

## PŘÍNOSY OPEN DAT, SOUVISEJÍCÍCH PORTÁLŮ A APLIKACÍ PRO VEŘEJNÝ SEKTOR

### THE BENEFITS OF OPEN DATA, RELATED PORTALS AND APPLICATIONS FOR THE PUBLIC SECTOR

Renáta Máchová, Martin Lněnička

**Abstrakt:** *Ekonomika řízená daty slibuje vytvoření nových ekonomických aktivit, přínosů a možností pro další růst. Větší dostupnost těchto dat také nabízí nové možnosti pro vývoj moderních služeb, ale rovněž otevírá nové výzvy jako je rostoucí objem příležitostí pro publikaci datových sad a jejich integraci v rámci vysoce škálovatelných ekosystémů veřejných dat. V současnosti je tento problém řešen tvorbou open data portálů a iniciativ, které mají za cíl lépe pochopit význam open, big a linked dat a zhodnotit jejich přínosy, včetně definování návrhů jak je dále zvýšit. Díky tomu pak zvýšená dostupnost těchto dat dále akceleruje jejich znovu využívání tak, že mohou být používána pro vývoj webových a mobilních aplikací, které vytvářejí požadovanou přidanou hodnotu, nabízí nový pohled na data, poskytují přístup ke službám e-governmentu a zvyšují transparentnost. Proto je hlavním cílem tohoto příspěvku vymezit význam open, big a linked pro veřejný sektor a rozvoj e-governmentu. Dále se blíže soustředit na problematiku aplikací vyvíjených nad těmito daty a za účelem zjištění informací o tom, jak by se měla tato oblast dále zlepšovat a zejména jaké aplikace by mohly mít největší potenciál, provést mezi cílovou skupinou vysokoškolských studentů průzkum, který na základě jejich odpovědí a názorů poskytne podklady pro další směřování výzkumu.*

**Klíčová slova:** *open data, big data, open data portály, aplikace, veřejný sektor, průzkum*

**Abstract:** *The data-driven economy promises the creation of enormous amounts of economic activities, benefits and growth opportunities. The availability of new data raises opportunities for the development of modern services, but also open new challenging issues such as the increasing volume of datasets publication available and their*

*integration in large-scale public data ecosystems. Currently, this problem is solved by the development of open data portals, in order to better understand of open, big and linked data, and assess the various types of value they generate, and identify the required improvements for increasing their value. The greater availability of these data then catalyses reuse of them, so that they can be used for building useful web or mobile applications which leverage their benefits, allow insight, provide access to e-government services and support transparency. Therefore, the main of this paper is to define the importance and benefits of open, big and linked data for the public sector and the e-government development. Furthermore, take a closer focus on the issue of applications developed over these data and in order to provide information for the improvement of open data supply and the potential of these applications, to conduct a survey among the target group of Czech university students, which, based on their answers and also perceptions, may provide the basis for the future direction of research.*

**Keywords:** *open data, big data, open data portals, applications, public sector, survey*

**JEL Classification:** C 42, H 83, L 86

## 1 ÚVOD

E-government, jeho vývoj a přístupy k jeho měření se v průběhu let vyvíjí s ohledem na nové informační a komunikační technologie (ICT) a jejich využití při komunikaci, výměně a také získávání informací mezi jednotlivými subjekty. Veřejná správa a přístup k jejímu řízení dnes vychází hlavně z pojmů New Public Management a Open Government (OG), kdy oba směřují k trvalému zvyšování kvality svých služeb pro občany a podniky, větší nákladové efektivnosti a obecně otevřenosti veřejné správy. Mezi hlavní oblasti ovlivňující tento vývoj a související požadavky pak patří open data, big data, cloud computing, sociální média, one-stop portály, multikanálová strategie, vysokorychlostní připojení k internetu nebo chytré mobilní telefony.

V tomto příspěvku bude kladen důraz na význam open a big dat v kontextu OG jako nástroje pro zvyšování transparentnosti, podpory spolupráce mezi různými zúčastněnými stranami, podpory využívání Informací veřejné sektoru (Public Sector Information, PSI), zapojování občanů do návrhů a implementace různých strategií, které dále ovlivňují rozhodovací procesy,

a především vytváření různých aplikací nad těmito daty, které budou podporovat tyto procesy. Tento trend je pak součástí snahy o vytvoření ekonomiky řízené daty (Buchholtz et al., 2014; Foulonneau et al., 2014). Technologickou základnu pro tyto cíle tvoří e-government a jeho infrastruktura. Centrálním bodem jsou zde open data portály, kde instituce veřejného sektoru zpřístupňují PSI podle definic open dat a OG (Geiger a von Lucke, 2012; Kalampokis et al., 2013). Data veřejného sektoru se nazývají Open Government Data (OGD), když dále existují i Open Business Data (OBD), Open Citizen Data (OCD) a Open Science Data (OSD). S nimi dále úzce souvisí Linked Open Data (LOD), což jsou vysoce strukturovaná a propojitelná data splňující definované požadavky (Ermilov et al., 2013). Zveřejněním těchto dat je instituce veřejného sektoru vracejí zpět k občanům a podnikům, kteří nepřímo zaplatili jejich vytvoření ze svých daní (Kostovski et al., 2012). Díky rozvoji ICT a růstu kapacit pro ukládání a přenos dat však roste i jejich objem, který je institucemi veřejného sektoru vytvářen. Tato data se pak nazývají big data, když v kontextu studované problematiky jde o open (big) data (Cowan et al., 2014; Marton et al., 2013; Tsiavos et al., 2012). Marton et al. (2013) tvrdí, že základní koncepty open a big dat jsou stejné v tom, že jsou vytvářeny za určitým účelem a dále znovu využívány. Pro získání přidané hodnoty, resp. generování přínosů jsou pak zásadní kombinace těchto dat na různých úrovních, které umožňují vytvářet nové modely a aplikace (Cowan et al., 2014; Kalampokis et al., 2013). Ty jsou využívány pro zpracování dat novými metodami a algoritmy, např. z oblasti big data analytics (Elgendy a Elragal, 2014; World Bank, 2014).

Hansen et al. (2013) odhadují, že objemy dat produkované evropskými institucemi mají pro trh hodnotu v řádech desítek miliard Eur. Také Buchholtz et al. (2014) odhadují, že přímé a nepřímé agregované ekonomické dopady open (big) dat pro ekonomiku celé Evropské Unie (EU) jsou v řádech miliard Eur ročně. Zpřístupnění těchto dat veřejnosti však nutně nemusí znamenat zvýšení transparentnosti a efektivity veřejného sektoru, lepší občanskou kontrolu nebo větší zapojení občanů do rozhodovacích procesů (Lourenço, 2013). Kromě zhodnocení ekonomického významu open (big) dat existují i další problémy, které je nutné řešit, např. kde a jaká data zveřejnit, v jakých formátech, jak o tom informovat veřejnost, atd. (Van der Waal et al., 2014). Návržnost investic do tvorby souvisejících strategií, spouštění portálů a obecně nákladů na zveřejnění těchto dat by měla být hodnocena na základě aplikací vytvořených nad těmito daty (Foulonneau et al., 2014).

Této problematice se věnují i mezinárodní instituce. Organizace spojených národů (OSN) ve svých posledních zprávách z roku 2012 (United Nations, 2012) a 2014 (United Nations, 2014) zdůrazňuje význam OGD a jejich přínosy v kontextu podpory udržitelného rozvoje veřejného sektoru. Také Mezinárodní telekomunikační unie (International Telecommunication Union, ITU), Světová banka (World Bank, WB) nebo Světové ekonomické fórum (World Economic Forum, WEF) se ve svých aktuálních zprávách blíže věnují této problematice (International Telecommunication Union, 2014; World Bank, 2014; World Economic Forum, 2014), když nejdále je WB, která založila pracovní skupinu pro OGD a představila metodiku s postupy a doporučeními hodnotící připravenost států využívat tato data.

## 2 VYMEZENÍ PROBLEMATIKY, CÍL A POUŽITÉ METODY

V současné době lze sledovat stále rostoucí poptávku ze strany občanů o informace, které jsou v držení institucí veřejného sektoru. Souběžně vlády a další instituce veřejného sektoru stále více experimentují se zapojováním občanů do rozvojových aktivit, programů a komunitních partnerství s využitím nových komunikačních kanálů, zejména prostřednictvím sociálních médií. Během posledních několika let vzniklo velké množství výzkumných prací a studií, kde byly tyto aktivity zkoumány v kontextu e-governmentu a souvisejících přínosů a rizik. Těmto výzkumným otázkám se věnují zejména mezinárodní instituce jako je OSN (United Nations, 2014), WB (World Bank, 2014) nebo WEF (International Telecommunication Union, 2014). Zde je nutné zmínit to, že tyto instituce řeší danou oblast spíše obecně a definují především směry vývoje ve vyspělých státech. Konkrétní výzkumné otázky lze nalézt např. v Cowan et al. (2014), kde autoři řeší příležitosti pro zvyšování transparentnosti služeb e-governmentu. Nicméně i když se věnují některým vybraným aplikacím, tak vůbec neřeší zdroje dat a jejich možné kombinace a integrace. To je řešeno např. v Yang a Kankanhalli (2013), kde ale zase naopak chybí bližší specifiky konkrétního využití a jejich funkce. Tyto rozpory proto budou v tomto příspěvku dále diskutovány a navrženo jejich řešení.

Jádrem těchto přístupů je tak schopnost open dat, participativních nástrojů a interoperabilních platforem umožnit rychlou, efektivní a také kvalitní tvorbu a poskytování veřejných statků a služeb. Vše zaštiťují principy OG, kdy vlády po celém světě očekávají, že open data budou stimulovat významné sociální, ekonomické a environmentální přínosy pro společnost (Jetzek et al., 2014).

Nicméně příslib ekonomiky řízené daty leží především ve vytváření nových služeb nad těmito daty. Výzkumy dále ukázaly, že občané stále více využívají různé webové nebo mobilní aplikace pro rychlé nalezení informací a obecně podporu rozhodování. Verma a Gupta (2012) uvádějí, že více jak 50 % občanů již komunikovalo s nějakou veřejnou institucí prostřednictvím svého mobilního telefonu. V této oblasti je však stále mnoho nezodpovězených otázek. Kromě nevědomosti uživatelů o tom, že daná aplikace využívá open data, je pak hlavním problémem samotná tvorba aplikací, resp. shoda na tom, kdo by měl dané aplikace aktivně vytvářet a jaké zaměření by tyto aplikace měly mít. Většina vlád a institucí veřejného sektoru dnes spoléhá na to, že tyto aplikace a analýzy nad open daty bude vytvářet nějaký prostředník (Cranefield et al., 2014).

Proto je v rámci tohoto příspěvku představen průzkum, který nabídne odpověď na tyto otázky. Hlavním cílem je nejprve detailně popsat problematiku open, linked dat, datových portálů a s nimi souvisejících aplikací. Dále se blíže zaměřit na oblast aplikací vytvářených nad open daty, možností jejich propojování a integrace za účelem zvyšování přidané hodnoty těchto dat pro všechny zúčastněné strany. Za tímto účelem byl realizován průzkum, kde byli za cílovou skupinu zvoleni vysokoškolští studenti a zjišťovány jejich názory na tuto problematiku a její řešení. Cílem je sumarizovat přínosy, rizika a nové výzvy těchto přístupů tak, aby institucím veřejného sektoru mohly pomoci při implementaci těchto technologií pro podporu rozvoje e-governmentu a aby jejich data byla prezentována efektivně, kvalitně a transparentně. Jak bylo zmíněno již výše, tento příspěvek se tedy snaží konkretizovat tuto problematiku pro praktické využití a zároveň řešit rozpory, které se objevují v některých zmíněných výzkumných pracích a studiích. Definovaný cíl je pak řešen s využitím metody dotazníkového šetření, analýzy souvisejících literárních zdrojů a porovnání kritérií za účelem určení přínosů a rizik tohoto řešení pro veřejný sektor.

### **3 LITERÁRNÍ REŠERŠE A POZADÍ ŘEŠENÉHO PROBLÉMU**

#### **3.1 Open, big a linked data, výhody, nevýhody a výzvy**

Z důvodu stále rostoucích objemů dat produkovaných institucemi veřejného sektoru se model open dat vyvinul do modelu open (big) dat (Marton et al., 2013; United Nations, 2012). Open data jsou především o standardech pro zajištění toho, aby tato data byla zdarma pro každého, strojově čitelná

a vzájemně propojitelná, tzv. linked data nebo LOD. Kontext big dat pak spíše popisuje proces získávání, ukládání a analýzy těchto dat s cílem získat informace o hlavních trendech a vývoji ve sledované oblasti, resp. jde o distribuované architektury a jejich využití pro zpracování těchto dat (Chen et al., 2014; Marton et al., 2013). Linked data pak popisují způsob publikace vysoce strukturovaných dat, která mohou být vzájemně propojena a stát se použitelnější využíváním sémantických dotazů. Cílem je umožnit propojení a interakce mezi různými zdroji dat (Ermilov et al., 2013, Heath a Bizer, 2011). Přestože je dnes význam big data již uznáván, názory na jejich definici se stále liší. Obecně se jedná o data v takových objemech, že již není možné je zpracovávat běžnými nástroji a metodami v rozumném čase (Chen et al., 2014). Tyto změny jsou charakterizovány jako 3V (velocity, volume a variety), tzn. rychlost (data neustále přicházejí), objem (v řádově GB až TB) a rozmanitost (v různě strukturované a nestrukturované formě). Někteří autoři přidávají také další V (value, veracity a variability), tzn. hodnota, správnost (pravdivost) a proměnlivost (Demchenko et al., 2013).

Celý ekosystém kolem těchto dat netvoří jen samotná data, ale zahrnuje také vizi, vedení, právní rámec, instituce, infrastrukturu a komunity uživatelů (World Bank, 2014). Samotný životní cyklus těchto dat se podle Zuiderwijk a Janssen (2013) skládá ze všech aktivit mezi okamžikem, kdy začíná vytváření dat, až po okamžik, kdy jsou tato data diskutována, včetně aktivit pro jejich zveřejnění, nalezení a další využívání. Proto by měly být vytvářeny modely a strategie řešící zapojování jednotlivých subjektů do tohoto cyklu, kde jim bude umožněno vést oboustranný dialog a realizovat své požadavky (Geiger a von Lucke, 2012). Solar et al. (2012) navrhli hierarchický model tvořený oblastmi, podoblastmi a kritickými ukazateli, který má sloužit pro hodnocení stavu implementace open dat v souladu s jejich principy a zásadami. Hyland a Wood (2011) představili model skládající se ze 6 kroků, který obsahuje pokyny pro identifikaci, popis, převod, publikování a správu open dat. Gurstein (2011) také navrhl model skládající se ze 7 vrstev pro podporu efektivního využívání open dat a navazujících datových struktur. Hlavní problémy s těmito daty souvisí zejména s jejich získáváním, vyhledáváním, ukládáním, sdílením, analýzou a zpřístupněním (Elgendy a Elragal, 2014). To má za následek to, že je nutné vyvíjet nové technologie a platformy pro efektivní práci s těmito daty (Chen et al., 2014; Tien, 2013). Hlavní výzvou jsou datově náročné procesy řízení zpracování těchto dat a návrh nových algoritmů pro vyhledávání v těchto datech. Již byly realizovány různé studie, které zkoumaly paralelní architektury

pro zpracování dat jako např. Che et al. (2013), Demchenko et al. (2013) nebo Pavlo et al. (2009).

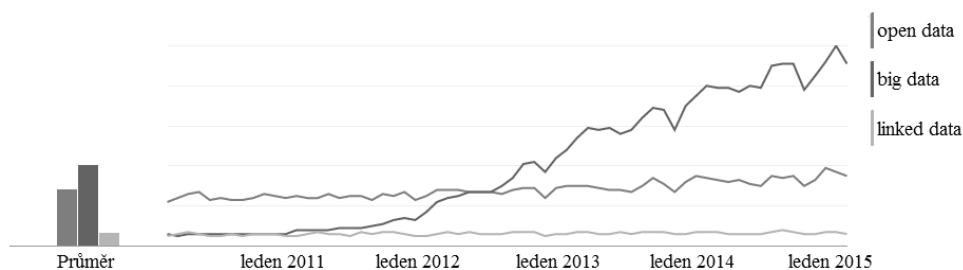
Většina literární zdrojů věnujících se znovu využívání těchto dat se zabývá jejich potenciálem a ekonomickou hodnotou (Geiger a von Lucke, 2012), zatímco literatura řešící iniciativy OG je cílena do oblasti vládních politik a legislativního prostředí, které má přispět k vytváření sociálního přínosu prostřednictvím spolupráce různých subjektů (Jetzek et al., 2013). Jak bylo zmíněno výše, zájem o koncept open dat je zde již několik let (Janssen et al., 2012) a dále se rozvíjí díky tlaku na zvyšování transparentnosti veřejné správy a požadavkům na analýzu big dat (Chen et al., 2014). Kucera a Chlapek (2014) definovali seznamy přínosů, které mohou být dosaženy zveřejněním OGD, a dále skupiny rizik, které je nutné hodnotit před tím, než jsou tato data zveřejněna. Cowan et al. (2014) také představili několik praktických příkladů za účelem ilustrace souvisejících rizik a možností open dat.

V návaznosti na principy open dat, pak jednotlivé státy začaly propagovat OG, vytvářet různé legislativní rámce a na ně navazující elektronické služby, především v souvislosti s portály, kde jsou tato data zveřejňována (United Nations, 2014). Tsiavos et al. (2012) zmiňují různé bariéry a konflikty při implementaci těchto rámců a proto zdůrazňují využívání open source nástrojů a otevřených architektur, které jsou zároveň zásadní pro snižování nákladů. Různí autoři zároveň potvrdili, že zveřejnění dat veřejného sektoru v otevřených formátech vytváří značné výhody pro občany, podniky, výzkumné pracovníky a dalšími zúčastněné subjekty za účelem pochopení a pohledu na fungování a rozhodování ve veřejné správě. To vše jak pak možné díky existenci pokročilých technologií pro zpracování těchto dat (Janssen et al., 2012). Kalampokis et al. (2013) tvrdí, že reálná hodnota a přínos OGD bude odhalen jen na základě analýz nad těmito daty, které jsou kombinací datový sad dříve uzavřených v oddělených databázích a datových skladech. Tyto zdroje dnes mohou být díky principům open a linked dat propojeny a zveřejněny za účelem získání přidané hodnoty, což povede k zefektivnění procesů, zrychlení rozhodování a především úsporám zejména investičních nákladů. Autoři navrhli koncept pro zpracování těchto dat včetně technických požadavků, který může být později implementován např. s využitím platformy Apache Hadoop. Hlavním trendem je zde především ekonomický potenciál těchto dat a jeho hodnocení (Yang a Kankanhalli, 2013). Jetzek et al. (2013) navrhli model zachycující možnosti přeměny open dat jako ekonomického zdroje do přidané hodnoty. Zároveň

tvrdí, že když jsou jednotlivé vztahy dobře pochopeny, tak je jednodušší nalézt potřebnou datovou sadou a realizovat postupy nutné pro získání přidané hodnoty. V roce 2014 model aktualizovali se zaměřením na generování udržitelné hodnoty ve sdílejší společnosti (Jetzek et al., 2014). V něm zdůrazňují vyšší míru spolupráce mezi veřejným a soukromým sektorem, především jejich motivaci, možnosti a schopnosti získat přidanou hodnotu z těchto dat. Geiger a von Lucke (2012) rovněž analyzovali přidanou hodnotu volně přístupných dat veřejné správy a diskutovali výzvy OGD pro veřejnou správu na jejich různých úrovních.

Trendy ve vyhledávání souvisejících pojmů open, big a linked data na Googlu od roku 2010 doposud jsou pomocí Trendů Google zachyceny na Obr. 1. Zatímco pojem open data roste mírně, o pojem big data je především v posledních několika letech velký zájem. Relativní stagnace pojmu linked data je způsobena tím, že tento pojem je další fází ve zveřejňování open dat, tudíž je nejdříve nutné pracovat se znalostí pojmu open data. Jak je u obrázku dále patrné, tak pojem open data roste pouze mírně, proto je nutné nadále zvyšovat povědomí o těchto datech a zapojovat různé subjekty do jejich využívání.

**Obr. 1: Vývoj vyhledávání sledovaných pojmů mezi roky 2010–2015.**



Zdroj: autoři.

### 3.2 Open a big data ve vývoji e-governmentu

Poslední dvě desetiletí byly ve světě ve znamení výrazného růstu odvětví ICT, související infrastruktury a služeb. Zároveň se vyvíjely i rámce pro hodnocení rozvoje e-governmentu. Počátky hodnocení a porovnávání e-governmentu ve vybraných státech světa lze zasadit do období konce 90. let minulého století a začátku nového tisíciletí. Na počátku se hodnotily počty televizorů nebo minimálně existence národního portálu. Postupně se dostaly do popředí pevné telefonní linky a vytáčené připojení, které byly postupně nahrazeny



mobilními telefony a bezdrátovým připojením k internetu. Zároveň s dostupností se zvyšovala i rychlost tohoto připojení. Další změny lze najít mezi roky 2008 a 2012 v návaznosti na globální ekonomickou krizi. V posledních letech se do popředí začaly dostávat nové trendy a technologie jako cloud computing, open data, big data, sociální média, vyšší míra zapojení občanů do rozhodovacích procesů, zelené vládnutí nebo úspory energie (International Telecommunication Union, 2014; United Nations, 2014; World Bank, 2014; World Economic Forum, 2014).

V posledních letech se pak vedle tradičních indexů jako E-Government Development Index (EGDI) a E-Participation Index (EPI) od OSN, Economist Intelligence Unit (EIU) index, ICT Development Index (IDI) od ITU, Networked Readiness Index (NRI), který je vydáván WEF nebo Waseda index, objevily i nové specializované indexy jako je např. Business Software Alliance (BSA) Global Cloud Computing Scorecard, Cloud Readiness Index (CRI), který vydává Asia Cloud Computing Association (ACCA), Open Data Barometer (ODB) a Web Index od World Wide Web Foundation (W3F) nebo také Global Open Data Index (GODI) od Open Knowledge Foundation (OKF).

Mezi základní globální rámce pro open data patří ODB a GODI. ODB a související zprávy mají za cíl odhalit skutečnou dostupnost a dopad open data iniciativ. Index analyzuje globální trendy a poskytuje srovnávací údaje o státech prostřednictvím metodiky kombinující primární data a sekundární ukazatele hodnotící různé dimenze připravenosti, implementace a dopadu open dat (World Wide Web Foundation, 2015). Každá z těchto dimenzí je samostatným sub-indexem, když připravenost s váhou 1/4 se hodnotí samostatně pro instituce veřejného sektoru, podniky a občany. Implementace s váhou 1/2 se skládá také ze tří částí: odpovědnost, inovace a sociální politika. Dopad s váhou 1/4 je pak hodnocen jako politický, ekonomický a sociální. První zpráva z roku 2013 hodnotila 77 států, druhá z roku 2014 celkem 86 států. Každý stát je zařazen do jedné ze 4 skupin: vysoká kapacita, rostoucí a pokročilý, omezená kapacita a jednostranné iniciativy. Podrobnosti lze pak nalézt v World Wide Web Foundation (2015). GODI hodnotí stav v oblasti dostupnosti a využívání OGD ve vybraných státech světa za účelem iniciovat diskusi mezi občany a institucemi veřejného sektoru (Open Knowledge, 2014). Tento index byl poprvé představen v roce 2013 a hodnotil 60 států. V roce 2014 byl jejich počet zvýšen na 97, když byla hodnocena dostupnost těchto 10 klíčových datových sad: výsledky voleb, obchodní rejstřík, prostorová data jako mapa státu, vládní výdaje, národní

statistiky, vládní rozpočet, legislativa, poštovní směrovací čísla, jízdní řády a nakonec emise znečišťujících látek. Každá z těchto datových sad je ohodnocena na základě 6 technických a 3 legislativních otázek s různou vahou. Více lze nalézt v Open Knowledge (2014). Jak bylo zmíněno již výše, tak i EU se zabývá hodnocením oblasti PSI. Proto za tímto účelem vytvořila PSI Scoreboard (PSIS), kde jsou hodnoceny všechny členské státy EU v oblasti open dat a znovu využívání PSI. Je hodnoceno 7 oblastí, které se dále skládají z celkem 25 indikátorů. V Tabulce 1 je zachycen přehled pořadí členských států EU na základě vybraných indexů. Rozmezí hodnot pro ODB je 0–100, pro GODI je 0–100 a pro PSIS je to 0–700. Data pochází z roku 2014. Nejlépe jsou v této oblasti hodnoceny Spojené království a Francie.

**Tab. 1: Porovnání členských států EU na základě vybraných indexů pro open data.**

Členský stát EU	ODB			GODI			PSIS	
	Hodnota	Pořadí		Hodnota	Pořadí		Pořadí	
		Globálně	V rámci EU		Globálně	V rámci EU	Hodnota	V rámci EU
Belgie	47,3	27	13	39	53	22	320	15
Bulharsko	N/A	N/A	N/A	41	51	20	215	28
Česká republika	58,1	17	11	66	12	6	290	19
Dánsko	70,1	9	5	83	2	2	440	11
Estonsko	60,2	13	8	N/A	N/A	N/A	480	10
Finsko	66,5	12	7	73	4	4	395	13
Francie	80,2	4	3	80	3	3	535	3
Chorvatsko	N/A	N/A	N/A	41	51	20	230	27
Irsko	40,7	32	16	48	36	16	485	8
Itálie	50,6	23	12	55	25	11	485	8
Kypr	N/A	N/A	N/A	21	87	26	265	23
Litva	N/A	N/A	N/A	32	70	25	250	24
Lotyšsko	N/A	N/A	N/A	51	34	15	315	16
Lucembursko	N/A	N/A	N/A	N/A	N/A	N/A	235	26
Maďarsko	38,3	33	17	48	36	16	300	18
Malta	N/A	N/A	N/A	52	31	13	245	25
Německo	67,6	11	6	69	9	5	400	12
Nizozemsko	75,8	6	4	64	16	8	505	4
Polsko	37	35	18	42	48	19	490	7
Portugalsko	46,1	29	14	47	39	18	360	14
Rakousko	58,5	16	10	59	22	10	500	5

Rumunsko	N/A	N/A	N/A	64	16	8	270	21
Řecko	40,8	31	15	38	54	23	500	5
Slovensko	N/A	N/A	N/A	35	61	24	280	20
Slovinsko	N/A	N/A	N/A	54	26	12	265	22
Spojené království	100	1	1	97	1	1	585	1
Španělsko	59,9	14	9	52	31	13	550	2
Švédsko	83,7	3	2	66	12	6	305	17

Zdroj: autoři.

E-government byl od počátku svého vývoje vždy implementován jako model zaměřený na služby, které instituce veřejného sektoru poskytují. V posledních letech se tento model změnil a je definován hlavně požadavky občanů v kontextu OG, kdy služby poskytované institucemi veřejného sektoru jsou navrhovány a implementovány pro občany (jednotlivce nebo i právní subjekty) jako uživatele těchto služeb, zatímco požadavky veřejného sektoru ustupují spíše do pozadí (Solar et al., 2012). Tyto nové služby musí splňovat různé podmínky (best practices) jako definuje např. Traunmüller (2014): přehledný design, jednoduchost ovládání, dodržování standardů použitelnosti, multikanálová podpora, interoperabilita, apod. Služby e-governmentu však i nadále směřují k úsporám nákladů, když zveřejňování open a big dat má potenciál tyto náklady dále snižovat a navíc získat přidanou hodnotu (Cranefield et al., 2014). Krishnan et al. (2013) pak navrhli model, který je zaměřen na vztahy mezi technologicko-organizačními faktory a úrovní e-governmentu. Jejich výsledky ukazují, že ICT infrastruktura, lidský kapitál a e-participace mají přímý vztah s úrovní e-governmentu. Machado a Parente de Oliveira (2011) prezentovali open data architekturu pro e-government, která umožňuje publikaci dat ve strukturovaných a strojově čitelných formátech s cílem umožnit občanům kombinovat tato data s dalšími zdroji a vytvářet nad nimi různé aplikace. Kalampokis et al. (2011) porovnali existující modely hodnotící úroveň e-governmentu a na základě jejich dekompozice navrhli model pro hodnocení úrovně implementace OGD s dalšími druhy dat (OBD, OCD, OSD), který má za cíl zvýšit přidanou hodnotu pro jejich uživatele. Höchtl a Reichstädter (2011) tvrdí, že zveřejňování OGD je dnes jednou z hlavních iniciativ e-governmentu a tato data by měla být považována za klíčový prvek strategického řízení informací ve veřejné správě. Dále zdůrazňují význam LOD a návrh modelů architektur, které by umožnily jejich šíření.

### 3.3 Open data portály, jejich vymezení a přínosy

Webové stránky jsou základní a nejjednodušší forma komunikace veřejného sektoru s dalšími subjekty v digitálním světě. Velké množství institucí veřejného sektoru tak provozuje vlastní portály, kde zveřejňují informace. Nicméně pouze málo z nich má vazby na open data nebo nabízí interaktivní prvky, které umožňují od uživatelů získávat zpětnou vazbu.

Prvním problémem, který je tedy nutné při práci s open a big data vyřešit, je jejich nalezení, a to v požadovaném zaměření (kategorii), formátu, aktuálnosti a dalšími atributy (metadaty), kterými jsou tato data definována (Kučera et al., 2013). V posledních několika letech vzniklo v souvislosti s prosazováním myšlenky OG rovněž velké množství datových portálů, kde lze prohledávat jejich datové katalogy, stahovat nebo nahrávat data, a využívat různé další služby včetně Application Programming Interface (API) nebo již existujících aplikací vytvořených nad těmito daty (Ubaldi, 2013; Van der Waal et al., 2014). Tyto portály jsou rozhraním mezi daty veřejné správy na jedné straně a jejich uživateli na straně druhé, a proto se řadí do datové infrastruktury veřejné správy (Verma a Gupta, 2012). Open data portály se zpravidla skládají z katalogů, které jsou kolekcí katalogových záznamů a metadat pro kolekce datových sad. Jsou spravovány buď pověřenou institucí, nevládní organizací nebo občanskou iniciativou. Každý portál nabízí různé datové sady členěné podle kategorií, institucí veřejného sektoru nebo jen na základě generovaných klíčových slov (Kučera et al., 2013). Cílem těchto portálů je nabídnout takové funkce, které může uživatel využívat ke snadnému přístupu, zobrazení a stažení dat (Foulonneau et al., 2014). Některé portály nabízejí i další funkce jako API nebo SPARQL, seznam aplikací vytvořených s využitím dané datové sady, diskuze a hodnocení datových sad nebo nástroje pro zobrazení dat. Důležitá je také možnost požádat o zveřejnění nové datové sady prostřednictvím formuláře (Ubaldi, 2013). Samotná data pak často nejsou přímou součástí portálu, a proto je na portálu vždy odkaz na jejich stažení včetně URL, kde mohou být data stažena. Každá datová sada tak může obsahovat více těchto odkazů (Cyganiak a Maali, 2015; Van der Waal et al., 2014).

První open data portály na národní úrovni byly spuštěny v roce 2009 (Spojené státy americké, USA) a v roce 2010 (Spojené království). V roce 2015 je jich již téměř 70 (Lněnička, 2015), když jejich kvalita je velmi rozdílná, a proto byl např. ve Lněnička (2015) navržen rámec pro jejich hodnocení. Kvalita open dat a obecně portálů, na kterých jsou tato data dostupná, je po fázi nalezení dat tou nejdůležitější. Kučera et al. (2013) diskutovali

některá omezení spojená s kvalitou datových sad a navrhli techniky pro její zlepšení. Aby mohli uživatelé hodnotit kvalitu dat, potřebují mít o datech co nejvíce informací. Proto by také open data měla být zveřejňována s podrobnými metadaty (Ubaldi, 2013). Každá datová sada spojená s metadaty by měla obsahovat informace o datové sadě (popis, datum zveřejnění, datum aktualizace, kategorie, klíčová slova a licence), informace o vydavateli (odpovědné instituci) a informace o pokrytí (časové, prostorové a granularita). Důležitý je rovněž formát, když data by měla být dostupná ve více formátech, což podporuje interoperabilitu a další integraci těchto dat (Verma a Gupta, 2012). Z hlediska zajištění interoperability metadat mezi různými open data portály a jejich rychlejší nalezení je možné využít slovník RDF nazvaný Data Catalogue Vocabulary (DCAT) od organizace World Wide Web Consortium (W3C), který definuje atributy pro popis datových sad (Cyganiak a Maali, 2015). Někteří autoři také navrhují vlastní slovníky DCAT RDF jako výměnný formát pro standardizaci popisu datových katalogů na různých portálech, jako např. Maali et al. (2010).

Při tvorbě těchto portálů by instituce veřejného sektoru měly spolupracovat také s projekty a komunitami věnujícími se této problematice (Cranefield et al., 2014). Kromě toho výzkumy věnující se interoperabilitě v e-governmentu a různým datovým integracím ukazují vysokou citlivost uživatelů na povahu informací a kvalitativních faktorů jako je např. komplexnost, odpovědnost, důvěryhodnost a hodnota hledaných dat (Dawes et al., 2004). Nicméně Yang a Kankanhalli (2013) tvrdí, že navzdory tomu, že instituce veřejného sektoru často aktivně propagují využívání těchto dat, odezva od zúčastněných subjektů je nedostatečná. Zuiderwijk a Janssen (2013) jsou pak toho názoru, že výstupy znovu využívání těchto dat by měly být více diskutovány, resp. jednotlivé postupy by měly být rovněž zveřejněny. Zároveň by měly být zkoumány i důvody, proč tato data nejsou více využívána. Janssen et al. (2012) zmiňují některé faktory bránící vyššímu využívání těchto dat jako je např. nedostatečné vysvětlení významu dat, špatná propagace nebo chybějící znalosti o tom, jak tato data zpracovat. Martin et al. (2013) definovali 7 kategorií rizik spojených s open daty: správa, ekonomické problémy, licence a právní rámce, charakteristiky dat, metadata, přístup a dovednosti. Stejně tak různé faktory od institucionálních po technické ovlivňují vývoj a implementaci OGD portálů na národní úrovni (Janssen et al., 2012). Často se však také stává, že vlády ve snaze zvyšovat vlastní důvěryhodnost kladou až přílišný důraz na tvorbu národního open data portálu a zcela opomíjejí další rozvoj infrastruktury pro přenos a ukládání dat

mezi jednotlivými institucemi (Ubaldi, 2013). Tato problematika je diskutována v Courmont (2012), kde se autor zaměřuje na politiku právního, ekonomického a technického rozhodování v kontextu dat. Někteří autoři také doporučují vybírat relevantní data na základě praktických faktorů jako je popularita, dostupnost a nekonfliktnost (Sayogo et al., 2014). Poslední fází je zapojování a spolupráce jednotlivých subjektů při využívání těchto dat a vytváření přidané hodnoty (Courmont, 2012; Geiger a von Lucke, 2012; Hansen et al., 2013; Jetzek et al., 2013; Kalampokis et al., 2011).

### **3.4 Aplikace využívající open data**

Aplikace jsou důležitou dimenzí open data iniciativ. Využití plného potenciálu principů open dat záleží především na tom, jak jsou tyto výstupy prezentovány veřejnosti. Nejschůdnější cestou jsou interaktivní webové a mobilní aplikace, které mají větší potenciál zasáhnout širší skupinu uživatelů, na rozdíl od různých webových katalogů a portálů, na kterých je nutné data vyhledat a zpravidla ještě zpracovat pro získání požadované informace (Lněnička a Máchová, 2014). Aplikace mohou být vyvíjeny institucemi veřejného sektoru, podniky, občany nebo různými komunitami. Každý z těchto subjektů má však jiné motivy a cíle, proč tyto aplikace vytvářet. Kromě komerčních cílů podniků jsou nejčastějším motivem transparentnost výdajů a řízení ve veřejném sektoru (Verma a Gupta, 2012). Většina těchto aplikací má za cíl pomoci uživatelům nalézt nějakou informaci, poskytnout službu nebo umožnit zobrazení vybraných datových sad v nějakém kontextu, např. přímo v mapě (Cranefield et al., 2014). Nejčastěji jsou vytvářeny jednoduché aplikace, které novým a atraktivním způsobem zobrazují open data. Tyto aplikace jsou pak často postaveny na nových obchodních modelech, kdy provoz těchto aplikací je financován z reklamních výnosů a ne z plateb od uživatelů (World Bank, 2014).

Foulonneau et al. (2014) porovnali současné přístupy k podpoře vytváření aplikací nad open daty a role těchto dat v návrhu služeb založených na teoretickém rámci inovace služeb. Autoři dospěli k závěru, že tato data mohou aplikace využívat různými způsoby: služba je založena na datech, služba využívá data jako zdroj a služba je daty ověřena, ale tato data nejsou přímo použita nebo nejsou přímo viditelná ve službě. Autoři také zjistili, že většina již fungujících aplikací využívá pouze omezený počet nejvíce populárních datových sad. Cowan et al. (2014) definovali 6 různých oblastí pro vývoj aplikací, které by mohly zaujmout nejvíce nových uživatelů a podporovat transparentnost a další rozvoj veřejného sektoru. Heath a Bizer (2011) také vymezili 8 kategorií dat, když se zaměřili především na linked

data. Lněnička a Máchová (2014) pak definovali 10 různých oblastí možného zaměření aplikací, které pracují nad open daty. Nejčastějšími vstupními daty pro vývoj aplikací jsou vládní výdaje a další investiční výdaje, dopravní informace, geodata, energie, kvalita životního prostředí, seznamy různých institucí veřejného sektoru, atd. (Cowan et al., 2014; Yang a Kankanhalli, 2013).

Některé úspěšně implementované aplikace pracující nad open daty lze nalézt např. v Yang a Kankanhalli (2013). Oficiální národní open data portál Spojeného království ([data.gov.uk](http://data.gov.uk)) pak v dubnu 2015 registroval celkem 367 aplikací využívajících open data, když nejvíce jich bylo z oblasti dopravy (44), lokalizačních služeb (38), společnosti (29) a zdraví (18). Oficiální open data portál USA ([data.gov](http://data.gov)) pak nabízí téměř 70 aplikací, které jsou zdarma, když na další aplikace nabízí odkazy. Téměř 80 aplikací lze nalézt také na portálu vyvíjeném OKF ([publicdata.eu](http://publicdata.eu)), který poskytuje přístup k oficiálním datovým sadám z celé Evropy na jednom místě. Mezi další státy světa podporující vytváření těchto aplikací patří ještě např. Austrálie nebo Singapur. V České republice (ČR) je to pak např. portál Otevřená data ([otevrenadata.cz](http://otevrenadata.cz)), který spravuje Fond Otakara Motejla v rámci Fóra pro otevřená data a již dvakrát pořádal soutěž o nejlepší aplikaci nad open daty. Tento portál není oficiálním národním open data portálem, ten ČR stále nemá, nicméně nejbliže je mu CKAN portál ([cz.ckan.net](http://cz.ckan.net)).

#### **4 VÝSLEDKY A DISKUZE K DOTAZNÍKOVÉMU ŠETŘENÍ**

Cílem průzkumu bylo získat názory studentů na problematiku open dat a souvisejících pojmů, včetně jejich zájmu o konkrétní tematické zaměření dat, která by měla být zveřejňována a dále nad nimi vytvářeny různé webové a mobilní aplikace. Vysokoškolští studenti byli zvoleni jako cílová skupina s ohledem na povahu tohoto tématu, resp. minimální požadované znalosti pro práci s těmito daty. Jak bylo zmíněno již v rámci literární rešerše, tak největší přidaná hodnota je z těchto dat získána právě jejich vizualizací prostřednictvím aplikací. Průzkumu se zúčastnilo celkem 157 studentů Fakulty ekonomicko-správní Univerzity Pardubice. Podrobně je vše zachyceno v Tabulce 2, kde je EM (Ekonomika a management), HPS (Hospodářská politika a správa) a SII (Systémové inženýrství a informatika).

**Tab. 2: Struktura dotazovaných studentů.**

Obor / ročník	1.	2.	3.	4.	5.	Celkem
EM	15	19	11	5	4	54
HPS	8	3	8	3	5	27
SII	9	32	10	14	11	76
<b>Celkem</b>	<b>32</b>	<b>54</b>	<b>29</b>	<b>22</b>	<b>20</b>	<b>157</b>

Zdroj: autoři.

Dále bylo zjišťováno, zda studenti znají související pojmy open, linked a big data. Výsledky jsou zachyceny v Tabulce 3 jako % podíl ze všech odpovědí, když na každou otázku bylo možné odpovědět: a) ANO, b) NE, c) Slyšel(a), ale neznám přesný význam. Podle výsledků studenti nejvíce znají význam pojmů open data a big data, když povědomí o těchto pojmech roste podle ročníků. Pojem linked nebo LOD je u studentů nižších ročníků prakticky neznámý a i u vyšších ročníků je znalost tohoto pojmu méně než třetinová. Pojem open data je pak známější u studentů oborů EM a HPS, pojem big data u studentů SII, což souvisí se spíše technologickou orientací pojmu big data a jeho vazby na systémy pro zpracování těchto dat.

**Tab. 3: Znalosti významu jednotlivých pojmů.**

Ročník / pojem	Open data			Linked data			Big data		
	a) [%]	b) [%]	c) [%]	a) [%]	b) [%]	c) [%]	a) [%]	b) [%]	c) [%]
1.	9,4	40,6	50,0	0,0	78,1	21,9	6,3	68,8	25,0
2.	5,6	50,0	44,4	1,9	74,1	24,1	7,4	53,7	38,9
3.	27,6	31,0	41,4	10,3	55,2	34,5	20,7	27,6	51,7
4.	54,5	18,2	27,3	22,7	50,0	27,3	45,5	22,7	31,8
5.	70,0	10,0	20,0	30,0	40,0	30,0	65,0	20,0	15,0
<b>Průměr</b>	<b>33,4</b>	<b>30,0</b>	<b>36,6</b>	<b>13,0</b>	<b>59,5</b>	<b>27,6</b>	<b>29,0</b>	<b>38,6</b>	<b>32,5</b>

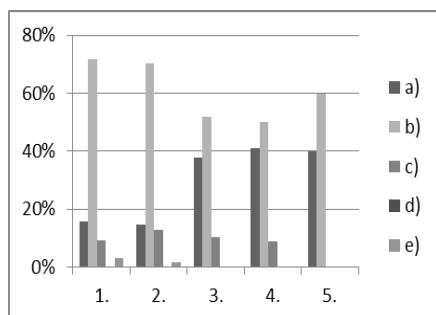
Zdroj: autoři.

Po této otázce byly studentům jednotlivé pojmy blíže vysvětleny a ti poté odpovídali na další otázky. První z nich zjišťovala jejich názor na to, zda by instituce veřejného sektoru měly častěji dávat volně k dispozici tato data (kromě citlivých dat, osobních údajů apod.). Na výběr byla jedna z těchto možností: a) Rozhodně ano, b) Spíše ano, c) Spíše ne, d) Rozhodně ne, e) Neumím se vyjádřit. Stejně možnosti byly na výběr i u další otázky zjišťující, zda by mělo zveřejňování open dat probíhat na všech úrovních veřejné správy, zejména pak na regionální úrovni krajů a lokální úrovni větších měst. Výsledky lze porovnat na grafech 1 a 2, kde na ose x je ročník



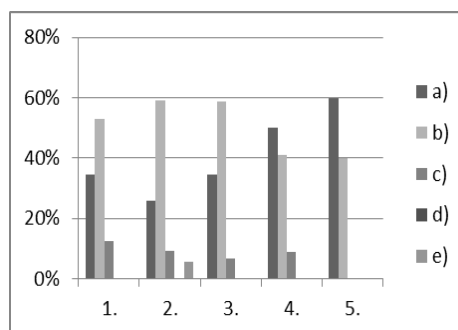
studia. K častějšímu zveřejňování open dat se studenti staví kladně, když podpora roste u vyšších ročníků. Celkový průměr pro všechny ročníky je 26,1 % pro rozhodně ano; 63,1 % pro spíše ano; 9,5 % pro spíše ne a 1,3 % se k tématu neumí vyjádřit. U ani jedné z těchto dvou otázek nikdo nevedl odpověď rozhodně ne, a to i přesto, že studentům byly vysvětleny i nevýhody a možná rizika těchto dat. Podpora pro zveřejňování open dat na všech úrovních veřejné správy je ještě vyšší než u předchozí otázky, když průměr pro všechny dotazované je 36,9 % pro rozhodně ano; 52,9 % pro spíše ano; 8,3 % pro spíše ne a 1,9 % studentů se k této otázce neumí vyjádřit.

**Graf 1: Zveřejňování open dat institucemi veřejné správy obecně.**



Zdroj: autoři.

**Graf 2: Zveřejňování open dat na všech úrovních veřejné správy.**



Zdroj: autoři.

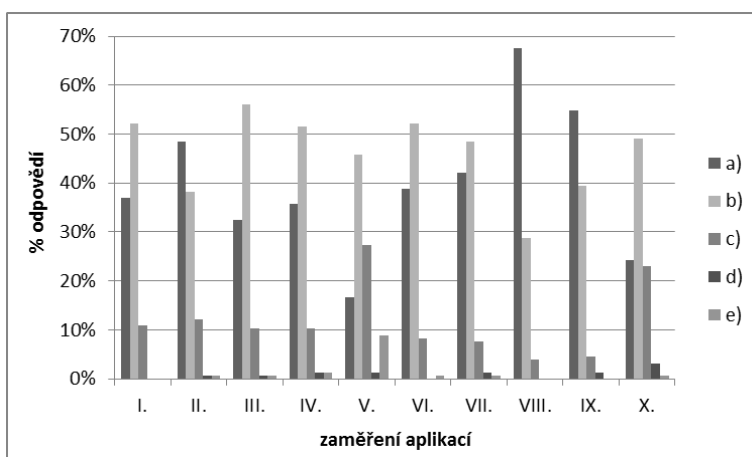
Další část dotazníku se pak věnovala webovým a mobilním aplikacím vytvářeným nad těmito daty. První otázka proto zjišťovala, zda se studenti již někdy setkali s těmito aplikacemi. Výsledky jsou zachyceny v Tabulce 4 jako % ze všech odpovědí pro daný ročník. Z výsledků je patrné, že většina studentů se již s těmito aplikacemi někdy setkala, když tyto aplikace jsou většinou zřetelně označeny např. logem OKF. Nicméně jak bylo zjištěno již výše, většina studentů především z nižších ročníků často neví, co si pod těmito pojmy představit a často po tom ani nepátrají a nevěnují pozornost zdrojům dat. Důvěryhodnost poskytovatele dat by přitom měla být zásadní pro práci s aplikací, resp. rozhodování se na základě jejich výstupů.

**Tab. 4: Znalosti v oblasti aplikací vytvářených nad open daty.**

Odpověď / ročník	1. [%]	2. [%]	3. [%]	4. [%]	5. [%]
ANO	37,5	24,1	41,4	68,2	75,0
NE	21,9	29,6	27,6	18,2	15,0
Nevěnuji pozornost zdrojům (označení) dat.	40,6	46,3	31,0	13,6	10,0

Zdroj: autoři.

Dále byli studenti dotazováni, jaké zaměření aplikací by podle jejich názoru zaujalo nejvíce nových uživatelů, resp. jaké aplikace by nově přivítali a měly by se vytvářet nad open daty. Studenti hodnotili celkem 10 různých oblastí možného zaměření dané aplikace, které byly definovány ve Lněnička a Máchová (2014). Ty jsou poté v grafu 3 označeny jako: I. (Doprava a prostorová data), II. (Finance a vládní výdaje), III. (Kriminalita a justice), IV. (Kultura), V. (Regionální a lokální zdroje), VI. (Školství), VII. (Turismus a sport), VIII. (Zaměstnanost a práce), IX. (Zdravotní péče a služby) a X. (Životní prostředí a udržitelný rozvoj). Názor na konkrétní zaměření aplikace mohli studenti vyjádřit výběrem jedné z možností: a) Rozhodně ano, b) Spíše ano, c) Spíše ne, d) Rozhodně ne, e) Neumím se vyjádřit. Zájem studentů je pak v grafu 3, stejně jako v grafu 4, vyjádřen jako průměr pro všechny ročníky, tzn. 157 studentů.

**Graf 3: Rozložení odpovědí podle možného zaměření aplikací nad open daty.**

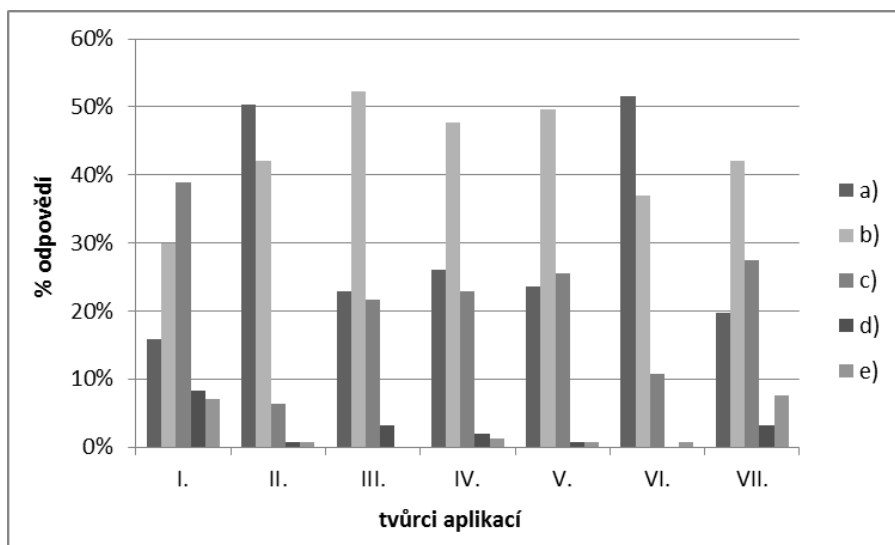
Zdroj: autoři.

Nežžádanějšími aplikacemi nad open daty jsou pro studenty ty se zaměřením na zaměstnanost a práci (67,5 %), zdravotní péče a služby (54,8 %) a finance a vládní výdaje (48,4 %), kde studenti vybrali odpověď rozhodně ano. Nejméně žádané jsou aplikace využívající regionální a lokální zdroje a také životní prostředí a udržitelný rozvoj. U oblasti V. rovněž téměř 9 % studentů uvedlo, že se neumí vyjádřit, což je pravděpodobně způsobeno tím, že vymezení této oblasti je pro ně nejasné. Proto bude tato oblast zřejmě v dalších výzkumech odstraněna. Odpovědi studentů z různých ročníků se výrazně nelišily, když studenti 1. ročníku vybírali odpověď za a) v průměru 42,8 %, 2. ročník (40,1 %), 3. ročník (39,7 %), 4. ročník (40,5 %) a 5. ročník (47 %). Studenti 1. a 2. ročníku také častěji uváděli, že se neumí vyjádřit. Studenti také mohli navrhnout nové oblasti, kde by aplikace nad open daty našly uplatnění. Mezi odpověďmi se objevily např. meteorologické modely počasí a povodně, které spadají do oblasti X., nebo bližší informace o konkrétních volených zastupitelích a jejich finanční situaci, což by mohla být podoblast oblasti II.

Nakonec byli studenti dotázáni na to, kdo by měl tyto aplikace aktivně vytvářet, resp. kdo by z nich mohl získat největší přidanou hodnotu, ať už pro komerční nebo nekomerční využívání. Studenti hodnotili celkem 7 možných tvůrců. Ti jsou poté v grafu 4 označeny jako: I. (Běžní občané), II. (Instituce veřejné správy), III. (Nadnárodní a větší podniky), IV. (Neziskové organizace), V. (Regionální a menší podniky), VI. (Vysoké školy a výzkumné ústavy) a VII. (Zájmové skupiny). Názor na konkrétní zaměření aplikace mohli studenti opět vyjádřit výběrem jedné z možností: a) Rozhodně ano, b) Spíše ano, c) Spíše ne, d) Rozhodně ne, e) Neumím se vyjádřit. Podle názoru dotazovaných studentů by aplikace nad open daty v první řadě měly vytvářet vysoké školy a výzkumné ústavy (51,6 %) a instituce veřejné správy (50,3 %). Spíše ano by pak měly aplikace vytvářet nadnárodní a větší podniky (52,2 %), regionální a menší podniky (49,7 %) a neziskové organizace (47,8 %). Zde se názory studentů mírně liší od současného trendu, kdy jsou tyto aplikace sice často vytvářeny studenty vysokých škol, ale hlavními tvůrci jsou jednotlivci nebo zájmové skupiny, často pak s podporou neziskových organizací, viz např. Fond Otakara Motejla v ČR. V soukromém sektoru je zájem o tato data a nad nimi vytvářenými aplikacemi stále minimální. Pokud jde o výsledky odpovědí studentů podle ročníků, tak studenti 1. ročníku vybírali odpověď za a) v průměru 19,1 %, 2. ročník (16,3 %), 3. ročník (20,7 %), 4. ročník (27,3 %) a 5. ročník (30,5 %). Odpověď b) je u všech ročníků kolem 30 % a odpověď za c) je u 1. ročníku

(21,9 %), u 2. ročníku (16,9 %), u 3. ročníku (14,1 %), u 4. ročníku (10,5 %) a u 5. ročníku (8 %).

**Graf 4: Rozložení odpovědí podle možných tvůrců aplikací nad open daty.**



Zdroj: autoři.

## 5 ZÁVĚR A DALŠÍ SMĚŘOVÁNÍ VÝZKUMU

Veřejný sektor vytváří, shromažďuje, zpracovává a šíří velké množství dat. Ta mohou být znovu využita a různě integrována pak vytvářet nové služby a produkty s přidanou hodnotou a potenciálně významným dopadem v globální ekonomice. Hlavním důvodem pro zveřejnění open dat je znovuobjevení hodnoty těchto dat prostřednictvím jejich využívání, propojování a kombinací s dalšími zdroji, podpora inovací, získávání zpětné vazby od jejich uživatelů, zvyšování transparentnosti, atd. Za tímto účelem jsou jednotlivými státy vytvářeny OG a open data iniciativy, jejichž výstupem zpravidla bývá vytvoření národního nebo lokálních open data portálů, kde pak mohou různé subjekty tato data prohledávat, stahovat, kombinovat je s dalšími zdroji a vytvářet z nich aplikace. Vlády a jednotlivé instituce veřejného sektoru by zároveň měly v této oblasti aktivně vystupovat a podporovat využívání těchto dat.

Jelikož je tato oblast stále nová, tak je nutné hledat odpovědi na související otázky. Proto byla v rámci tohoto příspěvku podrobně popsána situace v oblasti open, big a linked dat, včetně zhodnocení jejich přínosů, rizik a výzev v další práci s nimi. Dále byl přiblížen význam open data portálů a jejich fungování. Hlavní část se věnovala aplikacím vytvářeným nad open data, když bylo provedeno dotazníkové šetření, jehož účelem bylo nalézt odpovědi na otázky, kdo by měl dané aplikace aktivně vytvářet a jaké zaměření by tyto aplikace měly mít. Jako cílová skupina byli zvoleni vysokoškolští studenti a byly porovnávány jejich odpovědi v různých ročnících studia. Podle výsledků jsou známější pojmy open a big data, linked data výrazně méně. Zároveň znalost jednotlivých pojmů roste podle ročníků od 5–10 % v 1. ročníků do 65–70 % v 5. ročníku. Studenti také podporují zveřejňování těchto, a to na všech úrovních veřejné správy. Pokud jde o zaměření aplikací, které by studenty nejvíce zaujalo, tak se nejlépe umístily ty se zaměřením na zaměstnanost a práci, zdravotní péči a služby, a finance a vládní výdaje. Tato zaměření jsou mírně odlišná od již fungujících aplikací, viz open data portály Spojeného království nebo USA, z čehož lze usuzovat, že podle studentů je oblast aplikací zaměřených na dopravu a lokalizační služby již dostatečně pokryta. Podle názoru dotazovaných studentů by aplikace nad open data v první řadě měly vytvářet vysoké školy a výzkumné ústavy a dále pak instituce veřejné správy.

Závěrem je však nutné zmínit, že názory studentů z realizovaného průzkumu nemohou přímo ovlivnit praxi open dat, ale jako cílová skupina, která je problematice využívání a zpracování open dat nejbližší, viz požadavky na tato data nebo to, že dotazování studenti mají potřebné ekonomické i infromatické vzdělání, tak i s ohledem na budoucí vývoj v této oblasti je nutné brát jejich názory v úvahu, především jako budoucích uživatelů. Nicméně další výzkum bude směřovat přímo na jednotlivé instituce veřejného sektoru jako možné poskytovatele těchto dat a bude zjišťován i jejich názor na tuto problematiku. Zároveň bude další práce zaměřena na regionální a lokální úroveň, včetně open data portálů na této úrovni. Dále bude řešena také problematika integrace těchto dat a rozvoj potřebné infrastruktury. Pozornost bude věnována také oblasti velkých objemů open dat, tzv. open (big) data, více např. v Marton et al. (2013), a možnostem jejich zpracování.

**POUŽITÉ ZDROJE**

- [1] BUCHHOLTZ, S.; BUKOWSKI, M.; ŚNIEGOCKI, A. *Big and open data in Europe: A growth engine or a missed opportunity*. Varšava: demosEUROPA, 2014, 113 s. ISBN 978-83-925542-1-9.
- [2] CHE, D.; SAFRAN, M.; PENG, Z.. From Big Data to Big Data Mining: Challenges, Issues, and Opportunities. In: *Database Systems for Advanced Applications*. Springer, Berlin Heidelberg, 2013. s. 1-15.
- [3] CHEN, M.; MAO, S.; LIU, Y. Big Data: A Survey. *Mobile Networks and Applications*, 2014, vol. 19, no. 2, s. 171-209. ISSN 1572-8153.
- [4] COURMONT, A. How to Govern Open Data? The politics of open data portals. In: *Paper presented at the IPSA Conference*, Madrid, July 8-12, 2012. s. 1-17.
- [5] COWAN, D.; ALENCAR, P.; MCGARRY, F. Perspectives on Open Data: Issues and Opportunities. In: *Proceedings of the 2014 IEEE International Conference on Software Science, Technology and Engineering*. IEEE, 2014. s. 24-33.
- [6] CRANEFIELD, J.; ROBERTSON, O.; OLIVER, G. Value in the Mash: Exploring the Benefits, Barriers and Enablers of Open Data Apps. In: *Proceedings of the 22nd European Conference on Information Systems*, Tel Aviv, Israel, 2014, s. 1-15.
- [7] CYGANIAK, R.; MAALI, F. *Use Cases and Requirements for the Data Catalogue Vocabulary* [online]. 2015 [cit. 2015-04-04]. Dostupné z: <https://dvcs.w3.org/hg/gld/raw-file/default/dcat-ucr/index.html>
- [8] DAWES, S. S.; PARDO, T. A.; CRESSWELL, A. M. Designing Electronic Government Information Access Programs: A Holistic Approach. *Government Information Quarterly*, 2004, vol. 21, iss. 1, s. 3-23. ISSN 0740-624X.
- [9] DEMCHENKO, Y., et al. Addressing Big Data Issues in Scientific Data Infrastructure. In: *Proceedings of International Conference on Collaboration Technologies and Systems*, IEEE, 2013. s. 48-55.
- [10] ELGENDY, N.; ELRAGAL, A. Big Data Analytics: A Literature Review Paper. In: *Advances in Data Mining: Applications and Theoretical Aspects*. Springer, Heidelberg, 2014. s. 214-227.

- [11] ERMILOV, I., et al. Linked Open Data Statistics: Collection and Exploitation. In: *Knowledge Engineering and the Semantic Web*. Springer, Berlin Heidelberg, 2013. s. 242-249.
- [12] FOULONNEAU, M.; TURKI, S.; VIDOU, G.; MARTIN, S. Open data in Service design. *Electronic Journal of e-Government*, 2014, vol. 12, iss. 2, s. 99-107. ISSN 1479-439X.
- [13] GEIGER, C. P.; VON LUCKE, J. Open Government and (Linked) (Open) (Government) (Data). *eJournal of eDemocracy & Open Government*, 2012, vol. 4, no. 2, s. 265-278. ISSN 2075-9517.
- [14] GURSTEIN, M. B. Open data: Empowering the empowered or effective data use for everyone?. *First Monday*, 2011, vol. 16, no. 2.
- [15] HANSEN, H. S.; HVINGEL, L.; SCHRØDER, L. Open Government Data – A Key Element in the Digital Society. In: *Technology-Enabled Innovation for Democracy, Government and Governance*. Springer, Berlin, 2013. s. 167-180.
- [16] HEATH, T.; BIZER, C. Linked Data: Evolving the Web into a Global Data Space. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 2011, vol. 1, no. 1, s. 1-136.
- [17] HÖCHTL, J.; REICHSTÄDTER, P. Linked Open Data - A Means for Public Sector Information Management. In: *Proceedings of the 2<sup>nd</sup> International Conference on Electronic Government and the Information Systems Perspective: EGOVIS 2011*, Springer, Berlin Heidelberg, 2011. s. 330-343.
- [18] HYLAND, B.; WOOD, D. The Joy of Data-A Cookbook for Publishing Linked Government Data on the Web. In: *Linking Government Data*. Springer, New York, 2011. s. 3-26.
- [19] INTERNATIONAL TELECOMMUNICATION UNION. *Measuring the Information Society Report 2014*. Geneva: Telecommunication Development Bureau, 2014. 250 s. ISBN 978-92-61-15291-8.
- [20] JANSSEN, M.; CHARALABIDIS, Y.; ZUIDERWIJK, A. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 2012, vol. 29, no. 4, s. 258-268. ISSN 1934-8703.

- [21] JETZEK, T.; AVITAL, M.; BJØRN-ANDERSEN, N. Generating Sustainable Value from Open Data in a Sharing Society. In: *Proceedings of IFIP WG 8.6 International Conference on Transfer and Diffusion of IT: TDIT 2014*, Springer, Berlin Heidelberg, 2014. s. 62-82.
- [22] JETZEK, T.; AVITAL, M.; BJØRN-ANDERSEN, N. Generating Value from Open Government Data. In: *Proceedings of 34th International Conference on Information Systems: ICIS 2013*. Bepress, Berkeley, 2013. s. 1-20.
- [23] KALAMPOKIS, E.; TAMBOURIS, E.; TARABANIS, K. Linked Open Government Data Analytics. In: *Proceedings of the 12th IFIP WG 8.5 International Conference: EGOV 2013*. Springer, Berlin Heidelberg, 2013. s. 99-110.
- [24] KALAMPOKIS, E.; TAMBOURIS, E.; TARABANIS, K. Open Government Data: A stage model. In: *Proceedings of the 10th IFIP WG 8.5 International Conference: EGOV 2011*. Springer, Berlin Heidelberg, 2011. s. 235-246.
- [25] KOSTOVSKI, M.; JOVANOVIK, M.; TRAJANOV, D. Open Data Portal based on Semantic Web Technologies. In: *Proceedings of the 7th South East European Doctoral Student Conference*. University of Sheffield, Greece, 2012, s. 1-13.
- [26] KRISHNAN, S.; TEO, T. S. H.; LIM, J. E-Participation and E-Government Maturity: A Global Perspective. In: *Grand Successes and Failures in IT: Public and Private Sectors*. Springer, Berlin, 2013. s. 420-435.
- [27] KUCERA, J.; CHLAPEK, D. Benefits and Risks of Open Government Data. *Journal of Systems Integration*, 2014, vol. 5, no. 1, s. 30-41. ISSN 1804-2724.
- [28] KUČERA, J.; CHLAPEK, D.; NEČASKÝ, M. Open Government Data Catalogs: Current Approaches and Quality Perspective. In: *Technology-Enabled Innovation for Democracy, Government and Governance*. Springer, Berlin, 2013, s. 152-166.
- [29] LNĚNIČKA, M. An In-Depth Analysis of Open Data Portals as an Emerging Public E-Service. *International Journal of Social, Education, Economics and Management Engineering*, 2015, vol. 9, no. 2, s. 551-561.



- [30] LNĚNIČKA, M.; MÁCHOVÁ, R. Otevřená data a jejich vliv na zvyšování efektivnosti a rozhodování ve veřejné správě. *Advances in Informatics, Management and Economics: Pokroky v informatice, managementu a ekonomii*. Pardubice: Univerzita Pardubice, 2014, s. 66-87. ISBN 978-80-7395-811-4.
- [31] LOURENÇO, R. P. Open Government Portals Assessment: A Transparency for Accountability Perspective. In: *Proceedings of the 12th IFIP WG 8.5 International Conference: EGOV 2013*, Springer, Heidelberg, 2013, s. 62-74.
- [32] MAALI, F.; CYGANIAK, R.; PERISTERAS, V. Enabling Interoperability of Government Data Catalogues. In: *Proceedings of the 9th IFIP WG 8.5 International Conference (EGOV 2010)*, Springer, Heidelberg, 2010, s. 339-350.
- [33] MACHADO, A. L.; PARENTE DE OLIVEIRA, J. M. DIGO: An Open Data Architecture for E-government. In: *Proceedings of the 15th IEEE International Enterprise Distributed Object Computing Conference Workshops*, IEEE, 2011, s. 448-456.
- [34] MARTIN, S., et al. Open Data: Barriers, Risks and Opportunities. In: *Proceedings of the 13th European Conference on eGovernment: ECEG 2013*. Academic Conferences Limited, 2013, s. 301-309.
- [35] MARTON, A.; AVITAL, M.; BLEGIN JENSEN, T. Reframing Open Big Data. In: *Proceedings of the 21st European Conference on Information Systems*, Utrecht, Netherlands, 2013, s. 1-12.
- [36] OPEN KNOWLEDGE. *Global Open Data Index by Open Knowledge* [online]. 2014 [cit. 2015-04-13]. Dostupné z: <http://index.okfn.org/methodology/>
- [37] PAVLO, A., et al. A comparison of approaches to large-scale data analysis. In: *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data*. ACM, 2009. s. 165-178.
- [38] SAYOGO, D. S.; PARDO, T. A.; COOK, M. A Framework for Benchmarking Open Government Data Efforts. In: *Proceedings of the 47th Hawaii International Conference on System Sciences*. IEEE, 2014. s. 1896-1905.
- [39] SOLAR, M.; CONCHA, G.; MEIJUEIRO, L. A Model to Assess Open Government Data in Public Agencies. In: *Proceedings of the 11th IFIP*

WG 8.5 *International Conference: EGOV 2012*, Springer, Heidelberg, 2012. s. 210-221.

- [40] TIEN, J. M. Big Data: Unleashing Information. *Journal of Systems Science and Systems Engineering*, 2013, vol. 22, no. 2, s. 127-151. ISSN 1861-9576.
- [41] TRAUNMÜLLER, R. Open Government and Electronic Government: Some Considerations. In: *Proceedings of the 3rd International Conference on Electronic Government and the Information Systems Perspective: EGOVIS 2014*, Springer International Publishing, 2014. s. 201-207.
- [42] TSIAVOS, P.; STEFANEAS, P.; KAROUNOS, T. Financial Crisis and The Promised Land of Open Data. In: *Proceedings of the conference IPP2012: Big Data, Big Challenges?*. Oxford Internet Institute, 2012. s. 1-21.
- [43] UBALDI, B. Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives, *OECD Working Papers on Public Governance*, No. 22, OECD Publishing, 2013, 60 s.
- [44] UNITED NATIONS. *United Nations e-Government Survey 2012: E-Government for the People*. New York: UN Publishing Section, 2012. 143 s. ISBN 978-92-1-123190-8.
- [45] UNITED NATIONS. *United Nations e-Government Survey 2014: E-Government for the Future We Want*. New York: UN Publishing Section, 2014. 263 s. ISBN 978-92-1-123198-4.
- [46] VAN DER WAAL, S., et al. Lifting Open Data Portals to the Data Web. In: *Linked Open Data – Creating Knowledge Out of Interlinked Data*. Springer International Publishing. 2014, s. 175-195.
- [47] VERMA, N.; GUPTA, M. P. Open Government Data : More than Eighty Formats. In: *Proceedings of the 9th International Conference on E-Governance (ICEG 2012)*. CSI, 2012, s. 207-216.
- [48] WORLD BANK. *Open Data for Economic Growth* [online]. 2014 [cit. 2015-04-04]. Dostupné z: <https://openknowledge.worldbank.org/handle/10986/19997>
- [49] WORLD ECONOMIC FORUM. *The Global Information Technology Report 2014: Rewards and Risks of Big Data*. Geneva: SRO-Kundig, 2014. 343 s. ISBN 978-92-95044-63-0.

- [50] WORLD WIDE WEB FOUNDATION. Open Data Barometer Global Report [online]. 2015 [cit. 2015-04-04]. Dostupné z: <http://opendatabarometer.org/assets/downloads/Open%20Data%20Barometer%20-%20Global%20Report%20-%202nd%20Edition%20-%20PRINT.pdf>
- [51] YANG, Z.; KANKANHALLI, A. Innovation in Government Services: The Case of Open Data. In: *Grand Successes and Failures in IT: Public and Private Sectors*. Springer, Berlin, 2013. s. 644-651.
- [52] ZUIDERWIJK, A.; JANSSEN, M. A Coordination Theory Perspective to Improve the Use of Open Data in Policy-Making. In: *Proceedings of the 12th IFIP WG 8.5 International Conference: EGOV 2013*, Springer, Heidelberg, 2013. s. 38-49.

#### AUTOŘI:

**Ing. Renáta Máchová, Ph.D.**, Ústav systémového inženýrství a informatiky, Fakulta ekonomicko-správní, Univerzita Pardubice, Studentská 84, 53210 Pardubice, Česká republika, e-mail: [renata.machova@upce.cz](mailto:renata.machova@upce.cz)

**Ing. et Ing. Martin Lněnička**, Ústav systémového inženýrství a informatiky, Fakulta ekonomicko-správní, Univerzita Pardubice, Studentská 84, 53210 Pardubice, Česká republika, e-mail: [martin.lnenicka@gmail.com](mailto:martin.lnenicka@gmail.com)

#### AUTHORS:

**Ing. Renáta Máchová, Ph.D.**, Institute of System Engineering and Informatics, Faculty of Economics and Administration, University of Pardubice, Studentská 84, 53210 Pardubice, Czech Republic, e-mail: [renata.machova@upce.cz](mailto:renata.machova@upce.cz)

**Ing. et Ing. Martin Lněnička**, Institute of System Engineering and Informatics, Faculty of Economics and Administration, University of Pardubice, Studentská 84, 53210 Pardubice, Czech Republic, e-mail: [martin.lnenicka@gmail.com](mailto:martin.lnenicka@gmail.com)