

Požadavky na zpřístupnění dat o nákaze Covid-19

Vzhledem k tomu, že v průběhu epidemie Covid-19 vláda často velmi razantně zasahuje do základních práv a svobod, je nezbytné, aby občané chápali, z jakého důvodu je přistupováno k daným opatřením. Občanům by mělo být z odůvodnění také zřejmé, na základě jakých epidemiologických dat se vláda ke svým krokům uchyluje. Jen vládní opatření s dostatečným odůvodněním a odkazem na jasná data může vzbudit důvěru veřejnosti, že vláda šetří podstaty jejich práv a přemýšlí o přiměřenosti svých kroků. Vláda však anonymizovaná data v dostatečném rozsahu neposkytuje pro širokou veřejnost, ale ani pro vědecké a analytické účely. Data lze přitom anonymizovat - chránit osobní identitu, ale zároveň zachovat dostatečnou úroveň rozlišení, která by pomohla vědcům dospět k hlubšímu poznání nákazy Covid-19 a k souvisejícím analýzám.

Agregovaná data na úrovni okresů už nově jsou. Dává však smysl poskytování epidemiologických dat rozdělit na **dvě kvalitativní úrovně**, jak to běžně dělá ČSÚ:

a) Pro širokou veřejnost

- granularita na úrovni okresů postačuje
- doplnit o rozpad i podle profese, věku a zejména důvodu testu (podezření z nákazy, opakovaný test u nemocného apod.)
- jasně zdokumentovat, co jaký uveřejněný údaj znamená (máme-li tam např. dnes okres, je to bydliště testovaného, sídlo ústavu, hygienické stanice..?)
- neuveřejňovat týden zpětně jako dosud, ale denně, jak data přicházejí

Jedna informace tedy bude vypadat např. takto:

Den: 18.5., okres: Hodonín, důvod testu: preventivní, věk: 20 - 30, profese: zdravotní pracovník, počet testů: 12, pozitivní výsledek: 0 %.

Klíčem k dostatečné anonymizaci je zejména velikost jednotky (okres) a široké vymezení profese (zdravotník, nikoliv lékař). Dále postačí data uvádět v širokých kategoriích (např. věk 0-6,7-12,13-18,19-30, 30-40, 40-50, atd.). **Hlavním smyslem je podchytit zásah různých skupin - děti v MŠ, ZŠ, SŠ, důchodců atp. a tedy sledovat dopady různých opatření.** Pokud by tyto demografické skupiny šlo podchytit např. v popisu profese ("důchodce", "student ZŠ" apod.), mohou být věkové skupiny ještě širší.

b) **Pro zájemce z odborné veřejnosti, kteří podepíší nějaké NDA (zavážou se data dále nešířit, využívat pouze k výzkumným účelům)**, lze údaj o lokalitě dále zpřesnit např. na úroveň ORP - při zachování dostatečné anonymizace. Jedna informace tedy bude vypadat např. takto:

Den: 18.5., město: Hodonín, důvod testu: preventivní, věk: 20 - 30, profese: zdravotní pracovník, počet testů: 12, pozitivní výsledek: 0 %

Zde bude možné dovést konkrétní osobu jen v naprosto výjimečných případech a s vysokou znalostí místních reálií (např. jediný zdravotník v daném věkovém rozmezí a lokalitě). Pravděpodobnost takové situace je však mizivá, neboť nejmenší ORP v ČR mají cca 10 000

obyvatel. Aby se riziko dále omezilo, poskytují se údaje jen omezené skupině za striktních podmínek. Tak je to běžné třeba u práce výzkumníků s citlivými daty ČSÚ, VZP apod.

Dále lze zvážit poskytování určitých “vedlejších” dat, která nejsou přímo o nákaze Covid-19, ale která doplňují představu o epidemii - data, která umožní křížovou validaci dat, dopady Covid-19 do jiných oblastí, a která musí být bezpochyby součástí analýz pro rozhodování vlády, např. v ekonomických otázkách:

1) Denní sumy úmrtí evidovaných v Národním registru hrazených zdravotních služeb a Informačním systému infekční nemoci od začátku evidence, a to alespoň na úrovni 1 okresu nebo menší územní jednotky.

2) Denní souhrnné statistiky údajů evidovaných v Národním registru hrazených zdravotních služeb, podle kategorií hrazených zdravotních služeb, na úrovni 1 okresu či konkrétní nemocnice. Jinak řečeno, statistiky zdravotnických výkonů jednotlivých nemocnic. Je z toho poznat, jak se snížilo léčení jiných než Covid-19 nemocí v době epidemie v různých nemocnicích a oblastech. A do budoucna, zda to bude mít vliv na další nemocnost a úmrtí.

3) Zdrojová data použitá pro výpočet reprodukčního čísla R.

4) Veškeré analýzy (a statistiky), na základě kterých vláda rozhoduje, včetně zdrojových dat pro tyto analýzy.

Za Rekonstrukci státu připravili: Jiří Skuhrovec (EconLab), Michal Bláha (Hlídač státu), Lukáš Kraus (Frank Bold). S významným přispěním doc. Ing. Daniela Münicha, Ph.D. (výkonný ředitel think-tanku IDEA při CERGE-EI).