

Co jsou otevřená data?

Data zveřejněná na Internetu, která jsou

- úplná,
- snadno dostupná,
- strojově čitelná,
- používající standardy s volně dostupnou specifikací,
- zpřístupněna za jasně definovaných podmínek užití dat s minimem omezení,
- dostupná uživatelům při vynaložení minima možných nákladů.

Požadavky na podmínky užití otevřených dat:

- Neomezují jejich uživatele ve způsobu použití dat.
- Opravňují uživatele k jejich dalšímu šíření.
- Při dalším šíření musí mít všichni uživatelé stejná oprávnění s daty
- nakládat.
- Cílem je umožnit šíření dat pro nekomerční i komerční účely.

K čemu slouží otevřená data?

Otevřená data jsou informace, jejichž zveřejnění může sloužit široké veřejnosti. Data nepublikujeme proto, abychom si mohli „odškrtnout“, že jsme něco udělali pro transparentnost, ale proto, aby je někdo využil. Jedině tak získají přidanou hodnotu.

Jaká data je vhodné zveřejňovat

- 1) Data, která jsou často žádána podle zákona 106/1999. (Ušetříme si práci s odpovídáním.)
- 2) Data, o kterých si myslíme, že by někdo mohl použít (např. k vytvoření nějaké služby, či mobilní aplikace).

Způsoby zveřejňování

Nejjednodušší způsob je zveřejnění dat na webových stránkách v sekci OpenData jako soubor ke stažení. Vhodné jsou soubory např. xlsx nebo csv. Naopak nevhodné jsou např. pdf. Každou takovou tabulku je potřeba opatřit metadaty. Metadaty jsou doplňující informace - stručný popis, datum publikace, kontakt na správce, četnost aktualizace.

Kde taková data vzít

Pokud by kompletace jednoho datasetu (dataset = soubor obsahující data), znamenala měsíční práci jednoho úředníka, pravděpodobně se nevyplatí data zveřejňovat. Pokud ovšem víme, že o dataset stojí místní podnikatel, který nad ním postaví službu lákající občany na náš web, stojí taková investice za úvahu.

Současná praxe otevřených dat

- Pokud více OVM publikuje stejný typ datové sady (např. seznam smluv), každá opakovaně vymýšlí jiný formát (CSV, XLS(X), JSON, XML, ...) a ad-hoc datové schéma.
- Celkově je tedy současná situace pro potenciální uživatele datových sad i pro samotné OVM matoucí a nepřehledná.

- OVM nevědí, jaké datové sady mají publikovat a jak je mají publikovat.
- Uživatelé dat nevědí, kde mají datové sady hledat a je velmi obtížné až nemožné je hromadně zpracovávat. (Různé pojmenování datových sad, různé formáty a datová schémata.)

Proč katalogizovat otevřená data?

Bez katalogu je pro uživatele obtížné vyhledat, jaké datové sady OVM zveřejňují.

Pokud požadovanou datovou sadu uživatel nalezne, je pro něj bez metadat uvedených v katalogu obtížné zjistit

- kdo datovou sadu publikuje,
- zda a jak ji může využít,
- kdy byla publikována,
- jak s ní má pracovat,
- na koho se má obrátit v případě dotazů.

Národní katalog otevřených dat veřejné správy (NKOD)

10. 4. 2015 byl na portálu veřejné správy ČR <http://portal.gov.cz> spuštěn Národní katalog otevřených dat. Jedná se o pilotní provoz. Plný ostrý provoz se očekává Q2/2015

NKOD obsahuje katalogizační záznamy o datových sadách zveřejněných jednotlivými OVM včetně odkazů ke stažení dat.

Data ke stažení mohou být uložena:

- na webu příslušného OVM (poskytovatele dat)
- přímo na PVS v Rejstříku datových zdrojů (součást NKOD)

NKOD umožňuje katalogizaci následujícími způsoby:

- vyplnění formuláře na PVS a jeho odeslání přes ISDS
- automatické plnění metadaty z lokálních katalogů OVM (po předchozí registraci lokálního katalogu v NKOD)

Zveřejňovat mohou:

- jednotlivé datové sady
- odkaz na vlastní lokální katalog

Vzorový publikační plán

Vymezuje doporučené datové sady k publikaci v podobě otevřených dat a určuje podobu publikace.

V čem vzorový publikační plán pomůže?

- Stanoví doporučované datové sady pro všechny OVM v ČR.
- Ujednotí způsob publikace těchto sad napříč OVM (formáty, schémata, umístění v lokálním datovém katalogu/NKOD).
- Usnadní práci jednotlivých OVM při otevírání jejich dat. OVM nemusí analyzovat, jaké datové sady bude publikovat, nemusí navrhnout datové formáty a schémata a ani určovat, jakým způsobem budou datové sady publikovány.

Typy vzorových publikačních plánů

Různé typy OVM spravují různé datové sady a mají různé (technické, personální, ...) možnosti. Proto budou pro různé typy OVM vytvořeny různé vzorové publikační plány:

- ústřední orgány – ministerstva
- ústřední orgány – ostatní
- krajský úřad
- obce s rozšířenou působností
- obecní úřad – ostatní

Napříč všemi obcemi lze identifikovat řadu stejných datových sad, které obce spravují a lze je doporučit k plošné publikaci. Podobná situace je u krajů.

Při sestavování vzorového publikačního plánu postupujeme vzhledem k jejich možnostem a kapacitám. Pečlivě vybíráme seznam datových sad s maximálním potenciálem. Definujeme datová schémata s co nejjednodušším formátem.

Příklady doporučených datových sad

- Seznam smluv / objednávek / faktur
- Nabídka nepotřebného majetku
- Telefonní seznam
- Seznam poradních orgánů
- Volná pracovní místa
- Katalog knihovny
- Pořádané akce
- ...

V případě typu datové sady, která plošně existuje napříč různými OVM (např. seznam smluv či výsledky kontrol) se nevyplatí pracovat pouze s daty jedné OVM. Pokud ale každý OVM vytváří vlastní ad-hoc datové schéma pro datovou sadu, je plošnému využití prakticky zamezeno.

Validace vzorových publikačních plánů

Po vytvoření první verze vzorových publikačních plánů proběhne jejich validace vybranými validátory.

Pro každý druh OVM, pro který vytváříme specifický publikační plán, vybíráme několik validátorů z konkrétních OVM.

Vybraní validátoři budou připomínkovat vzorové publikační plány, zejména sestavení doporučených datových sad a jejich doporučených datových schémat a přínosy a rizika spojená s publikací doporučených datových sad.

Na základě připomínek bude vytvořena druhá verze, která bude vystavena na stránkách MVČR.

Role a činnosti při publikaci a katalogizaci otevřených dat

Role	Činnosti
<p>Poskytovatel dat (vedení)</p> <ul style="list-style-type: none"> - Právnická osoba, která je odpovědná za realizaci publikace (zpřístupnění) datové sady, např. formou souboru stažitelného z webových stránek. 	<ul style="list-style-type: none"> ✓ Rozhodnout o otevírání dat a obsadit roli koordinátora otevírání dat ✓ Schválit interní legislativu (směrnice, opatření, ...) – volitelná činnost ✓ Obsadit další role v kontextu OD ✓ Schválit publikační plán
<p>Koordinátor otevírání dat</p> <ul style="list-style-type: none"> - Osoba, která je v rámci organizace pověřená koordinací a řízením aktivit v oblasti otevřených dat. 	<ul style="list-style-type: none"> ✓ Připravit interní legislativu (směrnice, opatření, ...) - volitelná činnost ✓ Připravit publikační plán ✓ Zveřejnit publikační plán ✓ Rozhodnout o způsobu katalogizace (Národní katalog OD nebo i lokální katalog OD) ✓ Kontrolovat data připravená k publikaci ✓ Zajistit publikaci datových sad a katalogizačních záznamů (metadat) ✓ Komunikovat a reportovat publikaci otevřených dat
<p>Kurátor dat</p> <ul style="list-style-type: none"> - Osoba odpovědná za přípravu a údržbu určených datových sad a jejich katalogizačních záznamů (metadat). 	<ul style="list-style-type: none"> ✓ Navrhnout datové sady k publikaci ✓ Připravit datové sady a katalogizační záznamy k publikaci
<p>Správce katalogu otevřených dat</p> <ul style="list-style-type: none"> - Správce katalogu otevřených dat je subjekt odpovědný za přenos metadat do Národního katalogu otevřených dat. Případně také za přípravu a zajištění provozu lokálního datového katalogu. 	<ul style="list-style-type: none"> ✓ Připravit lokální katalog otevřených dat – volitelná činnost ✓ Spravovat katalogizační záznamy (v Národním katalogu OD nebo i lokálním katalogu OD)
<p>IT specialista</p> <ul style="list-style-type: none"> - Osoba se znalostmi a dovednostmi v oblasti informačních a komunikačních technologií. Poskytuje podporu ostatním rolím při publikaci otevřených dat, vyvíjí a testuje ETL procedury a realizuje převod dat do cílových formátů. 	<ul style="list-style-type: none"> ✓ Spolupracovat na přípravě datové sady ke zveřejnění ✓ Instalovat a provozovat lokální datový katalog – volitelná činnost

Standardizace publikace a katalogizace otevřených dat

Projekt OPLZZ na MV - doba řešení projektu: 1. 2. 2015 – 31. 10.2015

Cílem projektu je implementace principů otevřených dat do prostředí České republiky v souladu se strategickými dokumenty ČR a evropskou legislativou.

Publikace pod otevřenou licenci

Zveřejněná data jsou k ničemu, pokud zájemce nedostane svolení k jejich použití. Abychom zamezili nedorozuměním a nejistotám, je dobré podmínky užití explicitně zveřejnit (ideálně jako metadata).

Licencí pro otevřená data existuje celá řada, odborníci se ale poslední dobou shodují především na Creative Commons 4.0.

Jak použít Creative Commons 4.0

- Publikovat data.
- Na stránce Creative Commons vyplnit „bibliografické údaje“.
- Nezakazovat možnost upravovat dataset a dovolit ho použít pro komerční účely.