



Reg. č. projektu: CZ 1.04/ 4.1.00/A3.00004

## Open data

## Pracovní sešit

Materiál vznikl v rámci řešení projektu „**Vzdělávání v oblasti základních registrů a dalších kmenových projektů eGovernmentu**“, registrační číslo projektu: CZ 1.04/4.1.00/A3.00004, který je financován z prostředků Evropského sociálního fondu ČR, Operačního programu Lidské zdroje a zaměstnanost

Zpracovatel – Institut pro veřejnou správu Praha

Praha, červen 2014

# **OBSAH PRACOVNÍHO SEŠITU**

## **Tento pracovní sešit:**

- přináší doplňující informace k výkladu (např. slovníček terminologie související s obsahem kurzu, odkazy na užitečné webové stránky k tématu, seznam doporučené odborné literatury a podobně)
- slouží pro opakování a procvičování učiva probraného v teoretické části kurzu
- aktivizuje účastníky kurzu, usiluje o jejich participaci při plnění cílů výuky
- obsahuje zadání zpětnovazebních aktivit (různých typů cvičení), které účastníci kurzu řeší ve skupinách nebo individuálně
- může absolventy kurzu inspirovat k aktivitám nejen v rámci prezenční výuky, ale také při následném domácím samostudiu

## **TEORETICKÁ ČÁST:**

- Související legislativa
- Slovníček pojmů
- Užitečné webové stránky
- Doporučená odborná literatura

## **PRAKTICKÁ ČÁST:**

- Spojovačky
- Doplnovačky
- Kontrolní otázky
- Test



# A. TEORETICKÁ ČÁST

## SOUVISEJÍCÍ LEGISLATIVA



V České republice nejsou k červnu 2014 open data v legislativě nijak zakotvena. Zároveň nejsou ani omezena – s výjimkou chráněných informací (bezpečnost, osobní údaje a podobně). Nicméně připravuje se implementace Směrnice Evropského parlamentu a Rady 2013/37/ES, kterou se mění Směrnice Evropského parlamentu a Rady 2003/98/ES o opakovaném použití informací veřejného sektoru.

Implementace se pravděpodobně bude provádět prostřednictvím novel zákona o informačních systémech veřejné správy a zákona o svobodném přístupu k informacím. Příprava těchto novel je v době, kdy vznikl tento pracovní sešit, v začátcích, a tak není jasné, zda budou open data pro některé oblasti uložena jako povinnost, nebo zůstanou zcela dobrovolná. Zároveň se také diskutuje o odkladu případných povinností pro samosprávy až o dva roky. Jistým je nyní pouze to, že bude vymezen pojem „open data“ (možná v překladu „otevřená data“) a bude zřízen státní katalog open dat. Diskutována je také možnost, že čerpání některých dotací bude podmíněno zveřejněním open dat.

- Zákon č. 101/2000 Sb., o ochraně osobních údajů a o změně některých zákonů
- Zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon)
- Zákon č. 365/2000 Sb., o informačních systémech veřejné správy
- Zákon č. 106/1999 Sb., o svobodném přístupu k informacím
- Směrnice Evropského parlamentu a Rady 2003/98/ES o opakovaném použití informací veřejného sektoru ve znění pozdější úpravy Směrnicí Evropského parlamentu a Rady 2013/37/ES

## SLOVNÍČEK POJMŮ



Mezi významné výstupy tohoto kurzu patří mimo jiné osvojení terminologie týkající se open dat. Vybrali jsme pro Vás několik důležitých pojmů a nabízíme jejich stručné definice.

### **ANONYMIZACE OSOBNÍCH ÚDAJŮ:**

V případě, že data spadají pod ochranu osobních údajů, je třeba z nich osobní údaje vypustit, tedy tzv. anonymizovat. To se dá provést neuvedením těch specifických údajů, které umožňují jednoznačnou identifikaci osoby.

### **DATASETY:**

Open data jsou v základní podobě zveřejňována v datových souborech, např. v excelových tabulkách a CSV. Tyto soubory se nazývají datasety (nebo také datové sady či datové množiny).

### **DATOVÉ SCHRÁNKY:**

Slouží pro komunikaci v oblasti veřejné správy. Jejich prostřednictvím lze činit podání kterémukoliv úřadu. Úřady prostřednictvím datových schránek doručují své písemnosti příslušným adresátům (fyzickým nebo právnickým osobám) a komunikují s jinými orgány veřejné správy. Veškerým úkonům, které jsou prostřednictvím elektronické datové schránky činěny, je přiznána ekvivalence k úkonům činěným písemně.

### **KATALOGY DAT:**

Katalog dat obvykle shromažďuje informace o adrese, kde jsou ke stažení open data, co tato data obsahují, jak často jsou aktualizována, kdo je spravuje a řadu dalších informací o datech. Katalogů existuje ve světě velké množství a většina si mezi sebou vzájemně vyměňuje informace, takže přidáním záznamu do jednoho katalogu se záznam časem objeví také v ostatních.

Plánovaný státní katalog open data by měl být provozován na Portálu veřejné správy. Vkládání nových záznamů a úprava existujících by měla být prováděna prostřednictvím datových schránek (obdobně jako je tomu u popisů životních situací). Za každý záznam by měl být zodpovědný ten úřad, který jej založil. Ten také bude moci záznam v katalogu upravovat (příčemž díky zaslání aktualizace prostřednictvím datových schránek potvrdí, že je skutečně oprávněn záznam upravovat).

#### **OPEN DATA:**

Otevřená data. Data snadno dostupná, volně přístupná na internetu, bez omezení jejich dalšího využití, zveřejněná ve strukturované, strojově zpracovatelné (čitelné) podobě, pod otevřenou licencí (umožňující další zpracování těchto dat).

Cílem zveřejnění open dat je umožnit jejich opakované použití například pro datovou žurnalistiku nebo vznik webových a mobilních aplikací, které poskytují služby občanům a zvyšují transparentnost fungování veřejné správy.

#### **PROPOJITELNÁ OPEN DATA:**

Forma otevřených dat s pěti hvězdami. Za propojitelná je možné open data považovat v okamžiku, kdy má každý záznam v datech obsažený svůj jednoznačný identifikátor v podobě URL adresy. Na této URL adrese jsou k dispozici data o tomto záznamu ve strojově zpracovatelné podobě, která není určena pro běžné uživatele. Princip spočívá v tom, že formou hypertextových odkazů jsou propojena data různých úřadů z různých zemí. V této podobě je uživatel dat (programátor aplikace) schopen snadno zjistit další informace související, aniž by musel sám dohledávat, kdo a kde je zveřejnil. Jedná se tedy o stejný princip, jaký má současný web, kdy jsou webové stránky vzájemně provázány sítí hypertextových odkazů.

#### **OSOBNÍ ÚDAJE:**

Jakékoliv informace týkající se určeného nebo určitelného subjektu údajů. Subjekt údajů se považuje za určený nebo určitelný, jestliže lze subjekt údajů přímo či nepřímo identifikovat zejména na základě čísla, kódu nebo jednoho či více prvků, specifických pro jeho fyzickou, fyziologickou, psychickou, ekonomickou, kulturní nebo sociální identitu.

Úřad pro ochranu osobních údajů za osobní údaj považuje jakoukoli kombinaci údajů, které mohou vést k jednoznačné identifikaci konkrétní osoby.

## UŽITEČNÉ WEBOVÉ STRÁNKY



Internet může být zdrojem užitečných informací, a to také pro výkon veřejné správy. Proto přinášíme **několik dobrých tipů, co a kde lze na internetu najít (ve vztahu k obsahu tohoto kurzu):**

### OPEN DATA

#### **výhody open dat:**

přednáška britského premiéra Davida Camerona na konferenci TED (dostupné na [http://www.ted.com/talks/david\\_cameron](http://www.ted.com/talks/david_cameron) i s českými titulky)

#### **propojitelná otevřená data:**

experimentálně jsou do propojitelné podoby převáděna data na českých vysokých školách (seznam dat zpracovaných UK Praha je dostupný na <http://opendata.cz/cs/linked-data>)

#### **odborníci na open data:**

z akademické sféry (např. MFF UK a FIS VŠE - <http://opendata.cz>)

z neziskové sféry (např. Fond Otakara Motejla - <http://otevrenadata.cz> a <http://www.nasstat.cz/>)

#### **datová žurnalistika:**

v současnosti nejaktivnější datovou žurnalistiku v ČR má na svědomí redakce iHned.cz; její výstupy jsou dostupné na adrese <http://data.blog.ihned.cz/>

## licence:

ze standardizovaných licencí se jako nejvhodnější pro české podmínky považuje:

- Open Data Commons Public Domain Dedication and Licence (ODC PDDL) – v plném znění na <http://opendatacommons.org/licenses/pddl/1.0/>
- Creative Commons od verze 4.0, která by měla být v českém znění dostupná od září 2014 na <http://www.creativecommons.cz/>

## oficiální státní katalog otevřených dat veřejné správy ČR:

mohl by být spuštěn již od začátku roku 2015, ale jeho spuštění bude pravděpodobně navázáno na chystané legislativní změny, což může termín spuštění odsunout; prozatím je k dispozici například experimentální katalog na adrese <http://cz.ckan.net/>

## ÚŘAD PRO OCHRANU OSOBNÍCH ÚDAJŮ

<http://www.uouu.cz/>

- **judikatura:**

v této části naleznete rozhodnutí soudů, a to jak českých, tak i zahraničních (Evropského soudního dvora nebo Evropského soudu pro lidská práva, ale i některých národních soudů), která se týkají problematiky osobních údajů

- **právní předpisy:**

pod odkazem **Zákon o ochraně osobních údajů** je uvedena historie i aktuální text platného a účinného znění zákona č. 101/2000 Sb., který zahrnuje mezinárodní principy ochrany osobních údajů a k naplnění práva každého na ochranu před neoprávněným zasahováním do soukromí upravuje práva a povinnosti při zpracování osobních údajů

v **Evropské ochraně osobních údajů** naleznete nejvýznamnější mezinárodní i evropské právní předpisy upravující ochranu osobních údajů

na stránce **Oblasti zpracování osobních údajů** jsou uvedeny významné zvláštní zákonné agendy a s nimi související postupy týkající se zpracování osobních údajů

## SVOBODNÝ PŘÍSTUP K INFORMACÍM

na stránkách odboru veřejné správy, dozoru a kontroly Ministerstva vnitra na adrese <http://www.mvcr.cz/odk2/> naleznete rubriku „**Metodická pomoc obcím**“; zde mj. stanoviska odboru a metodická doporučení problematiky poskytování informací:

- Metodické doporučení č. 1 k postupu povinných subjektů podle zákona č. 106/1999 Sb., o svobodném přístupu k informacím - vyřizování žádostí o informace ze správních spisů
- Metodické doporučení č. 2 k postupu povinných subjektů podle zákona č. 106/1999 Sb., o svobodném přístupu k informacím - sazebník úhrad za poskytování informací obcemi jako povinnými subjekty

## SLOVNÍK NEJČASTĚJI POUŽÍVANÝCH POJMŮ VE VEŘEJNÉ SPRÁVĚ

<http://svs.institutpraha.cz/>

- **definice více než 1 600 pojmů** z oblasti veřejné správy

## PORTÁL VEŘEJNÉ SPRÁVY

<http://portal.gov.cz>

- plánovaný **státní katalog open dat** by měl být provozován na Portálu veřejné správy
- vkládání nových záznamů a úprava existujících by měla být prováděna prostřednictvím **datových schránek**



## **DOPORUČENÁ ODBORNÁ LITERATURA**



### **Otevřená data – příležitost pro Českou republiku:**

- <http://www.nasstat.cz/res/data/003406.pdf>

### **Metodika publikace otevřených dat veřejné správy ČR:**

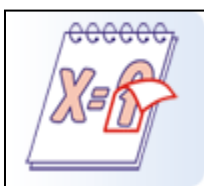
[http://www.korupce.cz/assets/partnerstvi-pro-otevrene-vladnuti/otevrena-data/Metodika\\_Publ\\_OpenData\\_verze\\_1\\_0.pdf](http://www.korupce.cz/assets/partnerstvi-pro-otevrene-vladnuti/otevrena-data/Metodika_Publ_OpenData_verze_1_0.pdf)

### **Publikace, které se věnují problematice ochrany osobních údajů v různých oblastech:**

- P. Mates, K. Neuwirt: „Právní úprava ochrany osobních údajů v ČR“
- JUDr. Pavel Mates: „Ochrana osobních údajů“
- JUDr. Alena Kučerová a kolektiv: „Zákon o ochraně osobních údajů - komentář“
- Miroslava Matoušová a kolektiv: „Ochrana osobních údajů v otázkách a odpovědích“
- JUDr. Jiří Maštalka: „Osobní údaje, právo a my“
- JUDr. Alena Kučerová, Mgr. František Nonnemann: „Ochrana osobních údajů v otázkách a odpovědích“
- JUDr. Alena Kučerová, Mgr. František Nonnemann: „Ochrana osobních údajů v praktických příkladech“
- JUDr. Václav Bartík, JUDr. Eva Janečková: „Ochrana osobních údajů v aplikační praxi“
- JUDr. Václav Bartík, JUDr. Eva Janečková: „Ochrana osobních údajů z pohledu zvláštních právních úprav“
- JUDr. Václav Bartík, JUDr. Eva Janečková: „Zpracování osobních údajů obcemi“
- Úřad pro ochranu osobních údajů: „Příručka o ochraně údajů při jejich zpracování policií“
- JUDr. Jakub Morávek: „Průvodce ochranou osobních údajů v pracovněprávní agendě“
- Pavel Mates: „Ochrana soukromí ve správním právu“

# B. PRAKTICKÁ ČÁST

## SPOJOVAČKY



### Úkol č. 1:

#### SOUVISEJÍCÍ LEGISLATIVA

K červnu 2014 nejsou open data v naší legislativě upravena žádný speciálním právním předpisem. Souvisí však s nimi některé stávající zákony. Spojte čarami čísla s příslušnými (zkrácenými) názvy těchto právních předpisů.

ČÍSLO ZÁKONA	NÁZEV ZÁKONA
101/2000 Sb.	o informačních systémech veřejné správy
121/2000 Sb.	autorský zákon
365/2000 Sb.	o svobodném přístupu k informacím
106/1999 Sb.	o ochraně osobních údajů

## Úkol č. 2:

### OPEN DATA – TERMINOLOGIE

Teoretická část tohoto pracovního sešitu obsahuje mimo jiné také definice několika pojmů souvisejících s problematikou open dat. Prostřednictvím tohoto cvičení si můžete vyzkoušet, zda tuto terminologii ovládáte. Spojte čarami pojem s příslušnou stručnou definicí.

POJEM	DEFINICE
propojitelná open data	obecné označení pro otevřená data
datové schránky	prostředek pro komunikaci s adresáty a s orgány veř. správy
katalogy dat	datové množiny (sady, soubory)
open data	otevřená data s pěti hvězdami
datasets	shromažďují informace o datech

## DOPLŇOVAČKY



### Úkol č. 3:

#### ČLENĚNÍ DAT PODLE 5 HVĚZD

Publikací dat rozumíme uložení dat na internet tak, aby byla volně přístupná. Forem, v jakých jsou data běžně zveřejňována, je celá řada. Z pohledu open dat je můžeme rozdělit například podle členění pěti hvězd. Toto členění uděluje datům 1 až 5 hvězd. Čím více hvězd, tím lépe.

Doplňte k následujícím charakteristikám různých forem dat příslušný počet hvězd.

..... hvězd patří formě propojitelných otevřených dat. Jedná se o podobu, kdy na konkrétní webové adrese jsou informace o jednotlivých záznamech dat.

..... hvězd jsou označením pro mezikrok k propojitelným datům. Tento mezikrok spočívá ve zveřejnění ve formátu založeném na značkovacím jazyce XML, nejčastěji RDF nebo TTL. Data v tomto formátu již mají vlastní URL adresu pro každý záznam, ačkoli na této adrese nic není a v prostředí webu neexistuje. Obdobně jsou využity odkazy na jiná data.

..... hvězd patří datům, která jsou v otevřeném formátu. To znamená, že jsou zcela nezávislá na výrobci konkrétního programového vybavení. Nejtypičtějším příkladem je formát CSV, tedy „čistý“ textový formát s hodnotami oddělenými např. středníkem. Uložení do formátu CSV podporují prakticky všechny kancelářské balíky (MS Office, Open Office).

..... hvězd náleží datům v nestrukturované podobě. Tedy v PDF, v MS Word nebo jiném textovém editoru. Tyto formáty jsou vhodné pouze pro zpracování člověkem, tedy čtení. Automatické zpracování je prakticky vyloučeno. Přestože by byla tato data zveřejněna pod otevřenou licenci, není možné je označit za open data, protože je není možné snadno znovu použít. Může jít ovšem o velmi užitečný doplněk k open datům.

..... hvězd označují data, která jsou nejen zveřejněna pod otevřenou licenci, ale současně jsou i ve strojově zpracovatelném formátu (jako je tabulka MS Excel). Na této úrovni je již možné začít hovořit o open datech.

#### Úkol č. 4:

### PŘÍNOSY OPEN DAT

S ohledem na kontext doplňte do následujícího článku chybějící výrazy.

Užití open dat lze rozdělit do dvou hlavních skupin: 1) datová ..... a 2) webové a mobilní aplikace. První skupina představuje dynamicky se rozvíjející obor, kde v dostupných datech hledají novináři zajímavé informace. V současnosti nejaktivnější je v tomto směru redakce iHned.cz. Webové a mobilní aplikace vytváří zejména soukromé subjekty, které vidí možnost ..... například z reklamy nebo z prodeje mobilních aplikací.

Z pohledu úřadu mohou open data přinést pokles počtu žádostí podle zákona č. .... Sb., o svobodném přístupu k informacím, a dotazů od novinářů, kteří si budou moci rychleji vše dohledat sami.

Z pohledu ekonomiky se jedná o vznik nových ..... míst, nárůst HDP a tím vyšší odvod daní.

## KONTROLNÍ OTÁZKY



Po absolvování tohoto kurzu byste měli umět odpovědět na následující **DESATERO KONTROLNÍCH OTÁZEK**:

1. Co označujeme pojmem „open data“?

2. Jaká open data lze považovat za propojitelná?

3. Jakým datům náleží 3 hvězdy?

4. V čem mohou být open data přínosná?

5. Uveďte příklady zajímavých aplikací open dat.

6. Které kroky by měl zahrnovat proces zveřejnění open dat?

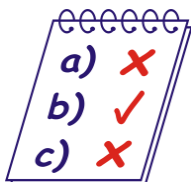
7. Vyjmenujte účinné právní předpisy související s problematikou open dat.

8. Definujte pojem „osobní údaje“.

9. Co by měly obsahovat katalogy dat?

10. Které subjekty v současnosti nabízí bezplatnou pomoc se zveřejňováním open dat?

## TEST



Následující test obsahuje uzavřené otázky. Ke každé testové otázce Vám nabízíme tři varianty odpovědi, z nichž je vždy pouze jedna správná. Ta je součástí nabídky vždy. K žádné otázce v tomto testu nejsou přiřazeny dvě nebo dokonce tři správné odpovědi.

### **1. Jaký formát dat je z pohledu open dat nejlepší?**

- a) CSV – data oddělená středníky
- b) PDF – naskenované tabulky
- c) RDF – data v jazyce XML s URL adresou pro každý záznam

### **2. Jaká data není možné zveřejnit bez anonymizace?**

- a) souhrnný počet obyvatel jednotlivých územních samosprávných celků
- b) hrubé měsíční příjmy jednotlivých domácností s adresami
- c) data o příjmech a výdajích rozpočtu úřadu

### **3. Jaká data mohou být zveřejněna jako open data?**

- a) vše, o co požádají novináři
- b) vše, co můžeme poskytnout v odpovědi na žádost dle zákona č. 106/1999 Sb., o svobodném přístupu k informacím
- c) vše, co máme na úřadě v tabulkových souborech

### **4. Jaký nástroj pomůže zájemcům o data k jejich nalezení?**

- a) katalog open dat
- b) anonymizace dat
- c) napsání licence na míru

**5. Co není výsledkem open dat?**

- a) datová žurnalistika
- b) mobilní a webové aplikace
- c) zrušení tiskového oddělení

**6. Co open data nemohou přinést?**

- a) příjem do rozpočtu z licencí na open data
- b) vyšší informovanost občanů
- c) podporu národní ekonomiky

**7. Co musí umožnit otevřená licence?**

- a) povolit komerční i nekomerční použití
- b) citovat jen se souhlasem úřadu, který data zveřejnil
- c) povolit stažení dat až po podpisu memoranda s Ministerstvem vnitra

**8. Jaký zákon musí být vždy zohledněn při analýze, zda data zveřejnit?**

- a) zákon č. 183/2006 Sb., o územním plánování a stavebním řádu
- b) zákon č. 634/1992 Sb., o ochraně spotřebitele
- c) zákon č. 101/2000 Sb., o ochraně osobních údajů

**9. Jaký krok zveřejnění open dat musí být první?**

- a) zápis do národního katalogu open dat
- b) analýza, jaká data zveřejnit
- c) popis dat

**10. Co může být považováno za osobní údaj?**

- a) cokoli, co identifikuje konkrétní právnickou osobu (společnost)
- b) cokoli, co identifikuje konkrétní orgán veřejné moci
- c) cokoli, co identifikuje konkrétní fyzickou osobu

