

MINISTERSTVO VNITRA  
ČESKÉ REPUBLIKY

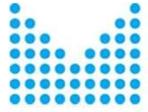
# Otevřená data nejen českou optikou

konference Otevřená data ve veřejné správě 2020

23.10.2020

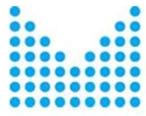
**Michal Kubáň**

odbor Hlavního architekta eGovernmentu MV ČR



# Struktura

- Legislativa
- Národní katalog otevřených dat
- Standardizace dat
- Školení a workshopy k otevřeným datům
- Články a propagace otevřených dat



# Legislativa

- Proces transpozice Směrnice EU o otevřených datech a opakovaném použití informací veřejného sektoru (2019/1024) do následujících předpisů:
  - zákona č. **106/1999 Sb.**, o svobodném přístupu k informacím
  - zákona č. **123/1998 Sb.**, o právu na informace o životním prostředí
  - zákon č. **130/2002 Sb.**, o podpoře výzkumu, experimentálního vývoje a inovací z veřejných prostředků a o změně některých souvisejících zákonů



# Datové soubory s vysokou hodnotou (HVDs)

Výběr z 6 kategorií:

Geoprostorové  
údaje

Pozorování  
Země a životní  
prostředí

Meteorologie

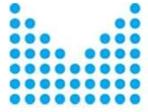
Statistika

Společnost a  
vlastnictví  
společnosti

Mobilita

Harmonogram:

- Studie hodnocení dopadu (RIA)  
**12/2019 – 08/2020**
- Veřejné konzultace MV  
**01/2020**
- Veřejné konzultace a slyšení  
**05/2020 – 06/2020**
- První draft návrhu seznamu HVDs  
**12/2020**
- Schválení HVDs  
**01/2021 – 03/2021**



# Datové soubory s vysokou hodnotou (HVDs)

## Pohled ČR:

- Požadavky na technickou specifikaci
- Potenciál HVDs nebude využit, dokud každá datová sada nebude obsahovat:
  - technickou dokumentaci
  - seznam atributů datové kvality

**OTEVŘENÁ DATA**

Články Novinky Národní katalog otevřených dat ▾ Další informace ▾

## Důležitost datové kvality datových sad s vysokou hodnotou

13. 2. 2020, Michal Kubáň

Publikace takzvaných datových sad s vysokou hodnotou, což je termín, který zavádí evropská směrnice o otevřených datech (2019/1024), nesplní očekávání, pokud nebude společně s jejich seznamem řádně řešena i jejich datová kvalita.

Je nezbytně nutné, aby pro každou datovou sadu na seznamu existovala technická specifikace a sada vyžadovaných atributů datové kvality. Musí být vynucováno dodržování technické specifikace a specifikovaných atributů datové kvality poskytovatelem dané datové sady v jednotlivých členských státech, jinak datové sady nebودou interoperabilní a jejich hodnota zůstane nevyužitelná. V tomto dokumentu identifikujeme relevantní atributy datové kvality a navrhujeme, jak by mohly být definovány a vyhodnocovány. Zádáme, aby byla každá budoucí definice datové sady s vysokou hodnotou doprovázena svou technickou specifikací, sadou vyžadovaných atributů datové kvality a mechanismem jejich sledování a vynucování jejich dodržování.

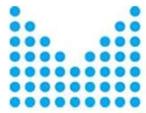
Požadavky v tomto dokumentu budou relevantní v okamžiku, kdy budou datové sady s vysokou hodnotou identifikovány. Pro zajištění přeshraniční interoperability poskytovaných dat tvořících jednu datovou sadu s vysokou hodnotou od jejich poskytovatele v jednotlivých členských státech je naprostě nedostačující datovou sadu pouze identifikovat. Je třeba zajistit, že po každou takovou datovou sadu vznikne řádná technická specifikace a jsou definovány požadované atributy datové kvality tak, aby členské státy mohly začít publikovat příslušná data interoperabilním způsobem, tak, aby data specifikaci a atributy dodržovaly. Komentáře a návrhy jsou vítány. V případě dotazů nás neváhejte [kontaktovat](#).

### Datové sady s vysokou hodnotou

Směrnice Evropského parlamentu a Rady (EU) 2019/1024 ze dne 20. června 2019 o otevřených datech a opakováném použití informací veřejného sektoru zavádí tzv. datové sady s vysokou hodnotou. Jedná se o datové sady, jejichž opakování použití různými zpracovateli dat bude mít zajímavé přínosy pro evropskou ekonomiku a společnost. V současné době je připravován jejich seznam. Po jeho schválení budou muset být datové sady uvedené v seznamu publikovány ve všech členských státech EU jako otevřená data, tj. ve strojově čitelném a otevřeném formátu, bez omezujících podmínek užití, volně na internetu a bezplatně.

Koncept datových sad s vysokou hodnotou se týká vyjmenovaných oblastí, kterými v současnosti jsou:

- Geoprostorové údaje
- Pozorování Země a životní prostředí
- Meteorologie
- Statistika
- Společnosti a vlastnictví společnosti
- Mobilita



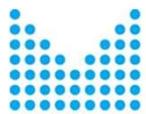
# Národní katalog otevřených dat (NKOD)

- Malý nárůst počtu poskytovatelů:

36 → 41  
2019 2020

- Problémy s kvalitou publikovaných dat:

- **22 %** poskytovatelů neuvádí podmínky užití
- **51 %** poskytovatelů nevyplňuje povinné atributy
- **41 %** poskytovatelů má některé datové sady nedostupné



# Kvalita otevřených dat

## Otevřené formální normy (OFN)

- technická doporučení
- zlepšují interoperabilitu dat
- jednoduší využití datové sady nezávisle na poskytovateli

The screenshot shows the 'Otevřená DATA' website interface. The top navigation bar includes a search bar and a login button. Below the header, a breadcrumb trail shows the current location: Umístění / Otevřené formální normy (OFN) / Historie / Otevřené formální normy (OFN). A sidebar on the left lists various links related to open data, with 'Otevřené formální normy' highlighted in a dark box. The main content area features a title 'Otevřené formální normy (OFN)' and a descriptive text about the purpose of OFN. A yellow callout box contains a lightbulb icon and text encouraging users to contribute to the GitHub repository. At the bottom, there's a section titled 'Význam a způsob použití otevřených formálních norem' with a list of 4 points explaining the use of OFN.

OTEVŘENÁ DATA

Hledat

Přihlásit se

Umístění / Otevřené formální normy (OFN) / Historie / Otevřené formální normy (OFN)

Co jsou otevřená data?

- Národní katalog otevřených dat (NKOD)
- Postup otevírání dat
- Vytvoření publikačního plánu
- Publikace datové sady v otevřené podobě
- Zprovoznění lokálního katalogu otevřených dat a jeho registrace v NKOD
- Špatná praxe v otevřených datech

Otevřené formální normy

Vzdělávání

Legislativa

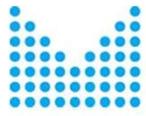
Kontakt

Otevřené formální normy (OFN)

Na otevřených formálních normách pracujeme na GitHubu. Navrhujte nové normy a přispivejte do existujících!

Otevřená formální norma (OFN) specifikuje, jak publikovat data o konkrétních věcech, např. sportovištích, akcích apod. Smysl OFN a zamýšlený způsob použití je následující:

1. Pokud publikujete data, která významově odpovídají tomu, co modelujeme, použijte tuto specifikaci, ať je to všude stejně a dá se s daty jednotně pracovat, nezávisle na zdroji.
2. Všechny položky jsou nepovinné, tj. pokud některé nemáte, a objektivně nejsou potřeba, nic se neděje, prostě je v datech vyměňáte.
3. V OFN pro konkrétní datové sady ale máme „minimalistický“ příklad. Tím říkáme, že pokud nepublikujete/nemáte ani to co je něm, tak ta data zřejmě nebude dát smysl a tudíž je nikdo nebude schopen použít, což je proti smyslu otevřených dat.
4. Pokud máte položky navíc, tak je můžete libovolně přidat. V tom případě bychom byli rádi, když nám dáte vědět, abychom mohli zvážit, zda to není něco, u čeho má smysl uvažovat o zanesení do OFN pro osatní poskytovatele a zpracovatele těchto dat.



# Školení a workshopy k otevřeným datům

- Skupinové školení, workshopy a konzultace poskytujeme i online, například:

Úvodní školení

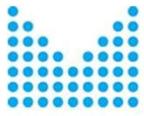
Příprava dat k publikaci

Katalogy a katalogizace

Konceptuální modelování a OFN

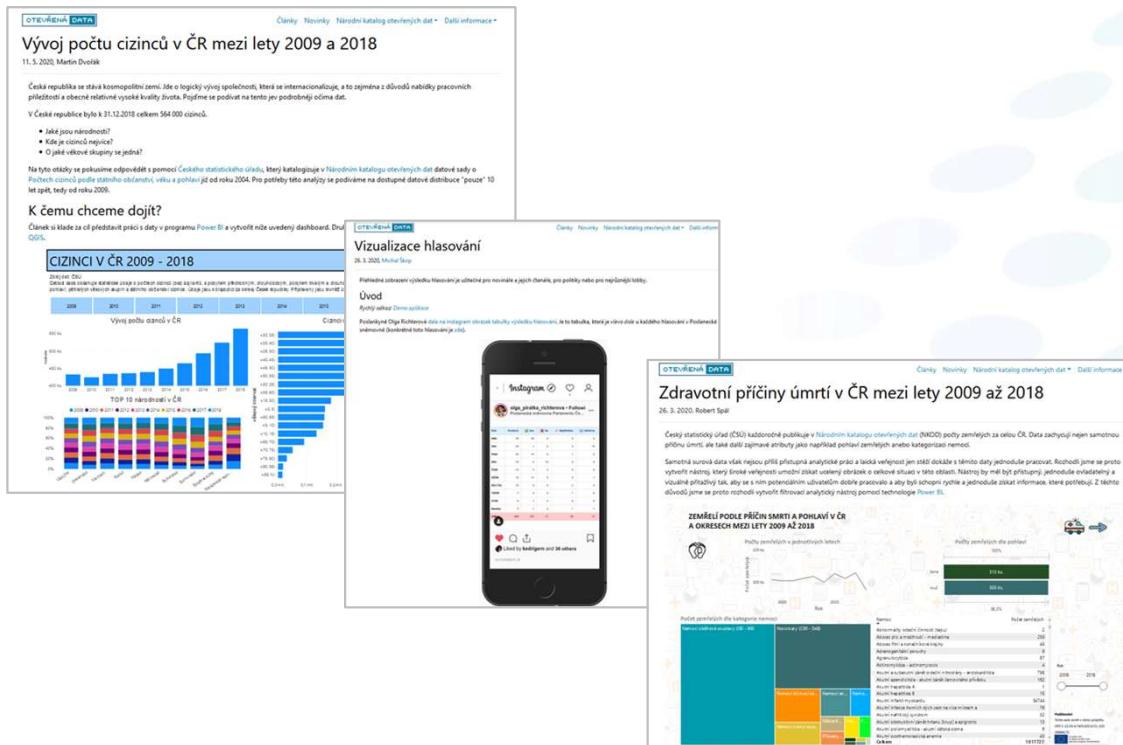
- Připravujeme eLearning pro **individuální vzdělávání**

- Příprava publikačního plánu (konec roku 2020)



# Články na portálu data.gov.cz

Cílem nové sekce je vzdělávat a informovat!



**Otevřená data**

## Články

Otevřená data a otevřené formální normy  
2. 10. 2020

Pro mnoho veřejných institucí se poskytování otevřených dat stalo běžním způsobem sdílení informací s veřejností. S rostoucím počtem poskytovatelů otevřených dat a s různorodostí dat, která jsou takto poskytována, se ukazuje, že nutným předpokladem použitelnosti dat je sjednocování jejich podoby. Na to pamatuje i legislativa, která v zákoně 106/1999 Sb. o svobodném přístupu k informacím zavádí pojem **otevřené formální normy (OFN)** jako písemně vydanou specifikaci požadavků na zajištění schopnosti různých programových vybavení vzájemně si poskytovat služby a efektivně spolupracovat.

V minulém dílu seriuje jsme se seznámili s pojmem **znalostní graf** (angl. *Knowledge Graph*). V tomto dílu si ukážeme, jakým způsobem lze znalostní grafy strojově reprezentovat v podobě vhodné pro jejich sdílení na webu jako otevřená data.

[Více >](#)

Série Znalostní grafy: Dil 2: Datový model RDF  
10. 8. 2020

**RDF model**

V květnu 2012 začal Google používat pojmenování **Knowledge Graph** (český **znalostní graf**) pro označení svého znalostního báze, kterou používá pro doplňování výsledků vyhledávání. Od té doby se pojem rozšířil do další oblasti a zášhlí i do světa otevřených dat. V tomto článku si znalostní grafy představíme a podíváme se, jak lze s nimi pracovat a jaké mohou najít uplatnění a dokonce jež již nachází v světě otevřených dat.

[Více >](#)

**Starší články**

Vývoj počtu cizinců v ČR mezi lety 2009 a 2018  
11. 5. 2020

Česká republika se stává kosmopolitní zemí. Jde o logický vývoj společnosti, která se internacionálizuje, a to zejména z důvodu nabídky pracovních příležitostí a obecně relativně vysoké kvality života. Pojďme se podívat na tento jev podrobněji očima dat.

[Více >](#)

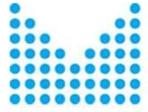
Kartogram ČR (choropleth, choropletová mapa)  
23. 4. 2020

Hackathony 2019  
9. 4. 2020

Hackathony se staly v posledních letech stále oblíbenějším nástrojem, jak přijít s rychlými návrhy nových procesů, aplikací a řešení, které mohou rozvinout prakticky jakoukoliv oblast veřejného nebo soukromého života. Jedná se prakticky o soutěž s cílem nalézt nejlepší řešení předem dané výzvy.

I když "hackovat" je možné prakticky cokoliv.





MINISTERSTVO VNITRA  
ČESKÉ REPUBLIKY

# Děkuji za pozornost

Michal Kubáň  
[otevrenadata@mvcr.cz](mailto:otevrenadata@mvcr.cz)