



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost



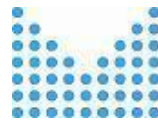
MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY

C1V6 - Analýza stavu datových sad s vysokou hodnotou v Česku

Vytvořeno v rámci projektu

**Rozvoj datových politik v oblasti zlepšování kvality
a interoperability dat veřejné správy**
CZ.03.4.74/0.0/0.0/15_025/0013983

Klíčová aktivita: 02 Analýza a aktualizace legislativy
Verze výstupu: 01



Relevantní právní předpisy:

- Směrnice Evropského parlamentu a Rady (EU) 2019/1024 ze dne 20. června 2019 o otevřených datech a opakovaném použití informací veřejného sektoru¹
- Prováděcí nařízení Komise (EU) 2023/138 ze dne 21. prosince 2022, kterým se stanoví seznam konkrétních datových souborů s vysokou hodnotou a opatření pro jejich zveřejnění a opakované použití² (dále jako „prováděcí nařízení“)
- Zákon č. 106/1999 Sb., o svobodném přístupu k informacím
- Národní právní předpisy specifikované v analýze

Popis výstupu C1V6:

Výstupem projektu je dokument shrnující dílčí analýzy, které budou vznikat v souvislosti se zaváděním seznamu datových sad s vysokou socio-ekonomickou hodnotou. Cílem výstupu je identifikace povinných subjektů odpovídajících za jednotlivé datové sady s vysokou socio-ekonomickou hodnotou podle toho, jak tyto datové sady definuje příloha prováděcího aktu ke směrnici EU č. 2019/1024. V návaznosti na to bude vytvořen přehled těchto datových sad, identifikace jejich dosavadního právního zakotvení na úrovni národního práva, identifikace příslušných subjektů, které mají tyto datové sady poskytovat a zhodnocení míry naplnění dílčích požadavků na tyto datové sady a jejich uveřejnění v ČR. Výstup bude připraven tak, aby byl prakticky použitelný jako podklad pro zajištění souladu s novou právní úpravou.

Forma výstupu:

Stěžejním dokumentem shrnujícím výsledky analýzy je tabulkový dokument, který obsahuje přehled subjektů, které zpracovávají data s vysokou socio-ekonomickou hodnotou, a který shrnuje současný stav publikování těchto dat v relevantních katalozích, především v Národním katalogu otevřených dat. Doprovodnou úlohu má tento textový dokument, jehož cílem je popsat proces vzniku výsledné analýzy a vysvětlit, jakým způsobem byla analýza koncipována.

¹ <https://eur-lex.europa.eu/legal-content/CS/TXT/?uri=CELEX:32019L1024>

² <https://eur-lex.europa.eu/legal-content/CS/TXT/?uri=CELEX:32023R0138>



Metodologie a popis výstupu:

Východiskem pro analýzu byla příloha prováděcího nařízení Komise (EU) 2023/138 ze dne 21. prosince 2022, kterým se stanoví seznam konkrétních datových souborů s vysokou hodnotou a opatření pro jejich zveřejnění a opakované použití. Příloha stanovuje šest oblastí – data sbíraná v rámci těchto oblastí jsou Evropskou unií považována za data s vysokou socio-ekonomickou hodnotou. Jedná se o následující témata:

1. Geoprostorové údaje
2. Pozorování Země a životní prostředí
3. Meteorologie
4. Statistika
5. Společnosti a vlastnictví
6. Mobilita

Každé z těchto oblastí odpovídá jeden list v tabulkovém dokumentu.

Data týkající se moří a oceánů se České republiky netýkají, proto jsou tyto datové sady označeny jako “(nesbíráme)” a dále se jimi analýza nezabývá.

V rámci uvedených šesti oblastí byly nejprve identifikovány subjekty, které v rámci své činnosti sbírají a zpracovávají odpovídající data.

Geoprostorové údaje	Pozorování Země a životní prostředí	Meteorologie
Český úřad katastrální a zeměměřičský	Český úřad katastrální a zeměměřičský	Ministerstvo životního prostředí
Ministerstvo zemědělství	Ministerstvo životního prostředí	
	Ministerstvo zemědělství	
	Ministerstvo pro místní rozvoj	

Statistika	Společnosti a vlastnictví	Mobilita
Český statistický úřad	Ministerstvo spravedlnosti	Ministerstvo dopravy
		Ministerstvo zemědělství
		Ministerstvo životního prostředí

V některých případech konkrétní data sbírají subjekty, které jsou podřízené hlavnímu subjektu odpovědnému za činnost v rámci oblasti definované Evropskou komisí. Jedná se o subjekty v oblastech Pozorování země a životního prostředí a Mobilita. Dalším údajem, který byl v rámci analýzy identifikován, tedy byly podřízené subjekty, nicméně jen v případě, pokud se sběru relevantních dat účastní. Tabulkový dokument tedy neslouží k vytvoření přehledu hierarchie jednotlivých subjektů v rámci dané instituce. Pokud je pole podřízeného subjektu u konkrétní datové sady prázdné, znamená to, že sběr dat



provádí přímo hlavní subjekt uvedený ve sloupci Subjekt. To stejné platí, pokud v dané kartě odpovídající jedné z šesti oblastí sloupec Podřízený subjekt zcela chybí.

Podle prováděcího nařízení musí subjekty publikovat příslušné datové sady formou otevřených dat v případě, že relevantní data sbírají. To znamená, že prováděcí nařízení nestanovuje novou povinnost začít sbírat všechna data, která jsou definovaná v příloze prováděcího nařízení. Jestliže subjekty příslušné datové sady nemají k dispozici, pak je nemusí začít vytvářet, a nevztahuje se na ně tedy povinnost publikovat je formou otevřených dat. Z tohoto důvodu je v dalším sloupci identifikován právní předpis, který subjektům stanovuje povinnost příslušná data sbírat, aby bylo zřejmé, zda mají subjekty povinnost datové sady publikovat formou otevřených dat.

V následujícím sloupci se nachází informace o publikaci datové sady v Národním katalogu otevřených dat. Informace v tomto sloupci značí nejen to, že daný subjekt sbírá příslušná data, ale zejména to, zda jsou již přístupná ve formátu otevřených dat ve smyslu § 3a odst. 5 zákona č. 106/1999 Sb., o svobodném přístupu k informacím.

Informace, zda jsou datové sady publikovány ve formátu otevřených dat v Národním katalogu otevřených dat, je zásadní. Následující analyzované údaje slouží k vytvoření přehledu o tom, zda jsou otevřená data v souladu s požadavky uvedenými v příloze prováděcího nařízení Komise, tedy zejména s požadavky na granularitu, frekvenci aktualizace a licenci.

V případě, že karta neobsahuje zvláštní sloupec s požadavky na granularitu, a je uvedeno, že jsou data publikována formou otevřených dat, pak to značí, že jsou zároveň splněny požadavky na granularitu. Jde například o oblast Statistiky, kde je popis požadované granularity natolik komplexní a provázaný s jinými kritérii, že nebylo účelné převádět ji nově do přehledové tabulky, která je výstupem této analýzy.

V případě, že je ve sloupci „licence v pořádku“ uvedeno „ano“, pak to znamená, že jsou otevřená data publikována na základě licence Creative Commons BY 4.0 nebo jakékoli rovnocenné či méně omezující otevřené licence tak, jak to vyžaduje prováděcí nařízení Komise. Obdobně je ve sloupci „frekvence aktualizace v pořádku“ uvedeno, zda jsou data aktualizovaná v periodicitě požadované prováděcím nařízením.

Pro přehlednost a snazší kontrolu je v tabulce uveden sloupec „Poznámka“, a dále také sloupec „Odkaz na NKOD“. V tomto posledním sloupci by měl být uveden URL odkaz na konkrétní datovou sadu publikovanou v Národním katalogu otevřených dat, která odpovídá popisu dle přílohy prováděcího nařízení Komise. Pro případ, že je předepsaná datová sada rozdělena na národní úrovni v katalogu do více datových sad (což neodporuje příloze prováděcího nařízení), jsou tyto rozděleny na jednotlivé řádky, a navíc popsány ve sloupci „název datové sady“ pro lepší přehlednost. Jedná se zejména o případy v oblasti mobility.



Shrnutí zjištění:

1. Geoprostorové údaje

Subjekty, které zpracovávají údaje patřící do této skupiny, jsou Český úřad zeměměřický a katastrální a Ministerstvo zemědělství. Z analýzy vyplývá, že požadované datové sady jsou k dispozici a jsou již publikovány v Národním katalogu otevřených dat³, případně v Geoportálu.⁴ U datových sad jsou tedy uvedeny odkazy na příslušnou datovou sadu, která splňuje náležitosti pro otevřená data. Výjimkou je datová sada „Zeměpisné názvy“, tato data však máme k dispozici. Jejich publikování formou otevřených dat je naplánováno na druhou polovinu roku 2023 v závislosti na novele zákona č. 200/1994 Sb., o zeměměřičství.

2. Pozorování Země a životní prostředí

Tato oblast je specifická tím, že v příloze k nařízení nebyly definovány jednotlivé datové sady tak, jak tomu bylo v případě jiných oblastí. Z předpisů, na které příloha odkazuje, tedy byla dovozena jednotlivá témata, v jejichž rámci je možné datové sady sbírat. Na základě toho pak jednotlivé subjekty samy definovaly příslušné datové sady, které tématům mohou odpovídat a něž by se mohlo jednat.

První část témat definovaných pro oblast Pozorování Země vychází ze směrnice INSPIRE, což znamená, že příslušné subjekty mají tyto datové sady k dispozici. Na základě analýzy bylo zjištěno, že převážná většina těchto datových sad však není publikována formou otevřených dat.

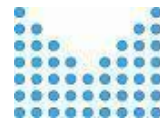
Frekvence aktualizace není přílohou stanovena. Je pouze uvedeno, že data mají být poskytována v „nejaktuálnější“ možné verzi. Sloupec „Frekvence aktualizace“ tedy v této kartě nevypovídá o předepsané povinnosti, ale pro lepší přehled je sloupec vyplněn údaji, které vypovídají o skutečnosti, jak často jsou data reálně aktualizována.

3. Meteorologie

V oblasti Meteorologie byl identifikován pouze jeden odpovědný subjekt, kterým je Ministerstvo životního prostředí, respektive Český hydrometeorologický ústav. Požadovaná data jsou sbírána, nicméně aktuálně nejsou publikována formou otevřených dat.

³ <https://data.gov.cz/datov%C3%A9-sady>

⁴ <https://geoportal.gov.cz/web/guest/map>



4. Statistika

Statistika je specifickou oblastí, neboť požadavky na granularitu dat jsou, vzhledem k jejich povaze, respektive vzhledem k možné rozmanitosti informací, velmi detailní. Jediným odpovědným subjektem v této sféře je Český statistický úřad. Dle analýzy bylo zjištěno, že Český statistický úřad publikuje v Národním katalogu otevřených dat velké množství datových sad. Nejedná se nicméně o všechny datové sady, které požaduje Komise v prováděcím nařízení. V rámci této oblasti také dochází k tomu, že datové sady publikované ve formátu otevřených dat jsou publikovány v jiném rozsahu, než požaduje nařízení – například je jedno téma definované přílohou nařízení rozčleněné do vícero datových sad, nebo je publikována pouze část požadovaných dat. O těchto rozdílech vede Český statistický úřad vlastní, detailní přehledovou tabulku, v jejímž rámci eviduje (ne)naplnění požadavků definovaných přílohou prováděcího nařízení.

5. Společnosti a vlastnictví

Datové sady spadající pod oblast Společnosti a vlastnictví spravuje Ministerstvo spravedlnosti a jsou v Česku k dispozici. Z analýzy je však zřejmé, že žádná datová sada v současné době není publikována formou otevřených dat.

6. Mobilita

Datové sady z oblasti Mobility zaštiťují Ministerstvo dopravy, Ministerstvo zemědělství a Ministerstvo životního prostředí, zejména pak subjekty jim podřízené. V rámci analýzy bylo zjištěno, že pouze čtyři z potřebných datových sad jsou publikovány formou otevřených dat v Národním katalogu otevřených dat.



Závěr:

Na základě provedené analýzy vznikl tabulkový soubor, jenž v jednotlivých kartách shrnuje zjištění o aktuální podobě publikování otevřených dat k příslušným oblastem, které byly specifikovány Evropskou komisí jako oblasti, do nichž spadají datové sady s vysokou socioekonomickou hodnotou.

Jednotlivé oblasti vykazují různou míru naplnění požadavků na publikování ve formátu otevřených dat. Podrobný přehled byl zpracován do tabulky a závěry plynoucí pro jednotlivé oblasti jsou uvedeny výše v tomto dokumentu.

Údaje, které nebyly zjištěny v rámci této analýzy, budou předmětem zkoumání v implementačním období předpisu. Zpracovaná tabulka poslouží jednotlivým odpovědným subjektům jako vodítko při naplňování požadavků Evropské komise a jako checklist při publikování otevřených dat.