

VYUŽITÍ ADMINISTRATIVNÍCH A GEOGRAFICKÝCH DAT PRO MONITOROVÁNÍ NEROVNOSTÍ NA PŘÍKLADU MOBILITY ŽÁKŮ ZÁKLADNÍCH ŠKOL

ADMINISTRATIVE AND GEOGRAPHIC DATA USE FOR MONITORING INEQUALITY IN EDUCATION: THE CASE OF STUDENT MOBILITY

Dominik Dvořák¹, Petr Meyer², Silvie R. Kučerová², Petr J. Gal¹

¹ Ústav výzkumu a rozvoje vzdělávání, Pedagogická fakulta, Univerzita Karlova, Praha

² Katedra geografie, Přírodovědecká fakulta, Univerzita Jana Evangelisty Purkyně
v Ústí nad Labem, Ústí nad Labem

Abstrakt

Administrativní data představují velkou výzvu pro pedagogický výzkum a jejich využití může pomoci zvýšit spravedlivost v přístupu ke kvalitnímu vzdělávání. Cílem tohoto příspěvku je představit metodologické postupy a problémy spojené s využitím dat českých centrálních orgánů na příkladu studia meziškolní mobility žáků, která je považována za významný indikátor pro školy se znevýhodněnou populací. Výsledky ukazují mj. nerovnost v přístupu k volbě školy mezi městskými a venkovskými obcemi.

Klíčová slova: *administrativní data, geografie vzdělávání, mobilita žáků, spravedlivost, Česko*

Abstract

Administrative data on schooling have enormous potential for educational research and for improved and more equitable access to quality teaching and learning. The purpose of this methodological paper is to examine methodological issues and risks related to use and merging of Czech government data for the study of inter-school mobility as an important indicator drawing attention to risk pupils and schools. The results feature among others the differences in school choice access and use in urban, suburban, and rural municipalities.

Keywords: *administrative data, geography of education, students' mobility, equity, Czechia*

1 Úvod

Množství dat, která jsou v současné době dostupná výzkumníkům, dramaticky roste. Rozvoj využívání velkých dat (big data) podle některých vědců přinese revoluci v sociálních vědách, jež zahrne nejen nové metodologie výzkumu, ale nové typy výzkumných otázek i aplikací v oblasti praxe. V tomto příspěvku představujeme zkušenosti s využíváním (mikro)dat ministerstva školství a otevřených dat České školní inspekce. Na základě našich předchozích zkušeností s využíváním dat o přestupech žáků středních škol (Dvořák et al., 2020) se nově zaměřujeme na mobilitu žáků základních škol ve venkovských a městských obcích.

2 Teorie

Administrativní data jsou typem „velkých dat“ – jde o data vytvářená státní správou k řízení a evidenci v rezortu. Nejsou tedy primárně určena pro výzkum, nicméně mají velký potenciál i pro monitorování nerovností ve vzdělávání (Connelly et al., 2016; Pattaro et al., 2020; Siddiqui, 2019). Panuje naděje, že dostupnost těchto dat posílí porozumění chování aktérů, efektivitu vzdělávací politiky,lepší nabídku vzdělávání i spokojenost aktérů (Kubáň, 2016). Shromažďování a výzkumné využívání velkého množství dat o jednotlivcích ovšem není bez filozofických, metodologických a etických problémů.

Redukce jedinečných životů lidí na několik čísel v evidenci a dostupnost velkých dat o chování jednotlivců (i pro soukromé firmy) znepokojovala intelektuály od G. Orwella přes M. Foucaulta po N. J. Harariho a další současníky. Ambivalenci role moderního státu s jeho „přebujenou byrokracií“, která zároveň v procesu modernizace osvobozuje, analyzoval u nás nedávno např. historik P. Himl (2019).

3 Metodologie výzkumu

Základním postupem při využívání administrativních dat je propojování (merging, linking) více databází. V našem příspěvku ukážeme postupy a přínosy

propojování rejstříkových, výkazových a matrikových data MŠMT ČR. Využíváme např. agregované informace i mikrodata o zápisech ve všech základních školách v České republice, rejstříková data o kapacitách a zřizovateli škol, a data z matrik o všech přestupech žáků mezi školami za tři školní roky. Dále demonstrujeme jejich propojování s databázemi sociogeografických dat o obcích a dalšími daty nesoucími prostorovou informaci (prostorová data).

Rámec našeho postupu tvoří model tzv. dvojitého diamantu, který představuje celkový výzkumný plán analýzy dat.

4 Výsledky

V metodologickém příspěvku slouží empirické výsledky pouze pro ilustraci práce s daty. Potenciál, postupy a problémy využití propojených administrativních a geografických dat budou ilustrovány na příkladu přestupů žáků základních škol. Data ukazují na dva základní jevy: prostorovou variabilitu mobility rodin žáků (stěhování, včetně mezinárodní migrace) podmíněnou různými sociálními charakteristikami obyvatel obcí. Druhá je rozložení změn školy bez změny bydliště, které může odpovídat revidované volbě školy. Specifickým fenoménem pak je přestup z neúplných do úplných základních škol, který v jsme analyzovali zvláště Zatímco český pedagogický výzkum se dlouhodobě zaměřuje na nerovnosti spojené se sociálním statutem, a v poslední době také o mezikrajské rozdíly v kvalitě vzdělávání, naše analýza přináší navíc také poznatky o rozdílech příležitostí k volbě školy mezi městským a venkovským prostorem. Vzhledem k užití explorativní analýzy (EDA) klademe důraz na vizualizace dat včetně map.

5 Závěry a diskuse

Ač z principu pracujeme s anonymizovanými daty (z hlediska osob), vyvstává řada etických a právních otázek spojených ať už s otevřenou identitou organizací, nebo s možností nepřímé identifikace žáků u malých škol (a takových je v Česku velké množství).

Metodologicky se práce s velkými daty liší od standardního kvantitativního výzkumu např. v tom, že výzkumný problém se přizpůsobuje dostupným datům. Z hlediska vědecké integrity pak může být nejasná hranice mezi explorativní analýzou a „rybařením“ v datech.

Přes uvedené i další problémy výzkum přináší cenné výsledky a využívání administrativních dat pro výzkumné účely do budoucna poroste, stejně jako u jiných typů „velkých“ dat (analytiky učení, žákovská nebo učitelská komunikace na sociálních sítích, geolokační data o denním pohybu žáků apod.). Limitem je mj. absence bezvýznamových identifikátorů aktérů, ale situaci může významně změnit ne/zavedení jednotného rezortního informačního systému ve školství.

Poděkování

Tato prezentace je výstupem grantu GA ČR 20-18545S. Děkujeme pracovníkům odboru školské statistiky, analýz a informační strategie MŠMT ČR, kteří nám poskytli administrativní data.

6 Literatura

- Connelly, R., Playford, C., Gayle, V., & Dibben, C. (2016). The role of administrative data in the big data revolution in social science research. *Social Science Research*, 59, 1–12. <https://doi.org/10.1016/j.ssresearch.2016.04.015>
- Dvořák, D., Meyer, P., Kučerová, S. R., Vyhnálek, J. & Šmíd, O. (2020). Changing place, changing track: Inter-school mobility of Czech secondary students. *Journal of Pedagogy*, 11(1), 83–105. <https://doi.org/10.2478/jped-2020-0005>
- Himl, P. (2019). *Pozorovat, popsat, stvořit*. Argo.
- Kubáň, M. (2016, říjen). Otevřená data v České republice a v zahraničí. Předneseno na Semináři ÚV ČR. https://opendata.gov.cz/_media/edu:2016-10-11-miku-od.pdf
- Pattaro, S., Bailey, N., & Dibben, Ch. (2020). Using linked longitudinal administrative data to identify social disadvantage. *Social Indicators Research*, 147, 865–895. <https://doi.org/10.1007/s11205-019-02173-1>