Shrnutí a doporučení

Pro potřeby dílčích i komplexních analýz je možné vytvořit jednotnou databázi všech veřejně dostupných datových souborů, které tato analýza identifikovala. Ty budou roztříděny dle zdroje dat, tematického zaměření, úrovně měření (žáci, školy, kraje, stát). Jednalo by se o datový archiv, který by byl veřejně dostupný (např. na stránkách ČŠI) v českém i anglickém jazyce. Vytvoření online datového archivu si vyžádá pouze činnost stávajících zaměstnanců.

V dnešní éře pokročilých analýz velkého objemu dat (*big data*) by komplexnost, porovnatelnost a systematičnost zvýšilo propojování všech dat do jednoho velkého datového souboru. Základním identifikátorem tohoto souboru by byl kód školy (např. REDIZO). Kód školy umožňuje snadnou párovatelnost, resp. propojitelnost různých datových sad. Druhým identifikátorem by byl kód obce, pro členěná statutární města a hlavní město Prahu pak kód místní části. Krokem k systematizaci by bylo pravidelné udržování této databáze, což by umožnilo ČŠI provádět sofistikované analýzy na velkém množství dat (například v programu R). Pro kvantitativní analýzu dat z této databáze pak nemusí být splněny zcela přísné akademické standardy, jako je reprezentativnost vzorku. Způsobem vytěžení všech dat by byl tzv. data mining. Jedná se o analytický způsob získávání netriviálních skrytých a potencionálně užitečných informací z velkého množství dat.