

**OZNÁMENÍ O ZJIŠTĚNÍ SKUTEČNOSTÍ,
KTERÉ NASVĚDČUJÍ O PŘÍPRAVĚ TRESTNÉHO ČINU**
dle § 180 (neoprávněné nakládání s osobními údaji) trestního zákoníku
v souladu s § 367 trestního zákoníku

Podezřelý:

Český statistický úřad – organizační složka státu, 10082 Praha 10 – Strašnice (Praha 10), Na padesátém 3268/81, IČ 00025593 (dále jen ČSÚ), případně jeho neznámý odpovědný pracovník.

Stručný popis skutkové podstaty:

Podezření se týká přípravy neoprávněného nakládání s osobními údaji dle § 180 trestního zákoníku, které spatřuji v tom, že při Sčítání lidu, domů a bytů 2011¹ (dále jen Sčítání lidu) ČSÚ plánuje neprovést anonymizaci dat a tedy plánuje s osobními údaji dále neoprávněné nakládat – zejména je uchovávat. Jedná se přitom o osobní údaje cca 10 milionů lidí v ČR včetně citlivých osobních údajů (např. národnost nebo náboženství), tedy případný trestní čin by byl velkého rozsahu.

Podrobný popis a zdůvodnění:

- 1) **Anonymizaci dat předepisuje zákon o sčítání lidu** (296/2009 Sb. v § 22, odst. 4). Nestanoví přesný postup anonymizace – volbu způsobu ponechává na tom, kdo má povinnost anonymizaci provést – tedy na ČSÚ. Přesto ale provedení anonymizace vyžaduje. Tuto povinnost v obecné rovině stanovuje i zákon 101/2000 Sb. (o ochraně osobních údajů) v § 5, odst. 4.
- 2) Anonymizace jako pojem není přímo definovaný a jeho gramatickým výkladem² dojdeme k tomu, že se jedná o **proces, který ze souborů (zejména osobních) údajů udělá soubor anonymních údajů.**
- 3) **Anonymní údaj** je již pojem definovaný v § 2 zákona č. 89/1995 Sb. (o státní statistické službě) a obdobně i podle zákona na ochranu osobních údajů (v § 4 zákona 101/2000 Sb.):

Citace § 2 zákona č. 89/1995 Sb.:

a) individuálním údajem údaj týkající se jednotlivé právnické nebo fyzické osoby, získaný pro statistické účely podle tohoto zákona,

¹ dle zákona 296/2009 Sb.

² tento výklad je podpořen i článkem vydaný Úřadem na ochranu osobních údajů, který uvádím v doplňujících informacích pod bodem 1)

- c) *anonymním údajem takový individuální údaj, který buď v původním tvaru nebo ani po provedeném zpracování neumožňuje přímou, popřípadě nepřímou identifikaci jednotlivé právnické nebo fyzické osoby,*
- d) *přímou identifikací identifikace jednotlivé právnické nebo fyzické osoby na základě jakéhokoliv identifikátoru, který lze použít bez nutnosti vynaložit ke zjištění identity dané osoby nepřiměřeného času a úsilí,*
- e) *nepřímou identifikací možnost odvodit totožnost jednotlivé právnické nebo fyzické osoby jiným způsobem než způsobem uvedeným v písmenu d),*

Citace § 4 zákona 101/2000 Sb.:

- a) *osobním údajem jakákoliv informace týkající se určeného nebo určitelného subjektu údajů. Subjekt údajů se považuje za určený nebo určitelný, jestliže lze subjekt údajů přímo či nepřímo identifikovat zejména na základě čísla, kódu nebo jednoho či více prvků, specifických pro jeho fyzickou, fyziologickou, psychickou, ekonomickou, kulturní nebo sociální identitu,*
- c) *anonymním údajem takový údaj, který buď v původním tvaru nebo po provedeném zpracování nelze vztáhnout k určenému nebo určitelnému subjektu údajů,*
- d) *subjektem údajů fyzická osoba, k níž se osobní údaje vztahují,*

Obě definice, ač jsou formulovány odlišně, se shodují na tom, že podstatnou vlastností anonymních údajů je, že **nesmí být možno žádným způsobem identifikovat osobu**³, které se údaj týká. Nakonec toto ČSÚ uvádí na svém webu⁴ (citují „...ale nikdo už nikdy nezjistí, jak se jmenuje, ani jaké má rodné číslo“). Zástupci ČSÚ to uvádí i v médiích.

4) ČSÚ, resp. jeho odpovědní pracovníci, se chystají provést „anonymizaci“ způsobem, kdy odstraní z údajů sčítacího dotazníků pouze **jméno, příjmení, rodné číslo, den a měsíc narození**. Doložit to lze např. těmito prohlášeními odpovědných pracovníků ČSÚ.

- Vyjádření Stanislava Drápala, místopředsedy ČSÚ, ve věci anonymizace dat ze Sčítání lidu⁵:
 - Jan Vobořil 9.3. 23:57: *Dobrý den, můžete mi říci, jaké údaje budou odstraněny z elektronických formulářů při jejich anonymizaci?*
 - Stanislav Drápal 14:07: *Dobrý den, při anonymizaci budou odstraněny údaje o jménu a příjmení, datu narození a rodné číslo. Papírové formuláře musí být podle zákona skartovány nejpozději do tří let od rozhodného okamžiku (půlnoc z 25. na 26. března). Veškeré údaje zůstanou pouze v elektronické podobě a ve stejném termínu musí být předány národnímu archivu.*
- Popis anonymizace dat za webu ČSÚ⁶:
 - *Jakmile budou papírové formuláře naskenovány, dojde před samotným zpracováním k jejich anonymizaci. To v praxi znamená, že budou z databází odstraněna všechna jména a rodná čísla.*

³ tento výklad je podpořen i článkem vydaný Úřadem na ochranu osobních údajů, který uvádím v doplňujících informacích pod bodem 1)

⁴ zdroj: http://www.czso.cz/sldb2011/redakce.nsf/i/scitani_lidu_bezpecnost_na_vsech_frontach

⁵ zdroj: <http://ekonom.ihned.cz/c1-51022320-ptejte-se-na-scitani-lidu-2011>

⁶ zdroj: http://www.czso.cz/sldb2011/redakce.nsf/i/scitani_lidu_bezpecnost_na_vsech_frontach

- Odpověď Barbory Serbusové, pracovnice ČSÚ na dotaz, jakým způsobem probíhá anonymizace dat při sčítání lidu 2011:
 - *Dle § 22 zákona č. 296/2009 Sb., o sčítání lidu, domů a bytů, budou sčítací formuláře po ukončení zpracování výsledků sčítání zařazeny do skartačního řízení. Skartace sčítacích formulářů proběhly dle zákona do 3 let od rozhodného okamžiku. Český statistický úřad však přepokládá, že formuláře budou skartovány dříve v návaznosti na postup zpracování výsledků. Sčítací formuláře budou převedeny do elektronické podoby (metodou optického snímání), a tyto formuláře spolu s elektronickými formuláři vyplněnými povinnými osobami budou ve stanoveném okamžiku zpracování dat anonymizovány. Neanonymizovaná data jsou nutná pro vyloučení duplicit v datovém souboru a pro konstrukci domácností. Proces anonymizace dat spočívá v odstranění jména a příjmení osoby, dále rodného čísla a měsíce a dne narození (z data narození bude vypočítán věk v dokončených letech).*
- Stejně informace podávají proškolení operátoři i na oficiální infolince Sčítání lidu.

- 5) V datech tak **z bude** mimo jiné **přesná adresa až na úroveň bytu, věk osoby v celých letech, pohlaví.**
- 6) Podle adresy, věku a pohlaví **není problém v mnoha případech identifikovat přímo konkrétní osobu.** V jednom bytě se málo kdy nachází více osob se stejným věkem v celých letech a pohlavím. Existuje i velké množství bytů či domů, kde bydlí pouze jedna osoba. Informace o bydlišti (ať trvalém nebo faktickém – stejně se ve většině případů shodují) jsou u celé řady osob veřejně dostupné⁷. Připouštím, že v některých případech identifikovat osobu nelze. Anonymizovat je ale třeba každý údaj, takže je irelevantní, u kolika dotazníků lze identifikovat osobu a u kolika nikoli. Po anonymizaci nesmí existovat ani jeden jediný dotazník, u kterého by bylo možné osobu identifikovat. To zjevně splněno není.
- 7) Z uvedených informací vyplývá, že uvedený postup odstranění uvedených údajů nespĺňuje kritéria anonymizace a **nelze jej za anonymizaci považovat.**

Doplňující informace:

- 1) Článek⁸ Úřadu na ochranu osobních údajů (dále jen ÚOOÚ) k pojmu osobní údaj z roku 2001 (aktualizován v květnu 2010), který vysvětluje mimo jiné pojem anonymizace. Důležité části jsou vyznačeny tučně:

K pojmu osobní údaj

Zákon č.101/2000 Sb., o ochraně osobních údajů a o změně některých zákonů, ve znění pozdějších předpisů, definuje v § 4 písmeno a) pojem osobní údaj takto: „Osobním údajem (se rozumí) jakákoliv informace týkající se určeného nebo určitelného subjektu údajů. Subjekt údajů se považuje za určený nebo určitelný, jestliže lze subjekt údajů přímo či nepřímo identifikovat zejména na základě čísla, kódu nebo jednoho či více prvků, specifických pro jeho fyzickou, fyziologickou, psychickou, ekonomickou, kulturní nebo sociální identitu.“

⁷ telefonní seznam, katastr nemovitostí, živnostenský rejstřík apod.

⁸ zdroj: <http://www.uoou.cz/uoou.aspx?menu=14&loc=365>

Úřad pro ochranu osobních údajů (dále jen „Úřad“) považuje za důležité úvodem upozornit na to, že nelze zaměňovat pojem osobní údaj a pojem projev osobní povahy, jak je upraven v § 11 a dalších občanského zákoníku. Jde o dva samostatné právní instituty, přičemž však projev osobní povahy může za určité situace obsahovat osobní údaj (např. podobizna, písemnost osobní povahy nebo zvukový projev) a za takové situace dochází k jejich překrývání.

Podle výše uvedené definice tedy v případě, pokud fyzická osoba může být ze shromážděných údajů nebo na jejich základě jiným způsobem přímo nebo nepřímo identifikována, pak tyto údaje jsou údaji osobními. Znakem osobního údaje tedy je, že vypovídá o subjektu údajů, který nelze zaměnit s jiným subjektem údajů. Definice osobního údaje je velmi široká, což ostatně vyplývá z rozmanitosti zpracovávání osobních údajů v praktickém životě.

Na definici osobních údajů však není možné pohlížet izolovaně, bez znalosti dalších okolností zpracování osobních údajů.

Jak je patrné, základním kritériem pro posouzení toho, zda se jedná o osobní údaj či nikoliv, je okolnost možné zjištění identity subjektu údajů (určenost nebo určitelnost). Je tedy nutno vycházet ze skutečnosti, zda správce či jiná osoba může vytvořit přímou vazbu mezi údajem a fyzickou osobou. Tento vztah se vytváří převážně pomocí přesného identifikátoru a/nebo pomocí několika jiných údajů, které jednoznačnou identifikaci fyzické osoby tvoří (přímá identifikace). **Při posuzování možnosti fyzickou osobu identifikovat je však nutno vycházet ze všech možností odlišit od sebe jednotlivé fyzické osoby. Tedy nikoliv pouze omezeným pohledem např. na obsah zpracovávaných dat v konkrétním zpracování (např. obsah databáze), ale také zjištěním dalších možností subjekt údajů identifikovat. Jedná-li se např. o správce, který má v držení soubor identifikující subjekty údajů, nemůže být žádný jiný soubor, umožňující propojení na soubor s identifikátory subjektů údajů, považován za anonymní, a to ani v případech, když by zpracovávaný soubor přímou identifikaci subjektu údajů neobsahoval. Obdobně však toto platí i o možnosti správce získat identifikátor ze souborů, které nejsou v jeho držení např. z veřejných registrů, od zpracovatele, s nímž má uzavřenu smlouvu apod. V těchto případech se, v souladu s definicí, jedná o nepřímou možnost zjištění identifikaci subjektu údajů (nepřímá identifikace). Není rozhodující, zda této možnosti správce má v úmyslu využít či nikoliv. Rovněž není rozhodující, že správce nesmí sdružovat osobní údaje shromážděné pro rozdílné účely.**

Příklad: Jedná-li se o správce, který provádí zpracování osobních údajů pro účely zdravotního pojištění, je zřejmé, že pro splnění tohoto účelu má v držení soubor údajů o zdravotních pojištěncích s jejich úplnou a jednoznačnou identifikací. Pokud by tento správce zpracovával další soubor údajů a tento jiný soubor by kterýmkoli znakem (údajem nebo údaji) umožnil propojení na soubor zdravotních pojištěnců a umožnil tak identifikovat subjekt údajů v tomto dalším souboru, pak se i v tomto případě jedná o zpracování osobních údajů.

Pro správce, kteří uchovávají soubory určených nebo určitelných (identifikovaných či identifikovatelných) fyzických osob, bude z těchto důvodů vždy poměrně obtížné vytvořit z takových souborů osobních údajů jiné soubory s údaji anonymizovanými. **Anonymizace je jedním ze způsobů ochrany osobních údajů. Pokud jsou pro určitá zpracování odstraněny údaje, jimiž jsou nebo mohou být subjekty údajů identifikovány, pak tato anonymizace musí být provedena způsobem, který zcela znemožní jakoukoliv zpětnou identifikaci subjektů údajů. Toto platí i pro předávání údajů jiným**

subjektům. Pokud by soubor údajů byl jedním správcem předán jako anonymní jinému správci, pak nový správce by musel opět splňovat požadavek, že nemá možnost k získaným údajům přiřadit další osobní údaje, pomocí nichž by bylo možné subjekty údajů identifikovat. Obdobně se tyto požadavky vztahují také na zpracovatele.

Příklad: Pokud by zdravotnické zařízení zpracovávalo anonymizované soubory subjektů údajů (pacientů), ale zároveň by mělo přístup k národnímu zdravotnímu registru, z něhož by bylo možno získat identifikaci těchto osob (a jednoznačně ji k nim přiřadit), pak by se nejednalo o zpracování osobních údajů anonymních.

Příklad: Pokud by statistický úřad zveřejnil anonymní údaje, ale kterýkoliv příjemce by měl možnost přiřadit k nim další informace (např. geografické) a tím umožnit jednoznačné rozlišení občanů či rodin, pak by se nejednalo o zveřejnění anonymních údajů.

Z výše uvedeného vyplývá, že subjekt údajů nemusí být identifikován pouze jménem, příjmením a adresou (přímá identifikace). Jakákoliv jiná kombinace osobních údajů, která umožní odlišit od sebe jednotlivé fyzické osoby (subjekty údajů), musí být považována za identifikaci subjektu údajů (nepřímá identifikace). Je však nutno vždy posoudit minimální kombinaci osobních údajů, která již identifikaci umožní. Nicméně např. kombinace osobních údajů jméno, příjmení, adresa a datum narození téměř vždy jednoznačně subjekt údajů identifikují. Ve většině případů zpracování osobních údajů je však znalost jména, příjmení a adresy dostatečným předpokladem pro určení nebo určitelnost subjektu údajů.

Příklad: V některých případech nemusí postačovat k rozlišení subjektů údajů pouze jméno, příjmení a adresa (např. více generací jedné rodiny a téhož jména a příjmení bydlících na stejné adrese). Pokud správce takovou možnost vzhledem k účelu zpracování (např. doručování zásilek) předpokládá, pak je možno k těmto osobním údajům přiřadit další údaje (např. datum narození), který jednoznačně rozlišení umožní. Pokud by však buď z rozsahu souboru, nebo z účelu zpracování možnost jednoznačného rozlišení subjektů údajů nevyplývala, nelze doplňující údaj (datum narození) považovat za součást nezbytných údajů pro identifikaci. Takový údaj by naopak mohl být posuzován jako nadbytečný pro splnění stanoveného účelu.

Ve smyslu definice je tedy určitelnost stav, kdy na základě osobních údajů, které má správce k dispozici, nebo z osobních údajů, k nimž má přístup, lze fyzickou osobu jednoznačně odlišit od jiných fyzických osob. Určenost je potom stav, kdy správce má k dispozici přímé identifikační údaje (např. jméno, příjmení a adresu) subjektu údajů.

Je nutno upozornit také na skutečnost, že osobní údaj se nemusí vždy vztahovat pouze k jedné fyzické osobě. V praktickém životě jsou situace, kdy se nějaký údaj vztahuje k více osobám nebo ke skupině osob – např. spoludávatelé nemovitosti, majitelé společného bankovního účtu. Jsou-li však jednotlivé osoby dále rozlišeny, pak se i v tomto případě jedná o osobní údaj.

Při zpracování údajů o právnických osobách se nepostupuje podle zákona o ochraně osobních údajů, údaje o právnické osobě jsou chráněny podle příslušných ustanovení obchodního zákoníku (např. obchodní jméno-firma, obchodní tajemství), popřípadě občanského zákoníku (§ 18 a související – obecná právní úprava právnických osob).

- 2) Nejedná se o problém pouze anonymizace formulářů konkrétních osob, ale i tzv. bytové formuláře, kde jsou identifikováni členové domácnosti (přesnou adresou a datem narození), přičemž tento formulář obsahuje další údaje (např. jejich příbuzenské vztahy), které se týkají této skupiny osob (osobní údaj se může vztahovat ke skupině osob – viz předposlední odstavce výše uvedeného článku ÚOOÚ).
- 3) ČSÚ byl na popsany problém upozorňován prostřednictvím e-mailu, oficiální telefonní infolinky, webové diskuse na Facebooku, kterou moderuje ČSÚ, ale jakoukoli změnu postupu odmítá nebo připomínku ignoruje.
- 4) Neprovedení anonymizace je závažné narušení bezpečnosti statistické akce Sčítání lidu. Uchováním dat, ze kterých lze i v budoucnosti přiřadit osobní údaje k jednotlivých osobám v mnoha případech s velkou pravděpodobností, vzniká velké riziko zneužití takovýchto dat. A to nejen v současné době – za současného právního systému a platné legislativy, ale i kdykoli v budoucnosti (do úmrtí všech sčítaných lidí).
- 5) ČSÚ informuje v rozporu se svým záměrem občany o tom, že údaje budou anonymizovány tak, že nebude možné přiřadit údaje ke konkrétní osobě. Např. na webu⁹ ČSÚ se píše:

Z výsledků sčítání tak pouze zjistíte, že např. v Bruntále v Jiráskově ulici žije žena, které je 35 let, má dvě děti, každý den dojíždí do práce vlakem do Opavy, vystudovala vysokou školu a narodila se v Ostravě, ale nikdo už nikdy nezjistí, jak se jmenuje, ani jaké má rodné číslo.

V případě provedení postupu, který ČSÚ považuje za anonymizaci, je toto tvrzení nepravdivé a zavádějící. Slovo „pouze“ vyvolává dojem, že nic jiného, než je uvedeno se ze výsledků sčítání nikdo nedozví. Přitom po provedení postupu dojde k zachování čísla domu, označení bytu a spousty dalších údajů. Jméno osoby lze v mnoha případech zjistit pomocí veřejně dostupných informací¹⁰, v některých případech není problém zjistit ani rodné číslo¹¹ (postup již byl popisován výše). Existují i neveřejné rejstříky – evidence obyvatel. Podstatná je možnost spárování údajů z dotazníků s konkrétní osobou a tedy i se všemi dalšími dostupnými informacemi o dané osobě, což se neslučuje s anonymizací.

- 6) Vyjádření ČSÚ k některým otázkám, z nichž jedna se týká anonymizace dat.

*Vážený pane ...,
pokud jde o proces anonymizace dat, obecně platí:*

...

Pokud jde o anonymizované údaje o jednotlivých osobách a jejich vazbu na adresu, tady byste mohl mít částečně pravdu, z těchto dat by se mohla dát v určitých případech (rodinné domy) identifikovat osoba. Tyto údaje však jednak nejsou nikomu poskytovány (to je i odpověď na jednu z Vašich otázek, přístup k neagregovaným údajům mají jen pracovníci našeho úřadu) a jednak v databázích jsou tyto adresy uloženy odděleně od vlastních dat za osoby.

...

⁹ http://www.scitani.cz/cz/media/scitani_lidu_bezpecnost_na_vsech_frontach.html

¹⁰ např. telefonní seznam, katastr nemovitostí, živnostenský rejstřík

¹¹ např. pomocí živnostenského rejstříku

Josef Škrabal

Ředitel odboru statistiky obyvatelstva

Zde jsou tedy nastíněny dva důvody, proč ČSÚ považuje tento stav za vyhovující.

1. **Údaje nejsou nikomu poskytovány.** To je důvod zjevně irelevantní, protože zákon požaduje provedení anonymizace a zničení původních dat. Tedy nestanovuje, že by data v neanonymizované podobě mohl ČSÚ dále držet, pokud je nebude nikomu poskytovat.
2. **Oddělené databáze.** Zaprvé zde ČSÚ nedeklaruje, že tyto databáze nepropojuje např. při vytváření statistických analýz (pokud by tak nečinil, lze pochybovat o užitečnosti takových dat), ani neuvádí konkrétní způsob oddělení. Zadruhé i tak je to irelevantní – viz první tučně vyznačená část zmiňovaného článku ÚOOÚ¹², podle které je třeba posuzovat anonymnost podle všech dat, která má daný subjekt k dispozici (ať je vlastní nebo se např. jedná o veřejně dostupné informace) a to bez ohledu na to, zda má v úmyslu propojování či nikoli. Tedy držení v oddělených datových souborech nemá na posuzování vliv ani v případě, že propojování datových souborů nemá v úmyslu.

Vyhotoveno dne 14. 3. 2011

¹² viz bod 1 doplňujících informací