



„Zabezpečení dat“

zápis z pracovní skupiny

Zápis shrnuje diskusi a hlavní argumenty v rámci pracovní skupiny k tématu „Zabezpečení dat“, která byla součástí konference „**Sběrem dat ke spravedlivému vzdělávání**“. Ta proběhla 21. října 2015 v Hrzánském paláci. Konference byla pořádána **Jiřím Dienstbierem, ministrem pro lidská práva, rovné příležitosti a legislativu, Nadací Open Society Fund Praha** a pod záštitou **Kateřiny Valachové, ministryně školství, mládeže a tělovýchovy**.

Zápis obsahuje diskuse během dvou setkání této pracovní skupiny na konferenci. Přestože vychází z poznámek zapisovatelky a stenozáznamu, nejedná se o přímý přepis diskuse. Po úvodních příspěvcích následuje diskuse s účastníky a účastnicemi konference. V rámci skupiny byly navrženy tyto okruhy otázek:

- Jak pracovat s citlivými údaji při jejich sběru?
- Jak bezpečně využívat citlivá data?

Na konec zápisu z každé pracovní skupiny jsou zařazeny příslušné závěry. Zápisy z dalších dvou pracovních skupin k tématu „jaká data potřebujeme sbírat“ a k metodologii sběru jsou obsaženy v samostatných dokumentech.

Moderátor:

David Beňák, *zástupce náměstkyně pro řízení, Sekce pro lidská práva, Úřad vlády ČR*

Úvodní příspěvky:

Klára Klementová, *oddělení analýzy a zpracování informací ve veřejném sektoru, Úřad pro ochranu osobních údajů*

Daniel Chytil, *Oddělení pracovních sil, migrace a rovných příležitostí, Český statistický úřad*

Pracovní skupina č. 1

D. Beňák uvedl účastníky do tématu a informoval o zaměření pracovní skupiny. Představil výše uvedené otázky a uvedl příspěvky Kláry Klementové a Daniela Chytila.

Klára Klementová

K. Klementová z oddělení analýzy a zpracování informací ve veřejném sektoru Úřadu pro ochranu osobních údajů seznámila účastníky pracovní skupiny se současným českým legislativním ošetřením problematiky zpracování citlivých a osobních údajů národnostních a etnických menšin, přičemž vyzdvihla především nutnost zohlednit jejich právo na soukromí. Zdůraznila také, že z pohledu zákona 101/2000 Sb. o ochraně osobních údajů a o změně některých zákonů je nutné rozlišovat mezi kategoriemi osobních údajů a citlivých údajů, což je významné pro legální ošetření sběru dat.

Osobním údajem je podle ustanovení § 4, písm. a) tohoto zákona jakákoli informace týkající se určené nebo určitelné fyzické osoby neboli subjektu údajů, na jejichž základě lze tyto osoby buď přímo, anebo nepřímo identifikovat (na základě čísla, kódu nebo jednoho či více prvků typických pro jeho psychickou, fyziologickou, kulturní anebo sociální identitu).

Citlivé údaje jsou potom podle ustanovení § 4, písm. b) téhož zákona jakékoliv informace týkající se určené nebo určitelné fyzické osoby neboli subjektu údajů, které vypovídají o jednom z jeho základních charakteristických znaků (tzn. o národnostním, etnickém anebo rasovém původu, o politických postojích, o členství v odborových organizacích, náboženství a filozofickém přesvědčení, o odsouzení za trestný čin, o zdravotním stavu nebo sexuálním životě subjektu údajů, přičemž může jít i o biometrický či genetický údaj).

Dále zdůraznila, že kategorie citlivých údajů jsou podmnožinou osobních údajů, přičemž pro zařazení údaje do této kategorie se údaj musí týkat identifikované nebo identifikovatelné fyzické osoby a musí splňovat parametry ustanovení § 4, písm. b). Je pro ně stanoven přísnější postup získávání a zpracovávání, jenž je určen konečným taxativním výčtem právních titulů, na jejichž základě je možné dané citlivé údaje zpracovávat. Tento výčet je uveden v § 9 zákona o ochraně osobních údajů, přičemž § 9 stanovuje také podmínky, za kterých lze tyto údaje zpracovávat, což je možné pouze s výslovným písemným souhlasem subjektu údajů. Evropským, mezinárodním i českým právem je garantována zvýšená ochrana citlivých údajů. Subjekt údajů musí být informován o základních parametrech zamýšleného zpracování, přičemž tyto informace musí být správcem poskytnuty nejpozději s udělením souhlasu subjektu ke zpracování. Správce údajů musí disponovat souhlasem subjektu údajů po celou dobu jejich zpracování, přičemž subjekt údajů má právo vzít souhlas se zpracováním kdykoliv zpět.

V souvislosti se zpracováním údajů národnostních a etnických menšin je z pohledu ÚOOÚ nepřijatelné vytváření jakýchkoli databází založených na národnostním či etnickém původu. Jediný přípustný nástroj pro tyto účely, například pro určení počtu žáků na ZŠ, kteří patří k nějakému etniku, je anonymní kvalifikovaný odhad; případně by například určení počtu žáků daného etnika mohlo být umožněno ještě na základě zákonného zmocnění, které však v současnosti neexistuje.

Daniel Chytil

D. Chytil, vedoucí Oddělení pracovních sil, migrace a rovných příležitostí Českého statistického úřadu, informoval především o problematičnosti a obtížnosti statistické práce v současném českém právním prostředí a zdůraznil skutečnost, že ČR má v současnosti nejpřísnější zákony na ochranu osobních údajů v EU.

Dále se věnoval problematice práce s citlivými a osobními daty v českém právním prostředí, především v kontextu pravidel pro zpracování dat ze sčítání lidu. Databáze ČSÚ ze sčítání se musí podle zákona o sčítání po třech letech natolik anonymizovat, že to v podstatě znemožňuje její využití pro další výzkumy. Data po uplynutí této doby neobsahují jména, adresy, rodná čísla ani data narození a není je možné ani propojit s žádnými dalšími údaji administrativních zdrojů dat, například jednotlivých ministerstev. Těchto zdrojů je v ČR obrovské množství, přičemž téměř všechny obsahují rodné číslo jako společný identifikátor. Stát má proto velmi dobré předpoklady nahradit rodné číslo nějakým anonymním identifikátorem, jež by nebylo možné použít k identifikaci konkrétních osob, a mohl by sloužit k propojení dat jednotlivých administrativních zdrojů. To ale kvůli právní úpravě ochrany osobních údajů není v současnosti možné. Funguje systém základních registrů, který ale neobsahuje funkci hromadného předávání údajů. Není proto možné například propojit údaje MŠMT s matrikami či údaji MPSV. To výrazně prodražuje statistická šetření, jež tak musí často probíhat opakovaně, což navíc zatěžuje i respondenty těchto šetření.

Pokud by toto chtěl stát zjednodušit a jednotlivé databáze propojit, s tím, že by údaje byly anonymizovány, musel by zásadně přehodnotit svůj přístup k ochraně osobních údajů. Například ve Švédsku toto propojení existuje, což statistickému úřadu umožňuje mimo jiné zjistit, kolik zástupců konkrétní etnické menšiny studuje, čerpá sociální dávky apod.; tato data jsou nicméně chráněna a běžný občan se k nim nedostane snáze nežli v ČR, kde propojování jednotlivých databází možné není, protože tato data v jednotlivých administrativních databázích již i v ČR existují. Nemožnost jejich propojování pak znemožňuje, například v kontextu školských matrik, sledovat dráhu dětí či vyhodnocovat úspěšnost konkrétních opatření například vůči sociálně znevýhodněným skupinám. Zdůraznil také, že tato data nejsou k dispozici ani o zástupcích většinové společnosti, a to nejen v případě osobních a citlivých dat, ale veškerých dat obecně, neboť například právě MŠMT je po každém sběru pouze jednorázově vyhodnotí a následně smaže, což je vysoce neefektivní.

Diskuse

D. Beňák poděkoval za úvodní příspěvky a shrnul základní třecí plochu problému, kterou identifikoval jako potřebu efektivní statistiky na jedné straně a potřebu spolehlivé ochrany osobních a citlivých údajů na straně druhé, přičemž zdůraznil, že základním imperativem, který je třeba brát v úvahu především, je veřejný zájem, jímž je v této oblasti jednoznačně ochrana osobních údajů, a tedy nikoli efektivní formulace politik, která by byla jedním z hlavních přínosů zefektivnění možností práce s těmito daty. Ústřední otázkou nicméně zůstává, jak umožnit efektivní statistiku při zaručení spolehlivé ochrany údajů.

Z. Doubravský se v tomto kontextu zeptal, proč nebude možné například u budoucího sčítání lidu v roce 2021 ze získaného datasetu následně vyškrtnout jméno a adresu a dále jej využívat jako data pro výzkum například romského etnika.

D. Chytil odpověděl, že to možné je a že tato data po anonymizaci jsou k dispozici i z minulého sčítání, přičemž jednotlivé proměnné z tohoto datasetu propojit lze, ale již není možné je propojit například se školskou matrikou romských dětí a dále pomocí nich sledovat například jejich vývoj v čase.

Dále pak D. Chytil informoval, že evropské nařízení o statistice v § 17 určuje, že statistické úřady mají mít přístup k administrativním zdrojům a mají se účastnit procesu sestavování dat, ale že iniciativy ČSÚ o přístup k nim většinou troskotají na souhlasu ÚOOÚ.

K. Klementová se v reakci na to zeptala, co považuje ČSÚ za hlavní problém: zda je to nutnost anonymizace dat po třech letech, nebo zda je hlavním problémem nemožnost propojení jednotlivých databází.

D. Chytil upřesnil, že jednoznačně nemožnost propojení jednotlivých databází a že za problém rozhodně nepovažuje fungování ÚOOÚ, ale legislativní ošetření problematiky, která je sice neefektivnější v EU v ochraně osobních údajů, ale znemožňuje efektivní plánování politik.

J. Macháčka vznesl dotaz na dopad připravované změny ochrany osobních údajů na evropské úrovni, na jejímž základě má dojít k nahrazení národních zákonů o ochraně osobních a citlivých údajů jednotným evropským nařízením. K. Klementová bohužel těmito informacemi nedisponovala.

D. Beňák shrnul dosavadní závěry z debaty. Je tedy možné s citlivými a osobními daty pracovat, ale je nutné disponovat souhlasem subjektů údajů dat, což je sice systémově možné, ale komplikované, protože to může snižovat validitu dat vzhledem k omezení zahrnutého vzorku.

I. M. Čonková v kontextu častých argumentů o obavách ze sběru etnických dat poukázala na historickou zkušenost s genocidou, která dle jejího názoru tyto obavy do velké míry způsobuje, přičemž tento strach považuje za relevantní i s přihlédnutím k popularitě různých extrémně pravicových hnutí. Položila proto otázku, jak zajistit, aby bylo možné efektivně získávat data pro tvoření účinnějších politik směřujících k zrovnoprávnění menšin a aby zároveň nebylo možné tato data zneužívat.

Dle D. Chytila by zajištění efektivního sběrů dat vyžadovalo především podpoření mediálních kampaní o možných přínosech sběru citlivých dat, jež by zároveň přesvědčivým způsobem vysvětlovaly, k čemu budou data sloužit, a zároveň poskytovaly jednoznačné informace o tom, jaké kroky budou podniknuty pro zabránění jejich zneužití (například informace o skartaci odpovědních archů či zabezpečení databází ČSÚ); je nicméně jasné, že jde o problematickou otázku. V hypotetickém případě, kdy by obyvatel uvedl, že je Rom, a pocházel by přitom z velmi malé obce, byl by teoreticky dohledatelný – zajištění nepřekonatelného zabezpečení tak samozřejmě v podstatě není reálné, ale to není možné zajistit ani v současné situaci, kdy příslušná data propojována nejsou.

M. Cichá upozornila na problematiku aplikace kvalifikovaného odhadu na případech odhadů počtu romských žáků ve školách a zdůraznila především skutečnost, že mnoho ředitelů nepřízná, že na škole romské žáky má, a ty kvůli tomu pak například nemohou být podpořeny stipendii. To považuje za jednu z hlavních překážek spolehlivosti využívání kvalifikovaného odhadu pro získávání etnických dat.

K. Klementová vyjádřila pochybnost nad možnými alternativami sčítání dětí ve školách kvalifikovaným odhadem; bez souhlasu subjektů dat by dle ní žádný jiný způsob nebyl možný, přičemž také vyjádřila pochybnost nad tím, kolik lidí bude ochotno tento souhlas pro sčítání svých dětí udělit.

Z. Doubravský zdůraznil, že mnoho lidí odmítá posuzovat lidi na základě etnické či jiné sociální příslušnosti.

D. Beňák v odpovědi apeloval na skutečnost, že kvalifikovaný odhad je v současnosti nejspolehlivějším zdrojem těchto dat, minimálně v kontextu určování etnické příslušnosti dětí ve školách – pochopitelně v případě, že jej provádí výzkumník, který neodmítá rozlišování žáků na základě etnické příslušnosti.

Závěry:

- 1) S citlivými daty pracovat lze, nicméně je potřeba dodržovat příslušná pravidla, která jsou určitou překážkou, ale ne nepřekonatelnou.
- 2) Pokud chceme, aby práce s těmito daty byla efektivnější a snazší, pak musí dojít k poměrně zásadní legislativní změně.
- 3) Osobní a citlivá data, i v případě usnadnění práce s nimi, efektivně zabezpečovat lze, ale s největší pravděpodobností není možné zaručit, že je za žádných okolností nebude možné zneužít; to ovšem nelze spolehlivě zaručit ani za současného stavu oddělení jednotlivých administrativních zdrojů dat.
- 4) Sběr těchto dat je nicméně důležitý a stávající stav je spíše v neprospěch Romů a dalších skupin.
- 5) K tomu, aby tento stav bylo možné zlepšovat, je potřeba vést vysvětlující kampaň, která by uceleně informovala o tom, proč je důležité dávat souhlas ke zpracování osobních údajů.

Pracovní skupina č. 2

Úvodní příspěvky se obsahově do velké míry kryjí s úvodními příspěvky z předchozí části. Kvůli některým odlišnostem byly však znovu zařazeny v celém jejich rozsahu.

D. Beňák uvedl účastníky do tématu a informoval o zaměření pracovní skupiny. Představil výše uvedené otázky, a uvedl příspěvky Kláry Klementové a Daniela Chytila.

Klára Klementová

K. Klementová z oddělení analýzy a zpracování informací ve veřejném sektoru Úřadu pro ochranu osobních údajů seznámila účastníky pracovní skupiny se současným českým legislativním ošetřením problematiky zpracování citlivých a osobních údajů národnostních a etnických menšin, přičemž vyzdvihla především nutnost zohlednit jejich právo na soukromí. Zdůraznila také, že z pohledu zákona 101/2000 Sb. o ochraně osobních údajů a o změně některých zákonů je nutné rozlišovat mezi kategoriemi osobních údajů a citlivých údajů, což je významné pro legální ošetření sběru dat.

Osobním údajem je podle ustanovení § 4, písm. a) tohoto zákona jakákoli informace týkající se určené nebo určitelné fyzické osoby neboli subjektu údajů, na jejichž základě lze tyto osoby buď přímo, anebo nepřímo identifikovat (na základě čísla, kódu nebo jednoho či více prvků typických pro jeho psychickou, fyziologickou, kulturní anebo sociální identitu).

Citlivé údaje jsou potom podle ustanovení § 4, písm. b) téhož zákona jakékoliv informace týkající se určené nebo určitelné fyzické osoby neboli subjektu údajů, které vypovídají o jednom z jeho základních charakteristických znaků (tzn. o národnostním, etnickém anebo rasovém původu, o politických postojích, o členství v odborových organizacích, náboženství a filozofickém přesvědčení, o odsouzení za trestný čin, o zdravotním stavu nebo sexuálním životě subjektu údajů, přičemž může jít i o biometrický či genetický údaj).

Dále zdůraznila, že kategorie citlivých údajů jsou podmnožinou osobních údajů, přičemž pro zařazení údaje do této kategorie se údaj musí týkat identifikované nebo identifikovatelné fyzické osoby a musí splňovat parametry ustanovení § 4, písm. b). Je pro ně stanoven přísnější postup získávání a zpracovávání, jež je určen konečným taxativním výčtem právních titulů, na jejichž základě je možné dané citlivé údaje zpracovávat. Tento výčet je uveden v § 9 zákona o ochraně osobních údajů, přičemž § 9 stanovuje také podmínky, za kterých lze tyto údaje zpracovávat, což je možné pouze s výslovným písemným souhlasem subjektu údajů. Evropským, mezinárod-

ním i českým právem je garantována zvýšená ochrana citlivých údajů. Subjekt údajů musí být informován o základních parametrech zamýšleného zpracování, přičemž tyto informace musí být správcem poskytnuty nejpozději s udělením souhlasu subjektu ke zpracování. Správce údajů musí disponovat souhlasem subjektu údajů po celou dobu jejich zpracování, přičemž subjekt údajů má právo vzít souhlas se zpracováním kdykoliv zpět.

V souvislosti se zpracováním údajů národnostních a etnických menšin je z pohledu ÚOOÚ nepřijatelné vytváření jakýchkoli databází založených na národnostním či etnickém původu. Jediný přípustný nástroj pro tyto účely, například pro určení počtu žáků na ZŠ, kteří patří k nějakému etniku, je anonymní kvalifikovaný odhad; případně by například určení počtu žáků daného etnika mohlo být umožněno ještě na základě zákonného zmocnění, které však v současnosti neexistuje.

Daniel Chytil

D. Chytil, vedoucí Oddělení pracovních sil, migrace a rovných příležitostí Českého statistického úřadu, informoval především o problematičnosti a obtížnosti statistické práce v současném českém právním prostředí a zdůraznil skutečnost, že ČR má v současnosti nejpřísnější zákony na ochranu osobních údajů z EU.

Dále se věnoval problematice práce s citlivými a osobními daty v českém právním prostředí, především v kontextu pravidel pro zpracování dat ze sčítání lidu. Databáze ČSÚ ze sčítání se musí podle zákona o sčítání po třech letech natolik anonymizovat, že to v podstatě znemožňuje její využití pro další výzkumy. Data po uplynutí této doby neobsahují jména, adresy, rodná čísla ani data narození a není je možné ani propojit s žádnými dalšími údaji administrativních zdrojů dat, například jednotlivých ministerstev. Těchto zdrojů je v ČR obrovské množství, přičemž téměř všechny obsahují rodné číslo jako společný identifikátor. Stát má proto velmi dobré předpoklady nahradit rodné číslo nějakým anonymním identifikátorem, jež by nebylo možné použít k identifikaci konkrétních osob, a mohl by sloužit k propojení dat jednotlivých administrativních zdrojů. To ale kvůli právní úpravě ochrany osobních údajů není v současnosti možné. Funguje systém základních registrů, který ale neobsahuje funkci hromadného předávání údajů. Není proto možné například propojit údaje MŠMT s matrikami či údaji MPSV. To výrazně prodražuje statistická šetření, jež tak musí často probíhat opakovaně, což navíc zatěžuje i respondenty těchto šetření.

Pokud by toto chtěl stát zjednodušit a jednotlivé databáze propojit, s tím, že by údaje byly anonymizovány, musel by zásadně přehodnotit svůj přístup k ochraně osobních údajů. Například ve Švédsku toto propojení existuje, což statistickému úřadu umožňuje například zjistit, kolik zástupců konkrétní etnické menšiny studuje, čerpá sociální dávky apod.; tato data jsou nicméně chráněna a běžný občan se k nim nedostane snáze nežli v ČR, kde propojování jednotlivých databází možné není, protože tato data v jednotlivých administrativních databázích již i v ČR existují. Nemožnost jejich propojování pak znemožňuje, například v kontextu školských matrik, sledovat dráhu dětí či vyhodnocovat úspěšnost konkrétních opatření například vůči sociálně znevýhodněným skupinám. Zdůraznil také, že tato data nejsou k dispozici ani o zástupcích většinové společnosti, a to nejen v případě osobních a citlivých dat, ale veškerých dat obecně, neboť například právě MŠMT je po každém sběru pouze jednorázově vyhodnotí a následně smaže, což je vysoce neefektivní.

V kontextu sběru etnických dat romské menšiny dále informoval o vyšší efektivitě sběru etnických dat v případech, kdy byli do sběru zapojeni romští asistenti, kteří respondentům podávali informace o přínosech uvedení etnické příslušnosti a odpovídali na případné pochybnosti a obavy respondentů. Považuje proto vedení vysvětlující kampaně za potenciálně velice přínosný

způsob usnadnění sběru etnických dat. Dále informoval také o obvyklé nevoli obyvatelstva a zdůraznil často se vyskytující otázku, proč stát lid sčítá, když danými informacemi již disponuje, a připomněl problematiku nemožnosti efektivně využívat administrativní zdroje dat pro účely statistických šetření. Závěrem pak zdůraznil, že je pro stát velice obtížné efektivně přistupovat k minoritě, když ji v podstatě není schopen rozlišit od majority, o níž také nemá údaje.

Diskuse

D. Beňák poděkoval za úvodní příspěvky a požádal o zprostředkování slovenské zkušenosti se získáváním etnických dat v konkrétních lokalitách zástupce Slovenského statistického úradu.

J. Kling Informoval o plánovaných výzkumech SŠÚ, který ve spolupráci se Zmocněncem pro Romy vytvořil datovou strukturu pro národní projekt k monitorování integrace Romů, v němž jsou vyčleněny finance i na sběr dat. Půjde především o výběrová šetření a plošná mapování romských koncentrací.

D. Chytil se v tomto kontextu zeptal, zdali SŠU měl pro tyto účely možnost využití dat ze sčítání jako opory výběru.

J. Kling informoval o záměru uskutečnit toto šetření při realizaci EU-SILC, přičemž ještě není rozhodnuto, zda to bude učiněno v rámci podvzorku či nadvzorku šetření EU-SILC v rámci Slovenska. Šetření v rámci nadvzorku by umožnilo srovnat data romské menšiny s všeobecnou populací, v níž jsou však pochopitelně zahrnuti i Romové. Co se týče romského vzorku, bude pravděpodobně využit datový rámec ze sčítání a Atlasu romských komunit, připraveného i za účasti SŠÚ, který se ale zapojoval pouze na metodické úrovni. Atlas romských komunit pracuje s údaji sbíranými na úrovni komunit a neobsahuje tak individuální data, tedy nevztahuje se na něj ochrana osobních údajů, protože nesbíral údaje o jednotlivcích a při sběru pracoval s expertními odhady. Práce na výzkumu EU-SILC však předpokládá práci s daty i na osobní úrovni a v případě romského vzorku bude vycházet ze stejného nastavení, jež má standardní EU-SILC, tak aby odpovídal všem zákonům o ochraně osobních údajů a o statistice. Na základě vyslechnutých příspěvků pak vyjádřil svůj předpoklad o tom, že české zákony budou pravděpodobně v těchto záležitostech přísnější než slovenské.

D. Chytil odpověděl, že v českém kontextu by například jako výběr pro EU-SILC stejně jako pro LFS mohlo sloužit pouze územní dělení podle registru sčítacích obvodů, z nichž se náhodně vybere a v jejichž rámci budou dále vyselektovány menší jednotky. ČSÚ tedy oporu výběru vzorku nemá ani ve sčítání, ani v evidenci, protože data ze sčítání není pro tyto účely možné použít. Jako další téma do debaty vnesl otázku možnosti identifikovat ve výběrovém šetření Romy a například je selektovat pro jejich vyšší zastoupení ve vzorku šetření (*sample booster*).

J. Kling v tomto kontextu informoval o šetření EU-MIDIS (Evropský průzkum o menšinách a diskriminaci), který v současnosti probíhá i v ČR a cílovou skupinou zde jsou primárně právě Romové. Sčítání z hlediska počtu Romů nedává příliš smysl (vzhledem k velmi nízkému počtu deklarování této etnické příslušnosti), nicméně předpokládá, že by data ze sčítání mohla být využita pro identifikaci prostorového rozložení populace Romů, na kterém by se mohlo dál stavět.

D. Chytil v reakci zdůraznil, že by to vyžadovalo přijmout předpoklad, že míra sebe-deklarace je územně všude stejná.

J. Kling uznal, že to není úplně neproblematické, nicméně pro výběr místa dotazování z toho lze vycházet vzhledem k tomu, že výzkum pak dále vybírá respondenty a domácnosti v této lokalitě

a umožní to tedy zpřesnění zaměření na romskou populaci, identifikovanou prostřednictvím expertního odhadu v takto vybrané lokaci i v případech nízké koncentrace této populace.

D. Beňák v kontextu debaty o dostupnosti dat pro výběr vzorku v českém prostředí poukázal na Gabalovu analýzu a Agenturu pro sociální začleňování, která nějakými daty také disponuje.

J. Kling uvedl, že v případě ambice zahrnout do vzorku i populaci žijící mimo sociálně vyloučené lokality považuje za nejvhodnější využít metodu „procházky“ v lokalitě, o které lze předpokládat, že se zde romská populace vyskytuje. Dále uvedl i některé problematické aspekty takového postupu, například vysokou časovou a finanční náročnost. Zdůraznil ale i možnost čerpat pro tyto účely strukturální fondy a svůj předpoklad o tom, že v českém případě by pravděpodobně bylo možné zajistit financování sběru dat i prostřednictvím technické pomoci. V souvislosti s omezenými možnostmi sběru osobních dat a využívání registrů navrhuje pro účely monitorování a hodnocení naplňování národní strategie pro integraci Romů více využívat jiné než administrativní zdroje dat, například právě výběrové průzkumy, které považuje za informačně bohaté a zároveň politicky mnohem méně problematické.

E. Klepáčková (IQ Roma servis, Katedra sociologie MU) se dotazovala na přístup k problematice sběru etnických dat v rámci evropských projektů a na metodologický a etický rozměr otázek, jak s těmito daty pracovat a nakládat. Důvodem této otázky byla především skutečnost, že podpora mnoha projektů je vázána právě na informaci o počtu podpořených zástupců určité menšiny, což s sebou nese problém, jak zajistit, aby to bylo v souladu s pravidly sběru osobních a citlivých dat.

K. Klementová apelovala na nutnost přihlídnutí k velkému riziku zneužitelnosti dat v případě, že tato obsahují jména. Proto zdůraznila kvalifikovaný odhad jako z jejího pohledu nejméně problematickou metodu sběru takových dat.

D. Beňák v tomto kontextu informoval o programu OP VVV, který jako jediný indikátor využívá právě počet podpořených romských dětí, přičemž vychází z předpokladu, že bude naplňován právě kvalifikovaným odhadem. Tedy, že příjemce podpory bude kvalifikovaně odhadovat počet podpořených romských dětí. Vyjádřil také své přesvědčení, že za předpokladu, že k tomu příjemci podpory budou přistupovat opravdu zodpovědně, mohlo by jít o jeden z nejpřesnějších údajů o počtu romských žáků. Získávání dat na individuální úrovni je samozřejmě problematické a obtížné, nicméně kvůli možnostem čerpání šablon budou takto získaná data pokrývat velmi významný vzorek populace, a pokud by se například prostřednictvím MZ podařilo promítnout data pojištěnoven do území, mohlo by to poskytovat velice komplexní obraz.

J. Kling zdůraznil, že jen velmi málo zemí sbírá nějaké údaje čistě o Romech, přičemž ale zároveň mají povinnost informovat o tom, jak se mění situace této skupiny obyvatel. Žádná ze zemí ale nemá dostatečně fundované informace o tom, kolik Romů podpořila a jak se změnila jejich situace. V tomto kontextu vyjádřil své přesvědčení o vhodnosti operování s expertními odhady i přesto, že nemají takovou váhu jako statistická data, protože jsou i tak schopna poskytovat velice relevantní informace. Podle jeho názoru lze s citlivými a osobními údaji pracovat i v současném legislativním nastavení. Uvedl také, že mnoho projektů, a to i poměrně rozsáhlých, jež tato data sbíraly, již uskutečněno bylo. Při práci s nimi je nicméně vždy potřeba disponovat souhlasem subjektu údajů a informovat ÚOOÚ o konkrétním nastavení, záměrech, metodách zabezpečení, atd.

Závěry

- 1) Metoda výběrových šetření může být zajímavým doplňkem systému monitorování strategie integrace Romů, s tím, že samozřejmě vyžaduje vyčlenění určitých finančních prostředků.
- 2) Jestliže existuje tolik překážek, jež brání sběru osobních a citlivých dat na úrovni jednotlivce, je vhodné zaměřit se na jiné metody, například na úrovni skupin či za pomoci expertních odhadů komunit, a nebojovat tak za každou cenu s překážkami, které jsou v současnosti pevně dané.
- 3) Je potřeba vést vysvětlovací kampaně o přínosech poskytování souhlasů subjektů dat a věnovat se způsobům zajištění maximálního zabezpečení těchto dat.
- 4) Je nicméně zároveň nutné v maximální možné míře respektovat veřejný zájem, kterým je v této otázce jednoznačně ochrana soukromí.