

Materiál Ministerstva vnitra



Export z Národní architektury eGovernmentu ČR

Obsah

Druh=Tématická oblast/Funkční celek| Název=Veřejný datový fond| Popis=Veřejný datový fond.|
Platípro=všechny OVM a OVS| Spravuje=Tomáš Šedivec| Gestor=Ministerstvo vnitra, Odbor hlavního architekta|
Platíkedni=30. září 2019

Veřejným datovým fondem (dále také jako VDF) se rozumí celek datového fondu ČR, obsahující údaje zveřejňované orgány veřejné správy jako veřejné rejstříky a údaje zveřejňované jako otevřená data s podporou centrálního nástroje, kterým je Národní katalog otevřených dat (NKOD).

Veřejné rejstříky (VDF-VR)

Veřejnými rejstříky právnických a fyzických osob podle zákona č. 304/2013 Sb., Zákon o veřejných rejstřících právnických a fyzických osob (dále jen „veřejný rejstřík“), se rozumí:

- spolkový rejstřík,
- nadační rejstřík,
- rejstřík ústavů,
- rejstřík společenství vlastníků jednotek,
- obchodní rejstřík a
- rejstřík obecně prospěšných společností.

Do veřejného rejstříku se zapisují zákonem stanovené údaje o právnických a fyzických osobách (dále jen „zapsaná osoba“). Veřejný rejstřík je informačním systémem veřejné správy. Veřejný rejstřík je veden v elektronické podobě. Veřejný rejstřík vede soud (dále jen „rejstříkový soud“).

Ministerstvo financí uveřejňuje způsobem umožňujícím dálkový přístup informace o osobách zapsaných v České republice a údaje o tom, ve kterém veřejném rejstříku jsou tyto osoby zapsány. Ministerstvo financí umožňuje získat o údajích vedených ve veřejných rejstřících elektronický opis.

Údaje z Obchodního rejstříku společně s údaji z Registru živnostenského podnikání (RŽP) a Registru ekonomických subjektů (RES), Registru plátců spotřební daně (SD) a dalších zdrojů publikuje Ministerstvo financí prostřednictvím Administrativního registru ekonomických subjektů (ARES). ARES je IS, který zpřístupňuje veřejné údaje o ekonomických subjektech z informačních systémů (zdrojů) veřejné správy. Obsahuje údaje ze základních (majoritních) zdrojů, které jsou formou odkazů doplněny údaji z dalších zdrojů. Při zpracování se používají též kontrolní zdroje.

Veřejné rejstříky a vzdálený přístup k jejich údajům prostřednictvím ARES předcházely Základním registrům a musí být s nimi sladěny a integrálně nově zařazeny do celkové koncepce eGovernmentu. Koncepce rozvoje veřejných rejstříků a jejich IS musí tedy naplňovat přinejmenším následující požadavky:

1. Všechny rejstříky musí obsahovat ztotožněné údaje FO a PO, pokud je bylo proti čemu ztotožnit (u rezidentů).
2. Všechny rejstříky vedle toho, že jsou publikovány na internetu, musí být dostupné jako otevřená data. A to přímo, nebo prostřednictvím OD k ARES.
3. Všechny rejstříky musí být pro potřeby agend OVS dostupné přes eGSB, a to buď sdruženě, prostřednictvím ARES, nebo přímo, pokud jejich veřejné údaje nejsou v ARES zahrnuty.
4. ARES, jako klíčový veřejný rejstřík s velkým hospodářským dopadem, musí být průběžně rozvíjen a inovován tak, aby odpovídal aktuálním potřebám občanské i podnikové veřejnosti ČR.

Otevřená data (VDF-OD)

Otevřenými daty se v rámci veřejného datového fondu se rozumí celek obsahující údaje zveřejňované orgány veřejné správy jako otevřená data a Národní katalog otevřených dat (NKOD). Otevřená data jsou:

1. Volně přístupná na webu jako datové soubory ke stažení ve strojově čitelném a otevřeném formátu - CSV, XML, JSON, RDF a další formáty s otevřenou specifikací

2. Opatřená podmínkami užití neomezuujícími jejich užití, viz návod na stanovení podmínek užití
3. Evidovaná v Národním katalogu otevřených dat (NKOD) jako datové sady opatřené přímými odkazy na datové soubory, které je tvoří
4. Úplný obsah databáze nebo agregovaná statistika
5. Opatřená dokumentací
6. Připravena s cílem co nejsnazšího strojového zpracování programátory apod.
7. Opatřená kontaktem na kurátora pro zpětnou vazbu (chyby, žádost o rozšíření, apod.)
8. Jsou publikovány dle otevřených formálních norem ve smyslu § 4b odst. 1 zákona č. 106/1999 Sb. o svobodném přístupu k informacím.

Pokud vaše datová sada nesplňuje všechny uvedené podmínky, nejedná se o otevřená data české veřejné správy podle zákona č. 106/1999 Sb., o svobodném přístupu k informacím, který definuje otevřená data ve svém § 3 odst. 11.

Otevřenými daty zejména není:

1. odkaz na vyhledávací formulář určený pro koncové uživatele
2. odkaz na novou stránku s dalšími informacemi
3. odkaz na veřejné mapové rozhraní GIS
4. odkaz na API umožňující přístup k jednotlivým záznamům či vyhledávání záznamů, nikoliv export kompletních dat
5. datový soubor ve formátu PDF
6. datový soubor ve formátu XLS formátovaný pro tisk nebo obsahující výpočty
7. datový soubor v pseudo-CSV formátu (např. jiný oddělovač než „“)

Pro zajištění publikace otevřených dat z IS je třeba:

1. Zajistit možnost získávání kompletních dat z IS v podobě datových souborů ve strojově čitelném a otevřeném formátu
2. Zajistit publikaci datových souborů na webu organizace nebo v jejím lokálním katalogu otevřených dat
3. Opatřit je lidsky čitelnou dokumentací a strojově čitelným schématem
4. Opatřit je otevřenými podmínkami užití
5. Zaregistrovat je v Národním katalogu otevřených dat (NKOD)

NKOD je nástroj, ve kterém jednotlivé orgány veřejné správy katalogizují jimi zveřejňovaná otevřená data a jiné orgány veřejné správy a veřejnost v něm otevřená data vyhledávají a získávají k nim přístup.

Existence NKOD a povinnost jednotlivých orgánů veřejné správy v něm katalogizovat svá otevřená data je dána zákonem č. 106/1999 Sb., o svobodném přístupu k informacím, který zavádí definici otevřených dat, NKOD a zmocňuje vládu vydat nařízení o seznamech, evidencích či registrech, jejichž veřejný obsah je povinně zveřejňován jako otevřená data, § 5 odst. 7 tohoto zákona pak zmocňuje orgány veřejné správy zveřejňovat i další informace jako otevřená data.

Nařízení vlády č. 425/2016 Sb. o seznamu informací zveřejňovaných jako otevřená data stanovuje, jaké informace mají být zveřejňovány orgány veřejné správy jako otevřená data povinně.

Pro naplnění závazků, které vyplývají ze strategických dokumentů, bude zveřejňování otevřených dat povinné pro všechny správce ISVS. Pro zamezení vytváření duplicitních informací ve veřejné správě, které mohou být zveřejňovány, bude zavedena povinnost orgánů veřejné správy je sdílet jako otevřená data.

Průběžná změna legislativního prostředí musí vést k postupnému rozšiřování povinností orgánů veřejné správy zveřejňovat informace reprezentované jako údaje evidované v ISVS, jejichž zveřejnění je možné, nebo jejich anonymizovanou podobu, souhrn či statistiku jako otevřená data s cílem dosáhnout plošné povinnosti. Je nutné výslovně zanést možnost zveřejňovat informace jako otevřená data do zvláštních předpisů upravujících zveřejňování údajů, v jejichž případě se postup podle zákona o svobodném přístupu k informacím nepoužije, např. do zákona o právu na informace o životním prostředí.

Je též nutné legislativně ukotvit povinnost orgánů veřejné správy využívat při výkonu svých agend otevřená data poskytovaná jinými orgány veřejné správy. To se týká především číselníků, kdy jsou stejné číselníky vytvářeny různými orgány veřejné správy.

Různá kvalita otevřených dat komplikuje jejich opakovatelnou použitelnost. Je též téměř nemožné sdílet veřejné údaje mezi orgány veřejné správy jako otevřená data, neboť neexistují žádné garance dostupnosti. Pro zlepšení současného stavu se VDF stane virtuálním distribuovaným datovým prostorem, ve kterém orgány veřejné správy sdílí s veřejností i mezi sebou navzájem veřejné údaje evidované v jimi spravovaných ISVS v podobě kvalitních otevřených dat. VDF-OD bude mít následující součásti:

- *Národní katalog otevřených dat (NKOD)* – bude zvýšena jeho uživatelská přívětivost i na bázi sémantického slovníku pojmů. Bude podporovat standard DCAT-AP[1].
- *Nástroj pro správu sémantického slovníku pojmů (NSSSP)* – umožní správu sémantického slovníku pojmů všemi orgány veřejné správy a popisovat jejich otevřená data za účelem sémantické harmonizace otevřených dat.
- *5* (pětihvězdičkový) veřejný datový fond (5VDF)* – bude podmnožinou otevřených dat, které jsou zveřejněny jako propojená otevřená data. Jeho součástí bude index datových entit, který pro každou datovou entitu reprezentovanou v 5VDF poskytuje základní údaje z 5VDF a odkazy na různé její reprezentace v otevřených datech jednotlivých orgánů veřejné správy v 5VDF.

Budou zajištěny kapacity pro vytvoření nových komponent VDF-OD. Softwarové nástroje pro realizaci VDF budou vytvořeny jako open-source s maximálním využitím existujících open-source komponent.

Bude vytvořeno HW a SW prostředí pro provoz komponent VDF-OD. Pro potřeby orgánů veřejné správy, které budou zveřejňovat otevřená data v 5VDF bude vybudováno sdílené úložiště 5VDF, které umožní zajistit a garantovat dostupnost pro ostatní orgány veřejné správy.

HW a SW prostředí pro VDF-OD bude zajištěno v NDC nebo eGovernment Cloudu. Správcem vytvořených prostředí bude Ministerstvo vnitra.

Možnosti přípravy IS pro otevřená data

Zásadní je mít přístup k datům IS. Provozovaný IS tedy musí:

1. umožňovat přístup k databázi nebo
2. mít možnost stahovat volitelně strukturovaná data (tabulky) z reportingového modulu systému nebo
3. nabídnout API, ze kterého se dají pravidelně získávat kompletní data v podobě datových souborů

Export dat nebo API v otevřeném formátu - preferované řešení

IS má rozhraní pro pravidelný export dat nebo neomezené API, které organizace může vytěžit a získat kompletní data v jednom z otevřených formátů.

- Pro tabulková data je to CSV
- Pro hierarchická data je to XML nebo JSON
- Pro grafová data je to RDF
- Pro geodata je to jeden z otevřených formátů pro geodata - GeoJSON, ESRI Shapefile, OGC GML, OGC GeoPackage.

Struktura dat je zdokumentována jednak lidsky čitelným dokumentem a jednak strojově čitelným schématem.

- Pro CSV je to schéma CSV on the Web
- Pro XML je to XML Schema
- Pro JSON je to JSON schema
- Pokud je použit vlastní slovník pro RDF, je popsán pomocí RDFS nebo OWL

V takovém případě jsou získaná kompletní data z IS připravena k publikaci formou otevřených dat.

Export dat nebo API v proprietárním formátu

Pokud se jedná o rozšíření existujícího IS, který neumožňuje export dat nebo nenabízí API ve strojově čitelném a otevřeném formátu a takovou úpravu nelze ve Vašem IS provést, využije se stávající export či API, které váš IS již umí (např. do MS Excel), a tento výstup se dále zpracuje do otevřeného formátu pomocí dalších nástrojů tak, aby bylo dosaženo stavu jako v případě přímého exportu do otevřeného formátu.

Následná příprava dat k publikaci v podobě otevřených dat

Data získaná z IS jedním z popsaných způsobů je následně třeba publikovat jako otevřená data. To znamená minimálně:

1. V případě API zajistit jeho vytěžení pro získání kompletních dat k publikaci (tj. případná přímá publikace API nenaplnuje podmínky otevřených dat)
2. Zajistit pravidelnou aktualizaci získaných dat (dle charakteru dat to může být ve frekvenci např. denně, měsíčně nebo ročně)
3. Publikovat získaná data na web ke stažení a následně publikovat každou jejich aktualizaci
4. Opatřit je dokumentací, podmínkami užití a kontaktem na kurátora
5. Katalogizovat je v Národním katalogu otevřených dat (NKOD)

K tomu lze využít nástrojů pro přípravu, publikaci a katalogizaci otevřených dat, jako je třeba LinkedPipes ETL.

Publikace otevřených dat by měla být zajištěna koncepčně na úrovni celé organizace. Kompletní postupy jsou k dispozici na Portálu otevřených dat, naleznete zde rovněž informace o školeních a workshopech, které Ministerstvo vnitra k této problematice poskytuje.

Pro příklad vhodného způsobu zveřejnění datové sady v lokálním katalogu se lze podívat na katalog České správy sociálního zabezpečení (ČSSZ).

Ochrana osobních údajů

Pokud jsou předmětem evidence informačního systému osobní údaje ve smyslu zákona č. 101/2000 Sb., o ochraně osobních údajů a nařízení (EU) č. 2016/679, Obecné nařízení o ochraně osobních údajů (GDPR), neznamená to, že nelze ze systému publikovat otevřená data. Ve formátu otevřených dat lze zveřejnit následující:

1. Pakliže se jedná o veřejnou evidenci či rejstřík a zvláštní právní předpis nařizuje zveřejnění informací, lze zveřejnit osobní údaje v podobě otevřených dat.
2. Anonymizace či Pseudonymizace. Z dat se odstraní osobní údaje a případně se nahradí bezvýznamovým umělým identifikátorem. Data bez osobních údajů se pak mohou zveřejnit v podobě otevřených dat. Ovšem pozor, v závislosti na charakteru dat je třeba zkontrolovat, zda data ve své kombinaci neumožňují identifikaci konkrétní osoby i po odstranění zjevných osobních údajů. Zejména se může jednat o kombinace jako město, věk a pohlaví nebo podobné.
3. Agregace. Data, která není možné nebo vhodné zveřejnit dle předchozího bodu, lze zveřejnit v agregované podobě. Tedy v podobě statistik. V případě zveřejnění statistik je žádoucí použít co nejjemnější možné členění. Tedy například počet přijatých podání uvádět raději po měsících namísto jen po letech.

Více informací naleznete na stránce o ochraně osobních údajů a GDPR ve vztahu k otevřeným datům.

Právní aspekty

Legislativní rámec otevřených dat v České republice tvoří jejich úprava obsažená v Zákoně č. 106/1999 Sb., o svobodném přístupu k informacím a v Nařízení vlády č. 425/2016 Sb., o seznamu informací zveřejňovaných jako otevřená data, které stanovuje vybraným orgánům veřejné správy povinnost zveřejňovat data z konkrétních jimi spravovaných informačních systémů ve formě otevřených dat.

Více informací o strategických dokumentech, akčních plánech a souvisejících předpisech ČR i EU naleznete na stránce věnované legislativnímu prostředí otevřených dat.

From:

<https://archi.gov.cz/> - Architektura eGovernmentu ČR

Permanent link:

https://archi.gov.cz/nap:verejny_datovy_fond?rev=1557922745

Last update: **2019/08/12 11:59**

