

MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY



Standardy publikace a katalogizace otevřených dat

Dušan Chlapek

David Hemala



Obsah prezentace

1. Vymezení otevřených dat
2. Standardy publikace otevřených dat
3. Role a procesy
4. Vzorové publikační plány
5. Datové sady a datové schéma
6. Katalogizační záznam a publikace datové sady
7. Nabídka konzultací



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.
- **Legislativa: Otevřená data jsou pouze ta data, která jsou evidována v Národním katalogu otevřených dat**



Standardy publikace a katalogizace OD VS ČR

OTEVŘENÁ DATA Standardy otevřených dat

Umístění: Standardy publikace a katalogizace otevřených dat veřejné správy ČR

Standardy publikace a katalogizace otevřených dat veřejné správy ČR

Tento portál je určen orgánům veřejné správy, které plánují otevřít svá data. Popisuje doporučený postup otevírání dat a nabízí vzorové publikační plány, které doporučují (pro jednotlivé typy orgánů veřejné správy) konkrétní datové sady k publikaci ve formátu otevřených dat. Portál rovněž informuje o úkolech připravených k tématu otevřených dat. Standardy publikace a katalogizace otevřených dat veřejné správy jsou k dispozici i jako dokument [Standardy publikace a katalogizace otevřených dat veřejné správy ČR](#).

Jak otevírat data?

- Vytvoření publikačního plánu
- Publikace datových sad v otevřené podobě
- Zprovoznění lokálního katalogu otevřených dat a jeho registrace v NKO

Co má obsahovat vzorový publikační plán?

- Ministerstva
- Centrální orgány
- Kraje
- Obce s rozšířenou působností
- Ostatní obce

Nástroje

- Národní katalog otevřených dat
- CKAN
- DKAN

Vzdělávání

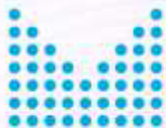
Proběhla školení k otevřeným datům a v listopadu byla uspořádána konference.

Standardy publikace a katalogizace otevřených dat VS ČR vznikly jako výstup projektu „Implementace strategií v oblasti otevřených dat veřejné správy ČR“ CZ.1.04/4.1.00/D9.00001. Projekt je financován v rámci Operačního programu lidské zdroje a zaměstnanost z Evropského sociálního fondu a státního rozpočtu ČR. Realizátorem projektu je Ministerstvo vnitra ČR.

esf evropský sociální fond v ČR **EVROPSKÁ UNIE** **OPERAČNÍ PROGRAM LIDSKÉ ZDROJE A ZAMĚSTNANOST** **PODPORUJEME VAŠI BUDOUCNOST** www.esfcr.cz

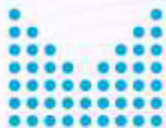
start.txt - Poslední úprava: 2016/05/14 14:03 (upraveno měno Dotu/Wiki)

<http://opendata.gov.cz>



Standardy pro publikaci otevřených dat

- Standardy se rozvíjeny také díky:
 - Implementací standardů do praxe (např.: ČOI, Moravskoslezský kraj, ČSSZ, NKÚ, aj.)
 - Validacím na workshopech - zejména připomínkováním vzorových publikačních plánů a možností integrace
 - Zpětné vazby ze školení a praktických cvičení
- Standardy budou i nadále rozšiřovány:
 - Plánované pracovní workshopy
 - Konference a školení
 - Sledování vývoje mezinárodních standardů
 - Legislativa
 - Monitoringem katalogizace datových sad do národního katalogu
 - Konzultacemi pro jednotlivé subjekty



Standardy pro publikaci otevřených dat

- Portál opendata.gov.cz
 - postupy a návody pro přípravu publikačních plánů,
 - postupy a návody pro publikaci a katalogizaci datových sad,
 - postupy a návody pro správu lokálního katalogu,
 - doporučené role, jejich činnosti a zodpovědnosti,
 - slovník pojmů a jejich kontext.
- Vzorová směrnice upravující role a jejich odpovědnosti při publikaci a katalogizaci otevřených dat
 - usnadní zavedení otevřených dat do interních procesů a jejich řídicí dokumentace,
 - východisko pro přípravu interní legislativy jednotlivých institucí.
- Předzpracované datové sady
 - vzorové publikační plány
 - seznam doporučených datových sad k otevření, příp. oblastí,
 - detailní popis doporučených datových sad,
 - vzory katalogizačních záznamů,
 - typické přínosy publikace doporučených datových sad,
 - typická rizika spojená s publikací doporučených datových sad a doporučené kroky k jejich zmírnění,
 - příklad publikovaných otevřených dat.

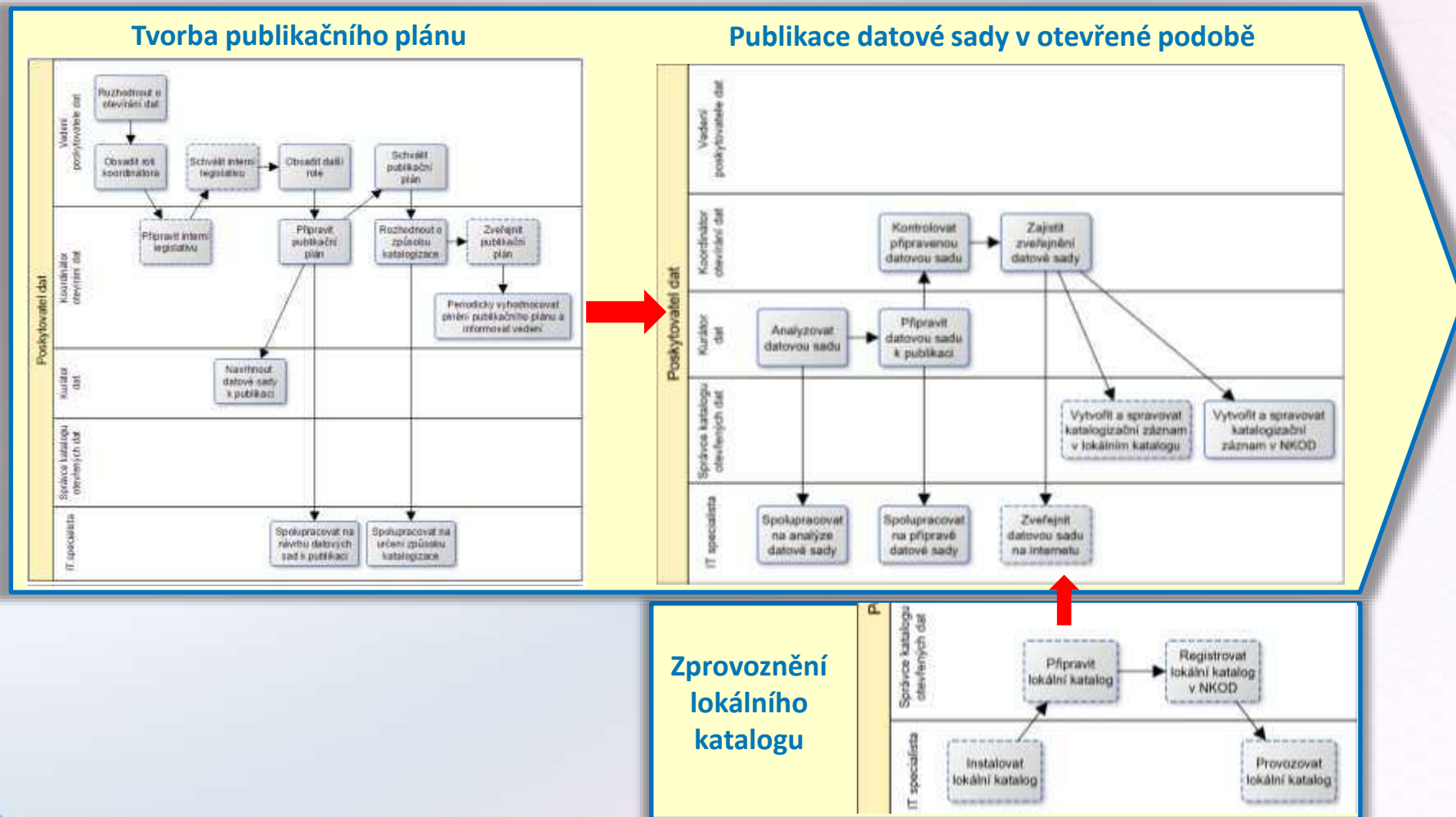


Role a činnosti při publikaci a katalogizaci otevřených dat

Role	Činnosti
Poskytovatel dat (vedení)	<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
Koordinátor otevírání 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
Kurátor dat	<ul style="list-style-type: none">✓ Navrhnout datové sady k publikaci✓ Připravit datové sady a katalogizační záznamy k publikaci
Správce katalogu otevřených dat	<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)
IT specialista	<ul style="list-style-type: none">✓ Spolupracovat na přípravě datové sady ke zveřejnění✓ Instalovat a provozovat lokální datový katalog – volitelná činnost



Stručný přehled procesů pro publikaci otevřených dat



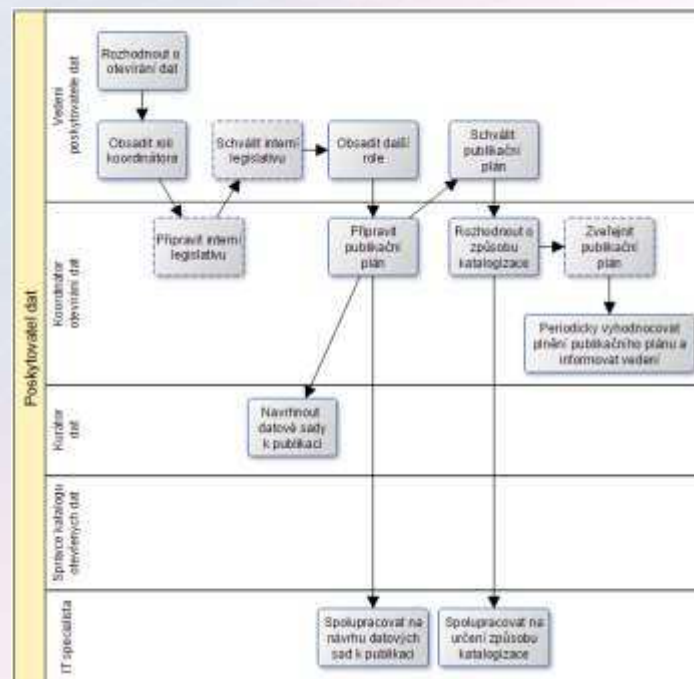


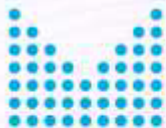
Popsané procesy a postupy spojené s vytvořením publikačního plánu

- Přípravné procesy na úrovni vedení
 - Rozhodnutí (politické) o otevírání dat
 - Obsazení rolí
 - Příprava a schválení interní legislativy
- Návrh datových sad a příprava publikačního plánu
- Schválení publikačního plánu
- Rozhodnutí o způsobu katalogizace otevřených dat
- Zveřejnění publikačního plánu
- Periodické vyhodnocování publikačního plánu a informování vedení

Legenda:

- obdélník s plným okrajem značí povinnou činnost
- obdélník s přerušovaným okrajem značí nepovinnou činnost
- šipky značí návaznost mezi činnostmi





Vzorové publikační plány

- Vzorový publikační plán vymezuje doporučené oblasti či doporučené datové sady k publikaci otevřených dat včetně typických rizik a přínosů
 - Vzorové publikační plány by měly být vytvořeny pro následující typy orgánů VS:
 - **Ústřední orgán státní správy – ministerstvo**
 - **Ústřední orgán státní správy – ostatní**
 - **Krajský úřad**
 - **Obce s rozšířenou působností**
 - **Obecní úřad – ostatní**
- Vzorový publikační plán zejména obsahuje:
 - Seznam doporučených datových sad k otevření, příp. oblastí, ze kterých datové sady vybírat
 - Detailní popis doporučených datových sad
 - Vzory katalogizačních záznamů
 - Typická rizika spojená s publikací doporučených datových sad a doporučené kroky k jejich zmírnění
 - Typické přínosy publikace doporučených datových sad a doporučené kroky pro jejich dosažení (upozornit u jednotlivých datových sad na možnost, jak přes publikaci otevřených dat mohou subjekty splnit svou zákonnou povinnost – např. „dálkový přístup“ k určeným datům)
- Vzorové publikační plány jsou k dispozici <http://opendata.gov.cz> v části „Vzorové publikační plány“.



Ukázka vzorového publikačního plánu

- Pro každou datovou sadu jsou v plánu uvedeny:
 - **název datové sady,**
 - **popis datové sady,**
 - **kurátor dat** zodpovědný za analýzu a přípravu datové sady,
 - stupeň otevřenosti,
 - datový formát,
 - **periodicita publikace** (aktualizace) datové sady,
 - podmínky užití datové sady,
 - plánovaný termín první publikace datové sady (ve zvoleném stupni otevřenosti),
 - **přínosy a rizika publikace datové sady** ve formátu otevřených dat.

OTEVŘENÁ DATA Standardy otevřených dat

Umožní: Standardy publikací a katalogizace otevřených dat veřejné správy ČR » Vzorové publikační plány » Vzorový publikační plán pro obce s rozšířenou působností

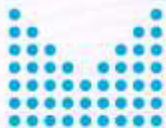
Vzorový publikační plán pro obce s rozšířenou působností

[Kompletní vzorový publikační plán ke stažení](#)

Seznam datových sad ze vzorového publikačního plánu

Název	Popis
Akce dobrovolných hasičů	Informace o pořádaných akcích (společenských, sportovních, ...) dobrovolnými hasiči
Aktuální opravy, rekonstrukce, uzavírky ulic	Aktuální opravy, rekonstrukce, uzavírky ulic
Aktuální poloha vozidel technické služby a zimní údržby	Aktuální poloha vozidel technické služby a zimní údržby

Název datové sady	Popis datové sady	Kurátor dat	Stupeň otevřenosti	Datový formát	Pravidelnost	Periodičita publikace	Podmínky užití	Plánovaný termín publikace
Akce dobrovolných hasičů	Informace o pořádaných akcích (společenských, sportovních, ...) dobrovolnými hasiči		3	CSV, XLS	přibližně	přibližně / ročně		
Aktuální opravy, rekonstrukce, uzavírky ulic	Aktuální opravy, rekonstrukce, uzavírky ulic	Odbor dopravního agenta	3	CSV, GPS lokace, HTML, webcam	pravidelně	reálné		
Aktuální poloha vozidel technické služby a zimní údržby	Aktuální poloha vozidel technické služby a zimní údržby			KML, GPS lokace	pravidelně	v reálném čase		
Bezbariérové přístupy objektů	Informace o bezbariérovém přístupu do objektů a způsobu zajištění pro zdravotně postižené.		3	PDF, CSV, HTML	pravidelně	městně		
Cestovní náklady zahraničí	Roční cestovní náklady souvisejících se zahraniční služební cestou.		3	CSV	pravidelně	ročně		



Datové sady a jejich přínosy

- Analýzu provádí koordinátor otevírání dat ve spolupráci s kurátory dat a je doporučen následující postup:
 - Pro datové sady uvedené ve vzorovém publikačním plánu jsou přínosy identifikovány a lze je snadno převzít.
 - Pokud jsou do seznamu přidány vlastní datové sady, ohodnocení je vhodné provést pásmově (0 – žádný přínos, 1 – nízký přínos, 2 – střední přínos, 3 – významný přínos).

Seznam přínosů

1. Posílení transparentnosti
2. Podpora ekonomického růstu
3. Zlepšení služeb veřejné správy / zlepšení kvality života
4. Podpora opětovného použití dat
5. Zlepšení vnímání veřejné správy veřejností
6. Zlepšení procesů a dat veřejné správy, zlepšení komunikace a spolupráce veřejné správy
7. Zamezení chyb vzniklých při práci s daty
8. Snížení počtu dotazů dle zák. č. 106/1999 Sb.
9. Pořádek ve vlastních datech
10. Zvýšení hodnoty dat

Seznam rizik

1. Zveřejnění dat v rozporu se zákonem
2. Porušení ochrany obchodního tajemství
3. Porušení ochrany osobních údajů a dobrého jména
4. Zveřejnění nevhodných dat či informací
5. Dezinterpretace dat
6. Absence konzumentů dat
7. Překrývání dat
8. Ohrožení bezpečnosti státu / majetku / osob

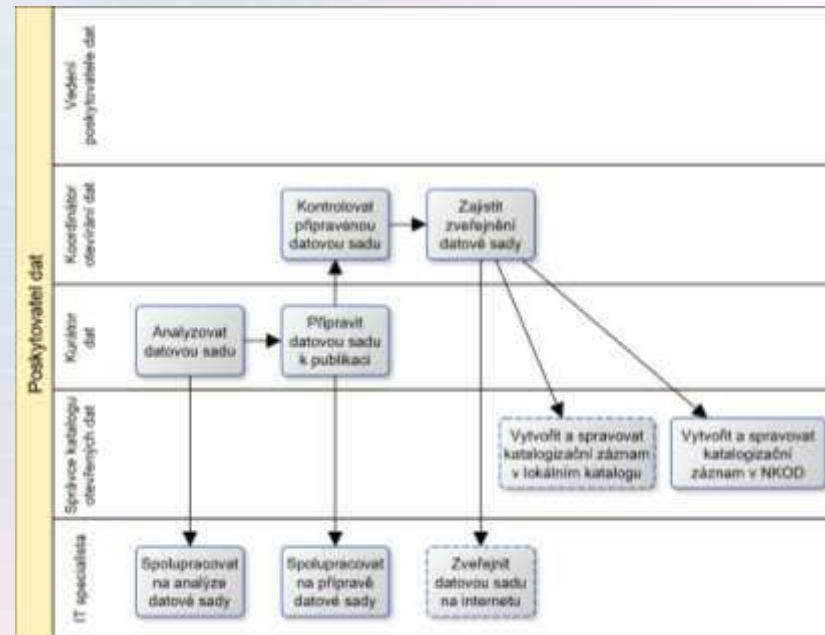


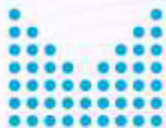
Popsané procesy a postupy spojené s publikací datové sady

- Po vytvoření publikačního plánu následuje postupné publikování jednotlivých datových sad dle stanoveného harmonogramu.
- Postup:
 - analýza datové sady a návržení:
 - stupně otevřenosti, na kterém bude datová sada zveřejněna (minimální stupeň 3 hvězd),
 - datového schématu určující strukturu, ve které bude datová sada zveřejněna,
 - katalogizačního záznamu datové sady,
 - příprava datové sady k publikaci,
 - publikace datové sady,
 - katalogizace datové sady.

Legenda:

- obdélník s plným okrajem značí povinnou činnost
- obdélník s přerušovaným okrajem značí nepovinnou činnost
- šipky značí návaznost mezi činnostmi





Použití vzorového datového schématu

- Předpřipraveno o datové sadě:
 - Dokumentace datového schématu pro formát CSV
 - Strojově čitelné datové schéma pro formát:
 - XML (XML Schema),
 - CSV (JSON Table Schema),
 - Metada datové sady (CSV)
 - Metada distribuce datové sady (CSV)

Co jsou otevřená data?
Stupně otevřenost
Důležité pojmy
Postup otevření dat
Vyročení publikačního plánu
Publikace datové sady v otevřené podobě
Zprovoznění lokálního katalogu otevřených dat (přílohy)
Vzorové publikační plány
Město Praha
Centrální orgány
Města
Obce s rozšířenou působností
Ostatní obce
Nástroje
Vzdělávání
Kontakt

Vzor vzorový publikační plán obce s rozšířenou působností

Vzorový publikační plán pro obce s rozšířenou působností

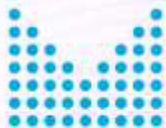
Kompletní vzorový publikační plán ke stažení

Seznam datových sad ze vzorového publikačního plánu

Název	Popis
Akce dobrovolných hasičů	Informace o pořádaných akcích (společenských, sportovních, ...) dobrovolnými hasiči.
Aktuální opravy, rekonstrukce, uzavírky ulic	Aktuální opravy, rekonstrukce, uzavírky ulic.
Aktuální poloha vozidel technické služby a zimní údržby	Aktuální poloha vozidel technické služby a zimní údržby.
Bezbariérové přístupy objektů	Informace o bezbariérovém přístupu do objektů a způsobu zpřístupnění pro zdravotně postižené.
Cestovní náklady - zahraničí	Rozpis cestovních nákladů souvisejících se zahraniční služební cestou.

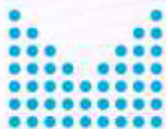
Dokumentace datového schématu pro formát CSV

Název atributu	Datový typ atributu	Popis atributu
publikující_organizace_ic	Řetězec	Organizace, která poskytuje data o přístupnosti objektu, IČ.
publikující_organizace_nazev	Řetězec	Organizace, která poskytuje data o přístupnosti objektu, Název.
objekt	Řetězec	Název objektu s bezbariérovým přístupem.
id_objektu	Řetězec	Identifikátor objektu.
adresa_kod_adm	Řetězec	Adresa objektu. ID adresního místa v Registru územních identifikátorů, adres a nemovitostí.
adresa_kod_obce	Řetězec	Adresa objektu. Obec - LAU 2 kód.
adresa_nazev_obce	Řetězec	Adresa objektu. Název obce.
adresa_nazev_morc	Řetězec	Adresa objektu. Název městského obvodu/městské části.



Použití vzorového datového schématu

- Praktická doporučení:
 - Nepřejmenovávat ani jinak nepředefinovávat datové prvky definované předpřipraveným datovým schématem
 - Pokud některý datový prvek ve své datové sadě evidujete s jiným názvem, doporučujeme jej ve vaší datové sadě přejmenovat na název definovaný předpřipraveným datovým schématem
 - Pokud některý datový prvek ve své datové sadě neevidujete, nemusíte jej doplňovat
 - Pokud některý datový prvek ve své datové sadě evidujete, ale strukturujete jej jinak, doporučujeme jej ve vaší datové sadě přestrukturovat. Pokud však na to nemáte kapacity, upravte datové schéma tak, aby odpovídalo vaší struktuře.
- Pokud využijete předpřipravené datové schéma beze změny, nemusíte jej kopírovat.
V katalogizačním záznamu datové sady (kde musí být datové schéma uvedeno), se budete na předpřipravené datové schéma pouze odkazovat.
- Pokud předpřipravené datové schéma potřebujete jakkoliv změnit, musíte si vytvořit vlastní kopii a tu upravit požadovaným způsobem.



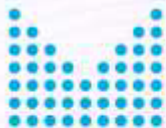
Primitivní datové typy

Datový typ	Popis	Specifikace – Předpokládá se, že budou odpovídat specifikaci datových typů z jazyka XML Schema, a to při použití i v jiných formátech než XML, tj. CSV, XML, RDF, JSON.
AnoNe	Tento datový typ reprezentuje binární hodnotu. Může nabývat hodnot true nebo false.	Předpokládá se, že bude odpovídat specifikaci datového typu xsd:boolean.
Celé číslo	Tento datový typ reprezentuje celočíselný údaj (bez desetinných míst).	Předpokládá se, že bude odpovídat specifikaci datového typu xsd:integer.
Čas	Tento datový typ reprezentuje časový údaj (bez data).	Předpokládá se, že bude odpovídat specifikaci datového typu xsd:time.
Datum	Tento datový typ reprezentuje datum (bez času).	Předpokládá se, že bude odpovídat specifikaci datového typu xsd:date.
Desetin né číslo	Tento datový typ reprezentuje číselný údaj s desetinnými místy.	Předpokládá se, že bude odpovídat specifikaci datového typu xsd:decimal.
Řetězec	Tento datový typ odpovídá kratšímu řetězci znaků (typicky kód, ne souvislý text) v kódování UTF-8.	Předpokládá se, že bude odpovídat specifikaci datového typu xsd:string.
Text	Tento datový typ odpovídá delšímu řetězci znaků (typicky souvislý text, nikoliv kód) v kódování UTF-8.	Předpokládá se, že bude odpovídat specifikaci datového typu xsd:string.
URL	Tento datový typ reprezentuje URL – Uniform Resource Locator.	Předpokládá se, že bude odpovídat specifikaci datového typu xsd:anyURI.



Příprava katalogizačního záznamu

- Struktura katalogizačního záznamu odpovídá evropskému standardu DCAT-AP, který rozlišuje datovou sadu a distribuci datové sady.
 - Datová sada je tvořena jednou či více distribucemi.
 - Distribuce obsahuje samotná data. Může to být datový soubor určený ke stažení nebo to může být datové API, které umožňuje přímé dotazování do obsahu datové sady.
- Katalogizační záznam se dělí na dvě části:
 - První část obsahuje metadata o samotné datové sadě. Jedná se např. o název a popis datové sady.
 - Druhá část obsahuje metadata o distribucích datové sady (datových zdrojích). Jedná se např. o specifikaci datového formátu datového zdroje či o jeho datové schéma.
- Pro doporučené datové sady ze vzorového publikačního plánu jsou již katalogizační záznamy předpřipraveny.
- Další podrobnosti v prezentaci „Lokální katalogy otevřených dat“ od Jakuba Klímka



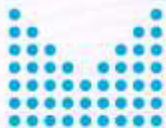
Publikace datové sady

- Před publikací se provede kontrola kompletnosti materiálů:
 - informace o zvoleném stupni otevřenosti datové sady,
 - strojově čitelný soubor s definicí datového schématu datové sady (pokud standardy vyžadují jeho vytvoření),
 - strojově čitelný datový soubor s obsahem datové sady či datový soubor s aktualizací obsahu datové sady,
 - informace o tom, že proběhla validace datového souboru (tj. kontrola, že datový soubor dodržuje všechna pravidla daná zvoleným datovým formátem a že jeho struktura odpovídá definovanému datovému schématu),
 - katalogizační záznam s vyplněnými všemi aktuálně známými atributy.
- V případě zveřejnění datových souborů tvořících datové sady na svých webových stránkách je nutné zveřejnit následující soubory na webu:
 - datový soubor s obsahem datové sady či datový soubor s aktualizací obsahu datové sady,
 - soubor s definicí datového schématu datové sady.
- Správce katalogu otevřených dat založí (v případě prvního zveřejnění datové sady) či upraví (v případě opakovaného zveřejnění) katalogizační záznam v určeném katalogu.
- Portál obsahuje podrobné návody jak:
 - zprovoznit a registrovat lokální katalog (podrobnosti v prezentaci „Lokální katalogy otevřených dat“ od Jakuba Klímka)
 - zveřejnit, editovat a vymazat datovou sadu v NKOD.



Možnost konzultace oha@mvcrcz

- Nabídka konzultací zejména při:
 - Analýze dat a tvorbě publikačního plánu
 - podpora při volbě stupně otevřenosti
 - podpora při validaci přínosů a rizik datových sad
 - podpora během tvorby harmonogramu otevírání dat
 - Tvorbě datových sad
 - podpora definování struktury dat
 - Publikaci datových sad
 - podpora při tvorbě a katalogizaci datových sad a případně jejich distribucí
 - konzultace způsobu spravování katalogizačního záznamu
 - Použití lokálního katalogu



MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY

isss[®]
LOCAL AND REGIONAL
INFORMATION SOCIETY

Ing. Dušan Chlapek, Ph.D.
chlapek@vse.cz

Ing. David Hemala
david.hemala@mpsv.cz

DĚKUJEME ZA POZORNOST