

Expertní seminář organizace CESSDA: Legální a etický rámec pro užití, znovuužití a archivaci nových typů dat

Viktor Dobrovolný, Michaela Kudrnáčová

Ve dnech 12. až 13. září 2017 proběhl každoroční Expertní seminář CESSDA¹, kterého se účastnil i Český sociálněvědní datový archiv (ČSDA) Sociologického ústavu Akademie věd, pořadatel loňského ročníku, který proběhl v Praze. Letos se seminář uskutečnil v norském Bergenu pod záštitou Norského centra pro výzkumná data (NSD)².

Seminář byl tematicky provázán s dílčími oblastmi projektu SERISS³, a to konkrétně se třemi částmi pracovního balíčku WP6 (6.1 až 6.3) „New forms of data – legal, ethical and quality issues“⁴. První z úkolů WP6 zpracovává zejména ČSDA a týká se legálních a etických problémů souvisejících s užitím velkých dat, se zaměřením na data ze sociálních sítí. Další části WP6 jsou věnovány legálním, etickým a kvalitativním problémům spojeným s užíváním administrativních dat a také otázce kvality obsahu datových souborů.

Zaměření semináře se týkalo legálního a etického rámce pro operace s novými typy dat. Důraz byl položen zejména na výměnu zkušeností s novými typy dat, přičemž nejvíce byla v tomto směru skloňována data ze sociálních médií. Velkým tématem bylo také obecné nařízení o ochraně osobních údajů EU, v originále General Data Protection Regulation (GDPR⁵), které vstoupí v platnost 25. května 2018 a týká se mimo jiné datových archivů. Mezi 23 účastníky, ze 13 institucí, byly zejména odborníci zajišťující management dat v rámci svých archivů.

V úvodní části semináře se rokovalo o definici nových typů dat. Na základě konsensu se došlo k závěru, že nové typy dat není možné redukovat pouze na data ze sociálních médií, ale je potřeba vzít v potaz všechna „big data“. V případě nových typů dat se tedy může jednat o již zmíněná data ze sociálních médií, ale i digitalizovaná starší data (například historické prameny), internetová data, biometrická data, geolokační data, biomarkery, administrativní data, data z registrů, audiovizuální záznamy a další.

Legální a etické problémy jsou spolu částečně svázané, přesto je potřeba se soustředit na obě oblasti jednotlivě. GDPR je novou, v jistém smyslu revoluční legislativou Evropské unie, která představuje právní úpravu ochrany osobních dat na celoevropské úrovni. Pro organizaci CESSDA je tento dokument o to důležitější, že jedním z cílů této organizace je harmonizovat jednotlivá etická doporučení, ale i zákonné normy všech členských států.

Datové archivy v současné době trápí jejich limity. Zvláště se jedná o to, jak přijímaná nová data s potenciálem pro archivaci dále zpracovávat a v jaké formě je dávat k dispozici veřej-

nosti. Nejprve je potřeba zvážit, jak a za jakým účelem jsou nová data v sociálních vědách využívána, tedy obrátit pozornost k odborníkům a výzkumníkům, kteří s těmito typy dat pracují, protože archivy by se měly v první řadě přizpůsobovat nabídce a poptávce. Na základě těchto faktorů musí být data bezpodmínečně shromažďována a popisována podle stejného standardu ve všech členských zemích organizace CESSDA, je tedy potřeba zaměřit pozornost tímto směrem a společnými silami zpracovat směrnici postavenou na konkrétních příkladech, která by sloužila jako stěžejní dokument pro všechny datové archivy členských zemí EU.

V souvislosti s touto problematikou si členové CESSDA kladou několik otázek, a to:

- *Jak vysoký je skutečný zájem o výzkumy postavené na nových datech?*
- *Co by se mělo archivovat, aby se umožnilo užití a znovuužití nových dat a zároveň se šetřilo finančními zdroji?*
- *Jaké jsou standardy kvality pro výzkum nových dat?*
- *Jsou datové archivy vybavené k tomu, pracovat s novými typy dat, nebo je třeba vyvinout nové metody, aby toho bylo dosaženo?*

V rámci jedné z expertních skupin vytvořených za účelem ucelenějšího a srozumitelnějšího výstupu se identifikovala témata s klíčovým významem pro sociálněvědní výzkum:

- Vytvoření přehledu národních harmonogramů implementace GDPR
- Vytvoření přehledu implementace GDPR do národního práva
- Ověření podmínek užití nových typů dat: jaká jsou pravidla a kdo je za ně zodpovědný
- Ověření podmínek, na jejichž základě je možné data shromažďovat, analyzovat a archivovat
- Zpracovat na etických a legálních problémech separátne, tj. vytvořit přehled doporučení pro obě oblasti zvlášť
- Vytvoření harmonogramu pro zakomponování zásad pro „legální, a etický rámec pro užití, znovuužití a archivaci nových typů dat
- Ustanovená pracovní skupina by se měla maximálně soustředit zejména na infrastrukturu (datové archivy) a až následně na vytvoření rad pro výzkumníky

CESSDA tak pokládá základy informační infrastruktury pro nové typy dat, která by jednou mohla svou komplexitou a mírou harmonizace vyvážit podobné iniciativy v USA.

poznámky

¹ Consortium of European Social Science Data Archives.

² Materiály z proběhlých seminářů jsou dostupné na webové adrese <https://tinyurl.com/y8jydp5w>.

³ Synergies for Europe's Research Infrastructures in the Social Sciences. Více informací je k dispozici na webových stránkách projektu <https://seriss.eu/about-seriss/project-overview/>.

⁴ Více informací ohledně WP6 je dostupných na webu <https://tinyurl.com/y8b3ppyu>.

⁵ Podrobnosti jsou dostupné na webu speciálně zřízeném Evropskou unií za účelem o GDPR informovat: <http://www.eugdpr.org/>.