

Nepatřičné položky v číselnících - ostatní, neuveřejněno, apod.

Číselníky (viz také [OFN Číselníky](#)) jsou jedny z nejdůležitějších druhů datových sad vůbec. Obsahují výčet hodnot a jejich kódů, které spolu s jinými číselníky dávají smysl ostatním datovým sadám, a umožňují různé datové sady propojit právě na základě hodnot ze společných číselníků. Příkladem může být číselník obcí a číselník pohlaví, které dávají význam datové sadě Průměrný plat podle obcí a pohlaví, která se na položky těchto číselníků odkazuje.

Problém

Problém nastává v okamžiku, kdy se do číselníků dostanou nesmyslné položky, které mají zástupný charakter. Nejčastěji se jedná o položky typu *Nezjištěno*, *Ostatní*, *Není apod.*, které jsou výsledkem špatného návrhu systému, který s tímto číselníkem pracuje, nebo ze kterého číselník pochází. Taková položka se pak v datech použije v následujících případech:

- Reprezentace faktu, že pro nějaký údaj není hodnota z tohoto číselníku známa (*Není*)
- Reprezentace faktu, že pro nějaká data číselník neobsahoval vhodnou položku (*Ostatní*)
- Reprezentace faktu, že pro nějaký údaj nebyla hodnota z tohoto číselníku zjištěna (*Nezjištěno*).

Nastává pak situace, že někdo chce číselník použít ve své aplikaci, a pak se například mezi obcemi objeví obec *Není* či obec *Ostatní*, což je pochopitelně nesmysl.

Řešení

Řešením této situace je především oprava návrhu systémů, které na takové položky spoléhají. Pokud se může stát, že pro nějaký údaj hodnotu z číselníku nezjišťuji, pak toto zaznamenám jako další vlastnost mých dat (*Zjištěno/Nezjištěno*), nikoliv přidáním hodnoty *Nezjištěno* do číselníku obcí, a podobně v ostatních situacích.

From:

<https://opendata.gov.cz/> - **Otevřená data**

Permanent link:

<https://opendata.gov.cz/%C5%A1patn%C3%A1-praxe:ostatn%C3%AD-polo%C5%BEky> 

Last update: **2019/10/04 12:54**