nicméně ve většině případů pouze shromažďují data stažená z národních portálů (přes API nebo jiné rozhraní). Nadnárodní úroveň je reprezentována EU. Národní datový portál je nejčastější řešení, které se ve světě objevuje (Máchová a Lněnička, 2017; Lněnička et al., 2021). Další úrovně jsou města, když se to týká hlavně těch větších, které mají dostatek prostředků pro provoz, a hlavně potenciál pro využitelnost otevřených dat. Lze nalézt i portály na úrovni regionů nebo portály pokrývající svazek měst nebo obcí, když tyto přístupy se liší v jednotlivých státech, s ohledem na jejich územní členění a historické a kulturní vazby.

Národní portál otevřených dat ČR lze nalézt na <https://data.gov.cz/>, když v dubnu 2022 zde bylo dostupných více jak 140 000 otevřených datových sad. Kromě nich zde lze nalézt velké množství informací pro poskytovatele a uživatele dat (Portál otevřených dat ČR, 2022).

Na úrovni krajů lze zmínit např. datový portál Královéhradeckého kraje, který je dostupný na <https://www.datakhk.cz/>. Na portálu lze nalézt v horní liště menu, které slouží pro navigaci po stránce. Je zde logo stránek a výběr z možností otevřených dat, aplikace a infografika, stav kraje v datech. Pro zapojení veřejnosti je zde možnost i zapojit se, což spočívá v možnosti sdílet svá data nebo zapojení se do hackathonu, kde je možnost i vývoje aplikací nad otevřenými daty. Poslední položka menu je zpětná vazba, která zahrnuje kontakt na tým portálu, možnost hodnocení portálu dotazníkem a sdílení postojů či preferencí uživatelem portálu.

V katalogu otevřených dat lze řadit datové sady dle relevance, data vytvoření, data aktualizace a názvu. Filtrovat je lze pomocí typu obsahu, poslední aktualizace, zdroje, kategorie a klíčových slov. Po rozkliknutí datové sady se nalevo objeví informace o poskytovateli datové sady a metadata o datové sadě, včetně licence. Lze si otevřít i další informace o datové sadě, vyfiltrovat data, stáhnout, přidat do oblíbených, sdílet nebo přeložit. V případě, že má uživatel zájem data použít, může si vybrat, co s nimi chce dělat. Lze vybrat z vytvoření mapy s těmito

daty, vytvořit příběh v ArcGIS StoryMaps, zobrazit zdroje API, všechna metadata nebo možnost otevřít data v ArcGIS online. Data lze zobrazit v mapě (kde to je možné), ale i z ní se lze jedním klikem dostat zpět na data v podobě tabulky (Datový portál Královéhradeckého kraje, 2022).

Na úrovni měst lze zmínit např. datový portál města Brna (Praha bude součástí analýzy v další kapitole), který je dostupný na <https://data.brno.cz/>. Tento portál kromě datového katalogu obsahuje i sekci pro aplikace vytvořené nad otevřenými daty rozdělené podle kategorií (je jich téměř 80) a sekci pro analýzy otevřených dat skrze dashboardy využívající nástroj Microsoft Power Business Intelligence pro vizualizaci dat (Datový portál města Brna, 2021).

2.5.2 Otevřená data a datové portály na úrovni EU

Stav a vývoj v oblasti otevřených dat na úrovni EU je již od roku 2015 každoročně hodnocen ve zprávách Open Data Maturity Report. Aktuální zpráva hodnotící rok 2021 si klade za cíl podpořit jednotlivé státy, aby lépe porozuměly úrovni jejich vyspělosti v této oblasti, zachytit jejich pokrok v čase, najít oblasti pro zlepšení a porovnat to vše s dalšími státy. Zpráva navíc poskytuje přehled osvědčených postupů úspěšně implementovaných v celé Evropě, které by mohly být dále využity v ostatních státech. Jsou hodnoceny tyto dimenze otevřených dat: 1) politika, 2) dopad, 3) datový portál a 4) kvalita (Van Hesteren a Van Knippenberg, 2021).

Portál evropských dat, dostupný na <https://data.europa.eu/>, lze používat ve všech oficiálních jazycích EU. Nabízí přehledy o počtech katalogů, států a souborů dat. Dále na něm lze nalézt oblíbené datové sady a prohledávat je. Pro jednoduchost jsou zde uvedeny kategorie dat jako jsou doprava, energie, hospodářství a finance, mezinárodní otázky, populace a společnost atd. V záložce data je na výběr ze souborů dat, vyhledávání SPARQL, statistiky, kvalita metadat, URI a slovníky EU. Důležitá jsou ovšem otevřená data. Ta je možná řadit podle otevřených dat, nebo vyhledat katalogy. Samotný řazení je možné podle naposledy upravených, relevance, dle názvu i sestupně a naposledy vytvořených dat. Filtrovat lze podle lokality, kterou lze vybrat i v mapě (Oficiální portál evropských dat, 2022). Nakonec lze zmínit, že tento portál je rovněž rozcestník pro přístup na národní portály států EU.

3 ANALÝZA VYBRANÝCH PORTÁLŮ

Tato práce se věnuje analýze datových portálů na úrovni regionů a měst. Zde je nutné zmínit, že přestože dále bude pracováno pouze s městy, tak hodně datových portálů měst lze považovat za regionální úroveň. Je to z toho důvodu, že menší města nebo obce nemají finanční prostředky na provoz vlastních portálů, a tak často využívají portál nejbližšího města, nebo mohou využít i národní portál.

Vlastní analýzu datových portálů lze provést na základě různých kritérií nebo hodnotících rámců (Máchová a Lněnička, 2017; Lněnička et al., 2021; Van Hesteren a Van Knippenberg, 2021). O ohledem na cíl této práce byla zvolena analýza s využitím indexu Global Open Data Index (GODI), jehož kritéria vychází ze základních požadavků na otevřená data, viz kapitola 1. Tento index se také liší od komplexních metodik a rámců, jako je Van Hesteren a Van Knippenberg (2021), hodí pro datové portály různých typů.

3.1 GLOBAL OPEN DATA INDEX

GODI je nezávislé hodnocení zveřejňování OGD, které bylo poprvé představeno v roce 2013. GODI umožňuje různým zúčastněným stranám v oblasti otevřených dat sledovat pokrok vlády v oblasti uvolňování otevřených dat. GODI také umožňuje vládám získat přímou zpětnou vazbu od uživatelů dat. Index poskytuje oběma stranám základ pro diskusi a analýzu ekosystému otevřených dat v jejich zemi a na mezinárodní úrovni. Nezabývá se dalšími aspekty společného rámce pro posuzování veřejně přístupných dat, jako je kontext, použití nebo dopad. Toto úzké zaměření mu umožňuje poskytovat standardizované a srovnatelné hodnocení otevřených dat po celém světě. Funguje na principu 11 otázek, jejichž odpověď se ohodnotí. Součet ohodnocení jednotlivých otázek poté vyústí v konkrétní číslo, na jehož základě je možné srovnat jednotlivé státy, resp. jejich datové portály (Methodology, 2016).

Otázky použité pro hodnocení portálů se v jednotlivých letech vyvíjely, když následující otázky vychází z metodiky z roku 2016. Pod každou otázkou je vysvětleno i bodové hodnocení, které bylo upraveno pro potřeby této práce, tzn., že bodové hodnocení může nabývat jen 3 hodnot.

1. Jsou data shromážděná vládou?

Řeší se zde, kdo provádí sběr dat, nebo zdali nejsou data shromažďována vůbec. Sběr údajů sám o sobě není charakteristickým znakem otevřených dat, nicméně, pokud si data nesbírá stát nebo třetí strana oficiálně pracující pro vládu, je možno nad

spolehlivostí takových dat debatovat a tázat se, zdali data nebyla upravena například pro vlastní účely.

Hodnocení: 1 bod za vlastní zpracování vládou nebo pověřenou organizací; 0,5 bodů za shromáždění třetí stranou; 0 pokud nelze dohledat.

2. Jsou data vůbec online?

Zde se hodnotí, jestli je vůbec reálné se k datům dostat jakýmkoliv způsobem.

Hodnocení: 1 pokud se lze k datům přes překážky dostat; 0 pokud k nim není možný přístup.

3. Jsou údaje k dispozici online, aniž by bylo nutné se k nim registrovat nebo žádat o přístup k nim?

V tomto případě je zjišťováno, jestli je možné se k datům dostat bez překážek v podobě registrace nebo jiných administrativních kroků.

Hodnocení: 1 bod, pokud je možné data najít a ihned bez překážek stáhnout; 0,5 bodu, pokud je třeba učít nějaký online krok k jejich získání; 0 pokud je třeba podstoupit kroky k jejich získání v reálném světě (fyzicky).

4. Jsou údaje k dispozici zdarma?

Data zdarma jsou taková data, za která není nutné platit.

Hodnocení: 1 pokud jsou data zdarma; 0 pokud je třeba za ně dát finanční obnos.

5. Jsou data uspořádána na jednom místě?

Najít data nemusí být vůbec snadné, obzvlášť, pokud je jich více a každá na jiném místě. Proto by měly na portálu otevřených dat existovat kategorie, pod kterými je možné nalézt vše k požadovanému tématu.

Hodnocení: 1 bod, pokud je možnost data třídit kategoriemi na jednom místě; 0,5 pokud existuje alespoň rozcestník, nebo seznam; 0 pokud je nutné ručně data hledat na vícero místech.

6. Jak náročné bylo nalezení dat?

V tomto bodě se hodnotí to, jak bylo náročné požadovaná data najít. Velmi obecně je otázka položena na počet kliků, které uživatel musel učinit.

Hodnocení: 1 bod, když je odkaz na otevřená data na hlavní stránce, ať už v podobě menu, jeho součásti, nebo jiného odkazu; 0,5 za předpokladu, že k otevřeným datům je cesta jen v určitých jazykových verzích či jiných překážkách; 0 pokud je nutná změna v URL adrese a standardní cestou je nelze dohledat.

7. Jsou data ke stažení najednou?

V tomto kritériu se řeší, jestli je možné data stáhnout z jedné URL adresy, ačkoli jsou v podsouborech kvůli velikosti, nebo rozlišení, nebo je třeba provést několik ručních kroků pro stažení či je možné načíst jen malou část datové sady najednou.

Hodnocení: 1 bod, pokud stačí 1 kliknutí na soubor pro stažení požadovaných dat; 0,5 bodu, pokud je třeba víc kliků, ale data se nachází na jedné URL adrese; 0 bodů, když je nutné procházet stránky pro stažení jedné datové sady.

8. Jsou data aktuální?

V tomto bodě se řeší, zdali je určitelné datum publikace dat a zdali jsou data aktuální. Ideální je odpovědět ano na obojí, neaktuálnost dat by totiž mohla být problémová v případném využití.

Hodnocení: 1 bod, pokud jsou data aktuální a lze nalézt datum aktualizace; 0,5 bodu, pokud lze nalézt datum poslední aktualizace, ale data jsou stará několik let; 0 bodů, pokud časové údaje nejsou k dispozici.

9. Jsou data licencovaná/veřejně přístupná?

Otevřená data musí být opatřena licenci, nebo alespoň prohlášením o možnosti užití. Pokud není zřejmé, že licence nebo podmínky použití jsou v souladu s definicí otevřených dat, tak se o otevřená data vůbec nemusí jednat.

Hodnocení: 1 bod, pokud je u dat uvedená licence, která je v souladu s definicí otevřených dat; 0,5 bodu, pokud není uvedena licence, ale je uvedeno alespoň prohlášení o užití, které splňuje definici otevřených dat; 0 bodů, pokud licence ani prohlášení uvedené není, nebo to není v souladu s definicí otevřených dat.

10. Jsou data v otevřených a strojově čitelných formátech souborů?

Strojové zpracování dat znamená, že je počítač dokáže sám číst, zpracovat a přistupovat k nim. Zároveň musí data být ve formátu, který lze zpracovat alespoň jedním bezplatným softwarovým nástrojem ideálně s otevřeným kódem.

Hodnocení: 1 bod, pokud jsou všechna data ve standardu umožňující strojové zpracování dostupným programovým vybavením, nebo pokud je nejvyšší zastoupení dokumentů ve strojově čitelném formátu; 0 bodů, pokud nesplňuje nic z výše uvedeného.

11. Kolik lidského úsilí je zapotřebí k použití dat?

Toto je poměrně subjektivní hodnocení. Ze své podstaty záleží, na co se daná data mají využít a dle toho je otázkou, nakolik vyhovují danému účelu. Pokud má uživatel v úmyslu data jen přečíst pro zajímavost, nutně mu nemusí vadit PDF soubor. Pokud má ale v plánu data zpracovávat a vytvářet analýzy, pak PDF soubor může tuto práci značně zkomplikovat.

Hodnocení: 1 bod, pokud jsou data ve více formátech; 0,5 bodu, pokud je nejvíce dat ve strojově čitelném formátu; 0 bodů, pokud jsou data ve formátech, které není snadné strojově přečíst.

3.2 POSTUP ANALÝZY

Jako vzorek pro analýzu byly zvoleny portály hlavních měst v Evropě. Prvním krokem bylo vytvoření seznamu států a jejich hlavních měst. Pro každé město byla snaha dohledat portál. Za tímto účelem byla zvolena strategie vyhledávání pomocí klíčových slov přes webový vyhledávač, tzn. název hlavního města + open data (portal). Dalším krokem bylo vyfiltrování portálů na základě kritéria a tím byla dostupnost anglické jazykové verze portálu, ze kterého byl výstupem počet portálů pro analýzu.

Hlavním krokem analýzy je porovnání portálů na základě kritérií definovaných v indexu GODI, včetně zvolené hodnotící škály, viz. předchozí kapitola. Z výsledků je následně vytvořeno pořadí a okomentovány jednotlivé segmenty a struktura portálů, tzn. všechno, co takové portály nabízí svým uživatelům s ohledem na kritéria indexu GODI.

3.3 VÝSLEDKY A ZJIŠTĚNÍ

3.3.1 Obecný popis portálů

Většina hodnocených portálů vypadá podobně. V horní části stránky se nachází zpravidla menu, díky kterému se lze dostat do datového katalogu, když většina portálů používá CKAN nebo DKAN. Menu zpravidla obsahuje také další položky podle zaměření portálu, které by mohly zaujmout i běžného návštěvníka stránek. Pod záložkou otevřených dat se tedy nachází katalog otevřených dat obsahující možnosti řazení a výběr z dat. Ten se nejčastěji provádí na základě

kategorií nebo poskytovatele datových sad. Nejčastěji zastoupené bývají finance obecně, ať už jde o rozpočet a následnou bilanci, tak i zobrazení HDP a dalšího.

Možná je ale i jiná varianta, a to portál pouze pro otevřená data. Ten má na hlavní stránce zpravidla jen odkazy na různé kategorie, což umožňuje rychlejší vyhledávání. Společné jsou pro tyto portály nástroje pro zpětnou vazbu, minimálně to je ve formě textu odkazující na email odpovědné osoby. Některé portály obsahují i informační sekci se články, případně možnost zažádat si o datovou sadu, která zatím není na portálu dostupná.

3.3.2 Vyhledání evropských portálů

Pro vyhledání evropských portálů byl využit internetový vyhledávač Google a kombinace slov město + open data (portal). Pokud odkaz existoval, byl zpravidla v nabídce prvních 5 odkazů. Další variantou bylo nalezení odkazu na národní portál otevřených dat, pod kterým město svá otevřená data zveřejňovalo. Jelikož takové portály ale neodpovídají zadání práce, tak jsou uvedeny pouze v první přehledové tabulce. Ze 46 hlavních měst nemělo 12 měst portál vůbec, resp. nebyl nalezen ani národní portál. Dále 13 měst využívalo portál na národní úrovni, tzn., že tam byla zpravidla sekce určená pro hlavní město. Pro 22 hlavních měst byl portál nalezen.

Pokud mělo město vlastní portál pro otevřená data, bylo ověřeno, zdali je dostupný v angličtině. To znamená, že do hlavní fáze analýzy postoupilo 11 portálů hlavních měst. Výsledná tabulka pro vyhodnocení portálů se skládala ze seznamu států Evropy, jejich hlavních měst, typu, resp. existence datového portálu hlavního města (ano = městský, národní, ne = nenalezen), URL, a dostupnosti anglické verze portálu, viz Tabulka 1.

Tabulka 1: Přehled evropských měst a jejich datových portálů. Zdroj: autor.

Stát	Hlavní město	Typ/nalezen	URL	Angl.
Albánie	Tirana	ano	https://opendata.tirana.al/	ne
Andorra	Andorra la Vella	ne		
Belgie	Brusel	ano	https://opendata.brussels.be/	ano
Bělorusko	Minsk	národní	https://opendata.by/	
Bosna a Hercegovina	Sarajevo	národní	https://bhas.gov.ba/	
Bulharsko	Sofia	ano	http://www.opensofia.info/	ne
Černá Hora	Podgorica	národní	https://data.gov.me/	
ČR	Praha	ano	https://opendata.praha.eu/	ano
Dánsko	Kodaň	národní	https://www.opendata.dk/city-of-copenhagen	
Estonsko	Tallinn	ano	https://avaandmed.tallinn.ee/	ano
Finsko	Helsinky	národní	https://hri.fi/	ano
Francie	Paříž	ano	https://opendata.paris.fr/	ne

Stát	Hlavní město	Typ/nalezen	URL	Angl.
Chorvatsko	Záhřeb	ano	http://data.zagreb.hr/	ne
Irsko	Dublin	ano	https://data.smartdublin.ie/	ano
Island	Reykjavík	ano	https://data-reykjavik.opendata.arcgis.com/	ne
Itálie	Řím	ano	https://dati.comune.roma.it/	ano
Kosovo	Priština	národní	https://opendatakosovo.org/	
Kypr	Nikósie	ne		
Lichtenštejnsko	Vaduz	ne		
Litva	Vilnius	ano	https://api.vilnius.lt/	ano
Lotyšsko	Riga	národní	https://data.gov.lv/	
Lucembursko	Lucemburk	ano	https://data.public.lu/	ano
Maďarsko	Budapešť	národní	http://www.opendata.hu/	
Malta	Valletta	národní	https://open.data.gov.mt/	
Moldavsko	Kišiněv	ne		
Monako	Monaco-Ville	ne		
Německo	Berlín	ano	https://daten.berlin.de/	ne
Nizozemsko	Amsterdam	ano	https://data.amsterdam.nl/	ne
Norsko	Oslo	ano	https://statistikkbanken.oslo.kommune.no/	ne
Polsko	Varšava	ne		
Portugalsko	Lisabon	ano	http://lisboaaberta.cm-lisboa.pt/	ne
Rakousko	Vídeň	ano	https://digitales.wien.gv.at/open-data/	ano
Rumunsko	Bukurešť	národní	https://data.gov.ro/	
Rusko	Moskva	ano	https://data.mos.ru/	ne
Řecko	Atény	národní	https://data.gov.gr/	
San Marino	San Marino	ne		
Severní Makedonie	Skopje	národní	https://data.gov.mk/	
Slovensko	Bratislava	ano	https://opendata.bratislava.sk/	ano
Slovinsko	Lublaň	ne		
Spojené království	Londýn	ano	https://data.london.gov.uk/	ano
Srbsko	Bělehrad	ne		
Španělsko	Madrid	ano	https://datos.madrid.es/	ne
Švédsko	Stockholm	ano	https://dataportalen.stockholm.se/	ne
Švýcarsko	úředně žádné (de facto Bern)	národní	https://opendata.swiss/	
Ukrajina	Kyjev	ne		
Vatikán	Vatikán	ne		

3.3.3 Hodnocení vybraných portálů

Po vyfiltrování měst, které je možné hodnotit, začalo samotné hodnocení. Hodnotilo se na základě metodiky GODI rozepsané výše a výsledky jsou v Tabulce 2. Nejvíce bodů, všech 11 mělo 5 měst, konkrétně Brusel, Praha, Dublin, Řím a Vilnius. Lze tedy tvrdit, že tato města splňují všechny požadavky na otevřená data vycházející z metodiky GODI.

Tabulka 2: Výsledné ohodnocení vybraných datových portálů na úrovni měst. Zdroj: autor.

Město	Kritérium indexu GODI											Součet bodů
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	
Brusel	1	1	1	1	1	1	1	1	1	1	1	11
Praha	1	1	1	1	1	1	1	1	1	1	1	11
Dublin	1	1	1	1	1	1	1	1	1	1	1	11
Řím	1	1	1	1	1	1	1	1	1	1	1	11
Vilnius	1	1	1	1	1	1	1	1	1	1	1	11
Helsinky	1	1	1	1	1	1	1	1	1	1	0,5	10,5
Bratislava	1	1	1	1	1	1	1	1	1	1	0,5	10,5
Londýn	1	1	1	1	1	1	1	1	1	1	0,5	10,5
Lucemburk	1	1	1	1	0,5	1	1	1	1	1	0,5	10
Vídeň	1	1	1	1	1	0,5	1	1	1	1	0,5	10
Tallinn	1	1	1	1	0,5	1	0,5	1	0,5	1	0,5	9

O půl bodu méně měla 3 města a to Helsinky, Bratislava a Londýn. Všechna tři města částečně selhala v posledním bodu hodnocení, a to v možnostech vyžití dat. Pro metodiku GODI bylo hodnocení založeno na předpokladu, že využití pro data může být různé. Obyčejný uživatel může mít zájem se na data podívat ve vizualizovaném prostředí či s možnostmi získat základní přehled o dané datové sadě, zpravidla jako metadata. Pro tyto účely je možné užití PDF formátu. Hlavním posláním otevřených dat je ovšem sloužit dál, což se provádí přes různé programy nebo rozhraní. Lze předpokládat, že málo lidí v naší zemi bude vědět, že existují soubory, které popisují dopravu u nás, jako jsou jízdní řády a podobně. Jenže mnohem větší část lidí zná jejich aplikační využití, jakým je např. Idos⁵. Z tohoto důvodu bylo uděleno hodnocení alespoň půl bodu za předpokladu, že sice nepřevažuje možnost získání dat ve vícero formátech, ale alespoň převažují datové sady ve formátech vhodných pro strojové čtení, což více odpovídá představě o reálném využití těchto dat dál.

⁵ Aplikace pro vyhledávání jízdních řádů v ČR.

O dalšího půl bodu, a tedy s celkovým počtem 10 bodů, je Lucemburk. Stejně jako Bratislava, Helsinky a Londýn ztratil půl bodu na prezentaci dat ve vícero formátech, ale ještě na zobrazení dat na jednom místě. Půl bodu dolů dostal kvůli nemožnosti data řadit podle kategorií. Takže všechny kategorie byly v jednom velkém seznamu a bylo nutné je složitě hledat na několika místech. Stejně si vedla Vídeň. Ta také ztratila půl bodu na využitelnosti dat, ale další půl bod ztratila na složitosti otevřená data najít. V anglické verzi portál zcela změnil strukturu a jsou k nalezení pouze články v anglickém jazyce o různých tématech. Zde bylo zvažováno 0 bodů, ale nakonec se v původním jazyce povedlo datový katalog najít, a proto portál dostal půl bodu v tomto kritériu.

Poslední místo zaujalo město Tallinn. To má speciální portál pouze pro otevřená dat. Půl bodu ztratilo na využitelnosti, dalšího půl bodu na licenci. U datové sady licenci nemělo uvedenou, ale získalo půl bodu alespoň za to, že na hlavní stránce mají prohlášení o možnostech užití otevřených dat, což se bere jako deklarace vůle. Další půl bod město dostalo v otázce stažení dat najednou. Data sice byla na jedné URL adrese, ale po kliknutí na odkaz se nepřešlo rovnou na stažení, ale na XML dokument v prohlížeči, což nenaplnuje podmínku pro udělení 1 bodu. Další půl bod byl stržen za neuspořádání dat na jednom místě. Kvůli nemožnosti data seřadit na základě jedné kategorie je třeba projít celý seznam dat, ale alespoň je ten seznam na jednom místě, za což portál dostal právě ten půl bod.

3.3.4 Z pohledu GODI kritérií

Jednotlivá města a to, jak si počínaly jejich portály, už je okomentováno. Zajímavější, kvůli nadhledu a vyvozování závěrů, je však pohled na to, jak si portály vedly v kritériích obecně. Všechny měly plný počet bodů u kritéria 1, což je získávání dat vládou. Všechny portály měly data získané státem, resp. nějakou institucí veřejného sektoru na úrovni města. Dále bylo možné se u všech webů k datům dostat, zpravidla bez překážek. Všechna data také byla zdarma, aktuální a ve strojově čitelných formátech. To je solidní výsledek vypovídající o tom, že úroveň dat z velké části splňuje samotnou definici otevřených dat.

Jeden portál dostal sníženou známku v šestém kritériu, které řeší náročnost nalezení dat. Pro šesté kritérium je to ale pořád dobrý výsledek. Stejně tak má jeden web snížené hodnocení i u kritérií 7 a další jeden u 9. Kritérium 7 řeší náročnost stažení dat a 9 uvedení licence. Tento fakt vypovídá o malých technických problémech, které jsou poměrně snadno řešitelné. Poslední dvě zmíněná kritéria jsou snadno řešitelná. Pro plný počet bodů u kritéria devět stačí při zveřejnění do šablony vložit odkaz na licenci. Kritérium sedm lze řešit taky poměrně snadno

malou programátorskou úpravou pro otevření odkazu nikoli na další stránce, ale pro stažení, což je kosmetická úprava v HTML kódu. Horší je řešení je u kritéria šest. Zde byl problém se k datům reálně dostat. Po změně jazykové verze nešlo nalézt otevřená data. To je závažná chyba a vzhledem k provedení je otázkou, zdali to tak není účelně, že se cizinci k těmto datům nedostanou. V takovém případě je ale otázkou, nakolik taková data splňují podmínku otevřenosti a ačkoli technické řešení není tak snadné jako u předchozích dvou bodů, zároveň ale není obtížné natolik, aby se nedalo zvládnout.

O malý kousek hůře si vedlo kritérium pět, které hodnotí uspořádání dat na jednom místě. Dva portály dostaly o půl bodu sníženou známku za absenci možného třídění dle kategorií. To zprvu nemusí působit nějak závažně, ale pokud má někdo zájem vytvořit aplikaci sledující např. ekonomický stav země, tak může mít v tomto stavu reálný problém všechna data, která se toho týkají, najít. To značně snižuje samotnou využitelnost takových dat, což je hlavní účel, ke kterému se vůbec zveřejňují. Zde je proto značná mezera, která může vést k velkému vylepšení datových portálů na úrovni měst.

Nejhůře si vedlo kritérium 11, které mělo za cíl hodnotit využitelnost dat jako takovou. Lze argumentovat, že nejvyšší využitelnost dat je tam, kde jsou data dostupná ve více formátech, a kde alespoň jeden formát podporuje možnost strojového čtení. Pokud jsou datové sady dostupné ve vícero formátech, je reálné, že je využije uživatel při snaze se informovat, stejně jako při vytváření aplikací, které mohou data „předložit“ široké veřejnosti, či být využita dalším rozhraním, které danému státu může přinést vyšší životní úroveň, ale hlavně zisk v podobě daní.

4 DOPORUČENÍ A AKTUÁLNÍ TRENDY

Tato kapitola se věnuje doporučením pro další rozvoj a vylepšení datových portálů na úrovni měst a regionů. Ta vychází především z výsledků analýzy provedené v předchozí kapitole a je dále doplněna o aktuální trendy objevující se v literatuře i v praxi.

První doporučení lze formulovat s ohledem na nutnost spuštění vlastního portálu, když pro menší státy je vhodnější mít jeden oficiální národní portál, kde může být sekce pro hlavní město. Takto lze ušetřit finanční prostředky a zároveň se uživatelům usnadní nalezení portálu, protože nemusí dohledávat další. Takovýto přístup je patrný i z Tabulky 1, kde pro rozlohou a počtem obyvatel menší státy nebyl nalezen portál pro hlavní město. Zároveň ale nebyl pro některé státy nalezen ani národní portál. Lze předpokládat, že tyto státy mají otevřená data přístupná např. přes portál veřejné správy nebo využívají nějaký nadnárodní portál, např. ten od EU.

S ohledem na propojení států v Evropě, kdy většina občanů ovládá nějaký cizí jazyk, především angličtinu, lze doporučit, aby portály měst byly dostupné i v tom jazyce. Jelikož jsou na portálech dostupná data v různých kategoriích, lze tímto způsobem získat i další potenciální zájemce o vytváření aplikací, které mohou přispět k rozvoji města (chytrá města).

Dále lze na základě GODI metodiky a jejích výsledků doporučit, aby portály nabízely možnost třídění dat na základě kategorií. Značně tím zvedají reálnou využitelnost dat. Stejně jako tím, že data budou publikovat alespoň ve dvou formátech. Jeden srozumitelný pro člověka s obrázky a nějakými statistickými údaji a agregovanými výstupy a druhý pro strojové zpracování. Cílem dat je v první řadě být znovu použity. Pokud je použije běžný občan, pro svou informovanost, má to také svoji cenu. Může to vést ke zvýšení životní úrovně, obecné informovanosti vedoucí k hlubšímu pochopení problematiky, a nakonec i spokojenosti obyvatel, protože uvidí, že stát, resp. město s nimi jedná jako se sobě rovnými. Ačkoli je tento přístup chvályhodný, stále je ale nakonec pro města primární další užití a zpracování strojem, vedoucí ke tvorbám aplikací či dalším službám založeným na otevřených datech. Tento způsob totiž může pomoci vzniku nové firmy (ve vazbě na OGD zpravidla jako start-up), která bude odvádět daně a zaměstnávat lidi. Je to hlavní ekonomický cíl otevřených dat, kterému je třeba pomoci tím, že případným tvůrcům nebudou házeny klacky pod nohy. Navíc shrnující dokument, který si přečte obyčejný uživatel může sloužit i programátorům pro pochopení dat.

Ve vazbě na OGD a jejich využitelnost ve prospěch měst je dnes nejčastěji diskutován koncept chytrých měst (smart cities), více informací lze nalézt např. v Hawken et al. (2020). Tento proces propojení moderních ICT a měst zahrnuje snahu o udržitelný rozvoj měst, zlepšení

kvality života a zefektivnění správy věcí veřejných. Lidem by mělo přinést kvalitnější život, úspory na energiích, ale i času, bezpečnější a rychlejší dopravu, více přírody ve městech, chytré řešení infrastruktury, rychlou komunikaci s úřady apod. Příkladem jsou např. kamery sledující provoz a reagující na jeho hustotu, svážení odpadů na základě informací ze senzorů atd.

Dalším trendem je základní start-upů, které jsou zpravidla založeny na moderních ICT a často využívají pro svoje činnosti i OGD, která rychle dokážou přetvářet do efektivnějších služeb či produktů, než kdyby ty samé činnosti realizoval veřejný sektor. Pro města je tedy výhodnější mít tyto firmy na svém území a vycházet jim vstříc.

4.1 BEST PRACTICES EU

Důležitý je tedy další rozvoj portálů na základě best practices. Například EU v roce 2020 vydala osvědčené postupy pro otevřená data (Open Data Best Practices in Europe, 2020):

1. Politika otevřených dat – vytvoření legislativy a strategií pro otevřená data. Dále vytvoření plánu implementace, aby byla strategie proveditelná a aby byly jasné odpovědnosti.
2. Portál otevřených dat – jeho spuštění, nabídka široké škály funkcí, a zaměření se na interakci mezi poskytovateli dat a uživateli prostřednictvím diskusních fór, systémů zpětné vazby pro konkrétní datové sady a systémů hodnocení.
3. Účinnost a dopad otevřených dat – monitorování a analýza opětovného použití otevřených dat; zajištění, aby byly ve správný čas k dispozici správné datové sady; komunikace s uživateli, pro pochopení jejich potřeby prostřednictvím interakcí s otevřenými daty.
4. Kvalita otevřených dat – školení vydavatelů a poskytovatelů dat, integrace všech opatření, návodů a možností školení do jedné platformy na portálu.

Všechna doporučení jsou zpravidla praktického rázu. Je to z důvodu, že v našich končinách ještě nejsou otevřená data běžnou součástí našich životů, a ještě jsme se nedostali k přemýšlení o otevřených datech v systémové logice a přístupu k nim. Proto lze každý z předchozích bodů doporučit a vnímat je jako postup pro zvýšení kvality otevřených dat.

Pro ČR je nutné zdůraznit především nutnost většího zapojení veřejnosti, připravit postupy pro získání podrobné zpětné vazby, a více se věnovat integraci jednotlivých systémů, ze kterých lze získávat data (ISVS) a publikovat je jako otevřená, protože v ČR je obecně velký problém s nekompatibilitou a převoditelností formátů dat.

4.2 BEST PRACTICES USA

Spojené státy americké (USA) mají doporučení rozdělená do 9 okruhů (U.S. Open Data Toolkit, 2018):

1. Sběr dat – způsob, jakým jsou data zpočátku shromažďována, může určovat kvalitu a hodnotu dat v průběhu životního cyklu, proto je nutná kontrola kvality na samém začátku.
2. Ochrana osobních údajů – před otevřením dat je nutné provést anonymizaci a deidentifikaci, protože je možné, že kombinací různých datových sad může dojít k identifikaci jednotlivců. Je doporučeno budování důvěry s komunitou ohledně používání dat.
3. Kvalita dat – zlepšení kvality dat vede k velké řadě sociálních, vědeckých, ekonomických a provozních výhod. Vysoce kvalitní data mohou vést k lepším veřejným službám, rozvoji vědeckého výzkumu a nárůstu obchodních příležitostí. Sloužit k tomu má vývoj systémů zpětné vazby od uživatelů k identifikaci problémů s kvalitou. Dalším bodem je využití výzev a soutěží ke zlepšení kvality dat.
4. Vydání dat – lze doporučit uvolnit jak nezpracovaná data, tak vylepšená data s transparentností ohledně přesnosti, kvality a původu. Také hackathony a výzvy mohou zapojit komunitu otevřených dat a další, aby našli nové způsoby, jak data využít.
5. Datové standardy – měly by být aplikovány před sběrem dat, protože pozdější převod na standardy může vést k určité ztrátě kvality dat.
6. Spolupráce mezi organizacemi – zveřejňujícími data i s těmi, kdo je využívají. Mohou to být i mezioborové spolupráce, které povedou k inovativním pracím.
7. Otevřená věda – v USA je to hnutí, jehož cílem je zpřístupnit data shromážděná prostřednictvím vědeckého výzkumu volně všem, osvědčilo se vyžadovat sdílení a zveřejňování dat jako podmínku financování výzkumu a pomáhat výzkumníkům splnit tento požadavek.
8. Spolupráce soukromého sektoru – soukromý sektor a vláda mají společný zájem na zajištění toho, aby vládní datové programy byly vysoce kvalitní, snadno dostupné a nákladově efektivní.

9. Veřejné zapojení – rozvinout komunity praktik kolem otevřených dat, které zahrnují zainteresované strany jak v rámci vlády, tak mimo ni. Webová analýza a demografická analýza mohou také poskytnout pohled na uživatele otevřených dat. Metriky pro hodnocení využívání otevřených dat zároveň mohou pomoci ukázat, jak efektivně jsou data agentur a výzkumníků využívána, a mohou agenturám pomoci upřednostnit hodnotu různých datových sad.

Při pohledu na EU a USA je od pohledu vidět značný rozdíl. Z tohoto důvodu je pro nás doporučení o ochraně osobních údajů reálně zastoupené primárně deidentifikací, spočívající v odebrání údajů, které by mohly daného člověka pomoci identifikovat a nahradit univerzálním identifikátorem. Zároveň si díky těmto doporučením můžeme domyslet, kam mohou teoreticky otevřená data směřovat a na základě toho upravovat už nyní, jak s nimi zacházíme, a už teď si lze domyslet, jaká bude nutná legislativa v budoucnosti. Proto lze doporučit především důraz na kvalitní legislativu okolo otevřených dat.

Dalším doporučením je to, že máme výzkumníky odměňovat či podmiňovat jejich výzkum právě zveřejněním dat. Je zcela představitelná rovina spolupráce na základě otevřených dat na úrovni státu, nicméně na úrovni vědy to otevírá zcela nový potenciál otevřených dat. Pokud by výzkumníci netratili za zveřejňování svých datových sad, můžeme reálně urychlit a zkvalitnit další výzkum, což může vést k velkému posunu v rámci technologií a dalších věd.

Dalším bodem, který se v malém měřítku už uplatňuje i u nás, jsou hackathony. Už fungují na úrovni některých krajů či měst, ale lze doporučit, že by měly více fungovat i na úrovni státu či EU. V tomto ohledu by mohly vzniknout převratné projekty, které z otevřených dat udělají nový atraktivní podobor. Problémem však často bývá přístup k vhodným datům, protože ta často nejsou připravena na tuto formu vytváření aplikací, zejména pokud jde o ČR a její ISVS.

Velký úspěch by mohlo mít i zveřejňování dat v surové formě a těch, které byly upraveny později, tzn. verzování datových sad. Díky čemuž se mohou zveřejňovat rychleji. Navíc zveřejnění těchto neupravených dat může vést k vytvoření standardů, které je možné zavést už při sběru dat. Neměl by být problém data zveřejnit nejdřív v reálném čase a následně je naformátovat a jednou za několik dní, týdnů apod. je zveřejnit v upravené formě. Například právě pro dopravní data by to mohlo být klíčem k tvorbě aplikací, které je využijí.

ZÁVĚR

Pokud někde mezi běžnými smrtelníky zazní pojem otevřená data, resp. open data, málokdo si dokáže představit, co se tím myslí, nebo proč by to mělo existovat. Osobně si myslím, že i díky minulým režimům není naše česká nátura zcela ochotná k otevřeným datům. Také je to práce navíc, a ne úplně levná záležitost. Ale tvářit se kvůli tomu, že jsou otevřená data nepodstatná, nebo se jim vyhýbat není cestou.

Občas se o otevřených datech mluví jako o „nové průmyslové revoluci“. Jenže otevřená data se od obvyklých surovin razantně odlišují. Dají se použít mnohokrát a jejich hodnota se právě znovupoužitím zvyšuje. Neubývá jich a jsou dostupná všem bez rozdílu. Jejich potenciál si zatím dost dobře uvědomuje málokdo, nicméně tím, co otevřená data dokázala už nyní, ve svém počátku, je z nich rozhodně cítit velký potenciál.

Cílem práce bylo popsat charakteristiky zdrojů otevřených dat a analyzovat datové portály na úrovni měst. Za tímto účelem byly nejdříve vymezeny základní pojmy týkající se otevřených dat, jejich historie, potenciálu, metodiky jejich publikace, právní legislativy, a zejména význam datových portálů. Následovala charakteristika zdrojů otevřených dat, která se blíže věnovala existujícím možnostem ukládání a správy otevřených dat, pětihvězdičkovému schématu, formátům otevřených dat, funkcím a službám datových portálů. V této kapitole byly podrobněji popsány a vyhodnoceny vybrané datové portály a aplikace nad otevřenými daty v ČR a EU, což lze považovat za první významný přínos této práce.

Hlavní přínosy lze nalézt v kapitole analyzující vybrané datové portály pro otevřená data na úrovni měst. Pro potřeby analýzy byla zvolena metodika indexu GODI a jeho 11 kritérií, pomocí kterých byly portály ohodnoceny. Pro všechna kritéria byl rovněž zformulován i vlastní postup hodnocení, tzn. přesné určení bodů a co daný portál musí splňovat, aby je získal. Následoval popis vlastního postupu jejich analýzy. Jako vzorek byla zvolena hlavní města států v Evropě. Ze 46 měst nemělo 12 portál vůbec, resp. nebyl nalezen ani národní portál. Dále 13 měst využívalo portál na národní úrovni, tzn., že tam byla zpravidla sekce určená pro hlavní město. Pro 22 hlavních měst byl portál nalezen a po aplikování podmínky dostupnosti anglické verze portálu bylo analyzováno celkem 11 datových portálů.

Tyto portály byly ohodnoceny podle kritérií indexu GODI, sestaveno jejich pořadí a každý portál byl popsán s ohledem na dosažené výsledky. Zdůrazněny byly zejména nedostatky těchto

portálů, aby bylo možné je potenciálně odstranit, případně se jich vyvarovat u jiných portálů. Konkrétním návrhům na jejich odstranění se pak věnovala další podkapitola.

Nakonec byly zformulovány závěry a doporučení pro další rozvoj a vylepšení datových portálů na úrovni měst a regionů. Ta vychází hlavně z výsledků provedené analýzy a jsou dále doplněna o aktuální trendy objevující se v literatuře i v praxi. Jedná se o využívání OGD pro rozvoj tzv. chytrých měst, která jsou založená na moderních ICT a datových tocích, což zveřejňování dat v otevřených formátech plně podporuje. Nakonec byla zmíněna i problematika start-upů a to, jak souvisí s otevřenými daty.

Závěrem lze zmínit, že přestože 11 ohodnocených portálů dosáhlo velmi dobrých výsledků, tak velké množství hlavních měst v Evropě buď svůj portál vůbec neprovozuje nebo využívá služeb národního portálu. Důvody těchto zjištění by tak měly být prozkoumány v další práci.

POUŽITÉ ZDROJE

BERNERS-LEE, Tim. Linked Data. W3C [online]. 2006 [cit. 2022-04-27]. Dostupné z: <https://www.w3.org/DesignIssues/LinkedData.html>

BOČEK, Jan, Jakub MRÁČEK a Jindřich MYNARZ. *Otevřená data: příležitost pro Českou republiku*. Praha: Fond Otakara Motejla, 2012, 20 s. ISBN 978-80-87725-03-0.

DAVIES, Tim, Stephen B. WALKER, Mor RUBINSTEIN a Fernando PERINI (Eds.). *The State of Open Data: Histories and Horizons*. Cape Town and Ottawa: African Minds and International Development Research Centre. 2019, 574 s. ISBN 978-1-55250-612-7.

Datový portál Královéhradeckého kraje [online]. Královéhradecký kraj, 2022 [cit. 2022-04-27]. Dostupné z: <https://www.datakhk.cz/>

Datový portál města Brna [online]. Statutární město Brno, 2021 [cit. 2022-04-27]. Dostupné z: <https://data.brno.cz/>

HARARI, Yuval N. *Sapiens: úchvatný i úděsný příběh lidstva*. Voznice: Leda, 2014. 532 s. ISBN 978-80-7335-388-9.

HAWKEN Scott, Hoon HAN a Chris PETTIT (Eds.). (2020). *Open Cities | Open Data: Collaborative Cities in the Information Era*. Singapore: Palgrave Macmillan, 2020, 418 s. ISBN 978-981-13-6605-5.

CHARALABIDIS, Yannis, Charalampos ALEXOPOULOS, Enrico FERRO, Marijn JANSSEN, Thomas LAMPOLTSHAMMER a Anneke ZUIDERWIJK. *The World of Open Data: Concepts, Methods, Tools and Experiences*. Cham: Springer, 2018, 229 s. ISBN 978-3-319-90850-2.

KLÍMEK, Jakub. Otevřená data a otevřené formální normy. *Portál otevřených dat* [online]. 2020 [cit. 2022-04-27]. Dostupné z: <https://data.gov.cz/články/otevřené-formální-normy-01-úvod>

KLÍMEK, Jakub. Národní katalog otevřených dat - výchozí bod práce s daty. *Portál otevřených dat* [online]. 2021 [cit. 2022-04-27]. Dostupné z: <https://data.gov.cz/články/národní-katalog-otevřených-dat-výchozí-bod-práce-s-daty>

KOMÁRKOVÁ Kateřina. *Otevřená data a jejich praktické využití*. Pardubice, 2017. 45 s. Bakalářská práce. Univerzita Pardubice, Fakulta ekonomicko-správní. Vedoucí práce Ing. Simona Pichová.

LNĚNIČKA, Martin, Renáta MÁCHOVÁ, Jolana VOLEJNÍKOVÁ, Veronika LINHARTOVÁ, Radka KNĚZÁČKOVÁ a Miloslav HUB. Enhancing transparency through open government data: The case of data portals and their features and capabilities. *Online Information Review*, 2021, 45(6), 1021-1038.

MÁCHOVÁ, Renáta a Martin LNĚNIČKA. Evaluating the quality of open data portals on the national level. *Journal of Theoretical and Applied Electronic Commerce Research*, 2017, 12(1), 21-41.

Methodology. Global Open Data Index [online]. UK: Open Knowledge Foundation, 2016 [cit. 2022-04-27]. Dostupné z: <https://index.okfn.org/methodology/>

Ministerstvo vnitra ČR. *Standardy publikace a katalogizace otevřených dat VS ČR* [online]. 2015 [cit. 2022-04-27]. Dostupné z: https://opendata.gov.cz/media/standardy_publicace_a_katalogizace_otevrenych_dat_vs_cr.pdf

MRÁČEK, Jakub, Jan BOČEK a Jáchym ČEPICKÝ. *Jak otevírat data?*. Praha: Fond Otakara Motejla, 2014, 113 s. ISBN 978-80-87725-15-3.

MYŠKA, Matěj, Radim POLČÁK, Libor KYNCL, Jaromír ŠAVELKA a Iveta SVIRÁKOVÁ. *Veřejné licence v České republice*. Brno: Masarykova univerzita, 2014, 189 s. ISBN 978-80-210-7193-3.

Oficiální portál evropských dat [online]. EU: Úřad pro publikace EU, 2022 [cit. 2022-04-27]. Dostupné z: <https://data.europa.eu/cs>

Open Data Best Practices in Europe. *Data Europa* [online]. EU: EU, 2020 [cit. 2022-04-29]. Dostupné z: <https://data.europa.eu/en/datastories/open-data-best-practices-europe>

Open Definition [online]. UK: Open Knowledge Foundation, 2014 [cit. 2022-04-27]. Dostupné z: <http://opendefinition.org/od/2.0/cz/>

Portál otevřených dat ČR [online]. Ministerstvo vnitra ČR, 2022 [cit. 2022-04-27]. Dostupné z: <https://data.gov.cz/>

Ten principles for opening up government information. Sunlight Foundation [online]. Washington, DC: The Sunlight Foundation, 2017 [cit. 2022-04-27]. Dostupné z: <https://sunlightfoundation.com/policy/documents/ten-open-data-principles/>

THORSBY, Jeffrey, Genie N.L. STOWERS, Kristen WOLSLEGEL a Ellie TUMBUAN. Understanding the content and features of open data portals in American cities. *Government Information Quarterly*, 2017, 34(1), 53-61.

U.S. Open Data Toolkit [online]. Washington, DC: The Center for Open Data Enterprise, 2018 [cit. 2022-04-30]. Dostupné z: <https://www.usopendat toolkit.org/best-practices-1>

VAN HESTEREN, Daphne a Laura VAN KNIPPENBERG. *Open Data Maturity Report 2021* [online]. Luxembourg: Publications Office of the European Union, 2021 [cit. 2022-04-27]. ISBN 978-92-78-42712-2. Dostupné z: https://data.europa.eu/sites/default/files/landscaping_insight_report_n7_2021.pdf

Zákonná definice otevřených dat [online]. ČR: Ministerstvo vnitra České republiky, 2020 [cit. 2022-04-27]. Dostupné z: https://opendata.gov.cz/definition_opendata